D3 – Report on evaluation of research and legal conditions
Kosovo

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

FNSNF
Swiss National Science Foundation

Deliverable Lead:	CeMI, IDM
Related Work package:	WP1

Author(s):	Lirika Demiri, in cooperation with: Pëllumb Kelmendi Arben Hajrullahu Nitë Bylykbashi-Deliu

Dissemination level:	Public (PU)
Submission date:	1st April 2016
Project Acronym:	SEEDS
Website:	http://www.seedsproject.ch
Call:	Scientific cooperation between Eastern Europe and Switzerland (SCOPES 2013-2016)
Start date of project:	1st May 2015
Duration:	24 months

Version History

Version	Date	Changes	Modified by
1.0	November 13, 2015	Released version	CPC
1.1	November 13, 2015	Revised version	FORS
2.0	February 17, 2016	Released version	CPC
2.1	April 29, 2016	Revised version	FORS
3.0	June 6, 2016	Released version	CPC

Acknowledgments

This report was developed as part of the "South-Eastern European Data Services" (SEEDS) (www.seedsproject.ch) project. The participant organizations of the SEEDS project are:

Name	Short Name	Country
Centre for Monitoring and Research, Podgorica	CeMI	Montenegro
Centre for Political Courage, Pristina	CPC	Kosovo
Institute for Democracy and Mediation, Tirana	IDM	Albania
Institute of Economic Sciences, Belgrade	IES	Serbia
Saints Cyril and Methodius University, Institute for Sociological, Political and Juridical Research, Skopje	ISPJR	Macedonia
Swiss Foundation for Research in Social Sciences, Lausanne	FORS	Switzerland
University of Ljubljana, Social Science Data Archive, Ljubljana	UL	Slovenia

Table of Contents

1. Introduction and methodology overview 4
1.1 The questionnaire for researchers 4
1.2 The research infrastructure questionnaire 5
1.3 The interview guide for research policy and funding institutions 5
Part I: Policy and legal framework 6
1. Basic features of the science system in Kosovo 6
1.1 Research units and key drivers of scientific research in the social sciences 6
1.2 Funding of research activity 7
2. Existing policy documents relevant to scientific research activity 7
3. Existing e-infrastructure in the public institutions 10
4. Legal framework 10
5. Interviews with research policy and funding institutions 12
Part II: Survey on production, preservation and use of research data among researchers . 15
1. Survey 15
1.1 Survey Participants 15
1.2. Production of data 18
1.4 Access to Data 25
1.5 Data Usage and Secondary Analysis 27
1.6 Data analysis 31
1.7 Perceived need for data archive 32
Part III: Survey on existing infrastructure and interest in data archiving 35
Annexes 40
Annex 1. List of interviews held 40
Annex 2. Questionnaires 41

1. Introduction and methodology overview

This report is focused on mapping out the current environment and opportunities related to the establishment of social science data services in Kosovo. The analysis of the current archival infrastructure, as well as of the policies related to data access, is based on findings from surveys conducted with three key groups of stakeholders that are directly involved and play an important role in establishing, using and preserving data archives. These three identified groups involve:

- Social science researchers;
- Representatives of research/data service institutions;
- Representatives of research policy and funding institutions.

The researchers are seen as potential contributors of data, as well as active users of the archives. The research infrastructure institutions are considered to be the main providers of professional and technical services regarding data collection, whereas the research policy and funding institutions are targeted as possible supporters and funders of establishing a data service.

Since each of the stakeholders constitutes a distinctive pillar of data archiving infrastructure, they were approached with three different survey instruments. With respect to survey instrument design, the South-Eastern European Data Services project (SEEDS) has gained from the experience of a similar set of surveys carried out previously by the SERSCIDA project.[1] Below, each survey instrument is briefly described and the main aspects of the data collection process are presented.

1.1 The questionnaire for researchers

The questionnaire for researchers is primarily focused on the current research practices and needs related to collecting empirical data, their preservation habits and use for secondary analysis. The survey was prepared using the LimeSurvey software. Social science researchers in Kosovo were contacted between July and August 2015 via e-mail and asked to fill in the questionnaire. The survey questions mostly refer to data preservation and dissemination practices, policies regarding data access within research institutions, as well as methodologies and software packages used for data collection and analysis. The researchers were also asked to provide their personal opinion regarding the importance of data sharing and to identify the barriers to conducting secondary analysis in Kosovo.

[1] For more information about SERSCIDA project: http://www.serscida.eu/en/

1.2 The research infrastructure questionnaire

Semi-structured interviews were conducted with the representatives of social science research institutions in Kosovo. All of the interviews were conducted face to face between August and November 2015. The contacted institutions included the Kosovo Agency of Statistics and other public research institutes. During the interviews, special emphasis was placed on long–term data preservation practices, standards used for documentation, existing technical capacities regarding data archiving inside these institutions, human resource capacities, as well as dissemination and funding policies. As the potential hosts of a data service such as the one the SEEDS project aims to establish, the representatives of research institutions were also asked to provide their opinions and recommendations on a possible national social science data service.

1.3 The interview guide for research policy and funding institutions

As the main responsible actor regarding research policies and funding, the Ministry of Education, Science and Technology (MEST), respectively the representative of the Department for Science and Technology was contacted directly. The discussion was primarily concentrated on the current policies and activities regarding long–term data preservation and initiatives for open access, as well as on existing infrastructure and financial obstacles and potentials for establishing a national data archiving service.

Part I: Policy and legal framework

1. Basic features of the science system in Kosovo

The MEST is the main institution responsible for developing, coordinating, financing and implementing national policies related to scientific-research activity in Kosovo. Two departments within this ministry are primarily in charge of regulating and organising research activity: the Department of Higher Education and the Department of Science and Technology. The main role of the Department of Science and Technology is to execute government policies that aim to create a sound institutional and financial infrastructure for the development of science and scientific research, as well as to encourage the development and application of modern technology in the Kosovo economy.[2]

Another important institution is the National Science Council established by the Assembly of Kosovo in 2007. As defined by the Law on Scientific – Research Activities, the National Science Council "shall attend to systematic development of scientific-research and technological activity in the Republic of Kosovo."[3] Furthermore, the Council is obliged to "assess, in comparative term, condition in scientific-research and technological activity, position and its development in the national and international degree; to propose and stimulate taking of measures for the progress of scientific -research activity."[4] The National Science Council has compiled the Kosovo National Research Programme.

It is also worth mentioning the Kosovo Accreditation Agency (KAA), which is the main responsible body for accrediting higher education institutions and ensuring the quality of existing programmes.

1.1 Research units and key drivers of scientific research in the social sciences

Kosovo has public universities in all seven of its largest cities, as well as 29 accredited private institutions of higher education and several public institutes with a research component.[5] The University of Prishtina is the main public higher education institution and much of the research conducted in the social sciences takes place there. However, as it is also shown in the survey results below, no organised institutional support is given to research projects, and most of the research activity is realised as a result of international funding or individual initiatives. The civil society sector and think tank organisations, in

[2] See, http://masht.rks-gov.net/departamenti-per-shkence-dhe-teknologji (accessed on 29 September 2015).
[3] Law No. 04/L-135 On Scientific Research Activity, Article 41, available at: http://www.kuvendikosoves.org/common/docs/ligjet/Law%20on%20Scientific%20Research%20Activities.pdf, (accessed on 29 September 2015).
[4] Ibid. Article 41.
[5] Kosovo Accreditation Agency, http://www.akreditimi-ks.org/new/index.php/en/download/higher-eduacion-institutions (accessed on 5 December 2015).

particular, have enriched the sphere of scientificresearch in the social sciences with numerous internationally funded research projects. In this respect, a large amount of data is produced and preserved in the NGO sector. The public institutes that are under the supervision of the Ministry of Education are also important actors engaging in research activity. Some of the public institutes include the Institute of Albanology, the Institute of History "Ali Hadri", and the Education Institute of Kosovo.

1.2 Funding of research activity

Public funding for research activity is defined by the 2013 Law on Scientific Research Activities 04-L/135, Art. 5, which stipulates that "[f]or providing the conditions and necessary means for scientific research activity, the Government of the Republic of Kosovo allocates zero point seven percent (0.7 %) of the local annual budget."[6] However, in Kosovo, only approximately 0.1% of Gross Domestic Product (GDP) is allocated to research and development, which represents the smallest level of expenditure for research activity in the region of South Eastern Europe.[7]

The Kosovo National Programme had foreseen a budget of €1 million in 2010, with a gradual increase reaching a peak of €6 million in 2015. However, the amount of € 580,000 allocated to implement the National Programme until 2014 is not even close to the necessary amount of the € 5,393,953 planned for financing 22 activities related to research, science, infrastructure, mobility, internationalisation of research, as well as establishing quality assurance mechanisms.

MEST has also established the National Research Fund in order to support different research projects with grants. However, until now there was no transparency and accountability in the process of fund allocation, and its financial sustainability is also uncertain.[8]

2. Existing policy documents relevant to scientific research activity

An important document outlining Kosovo's policy framework regarding scientific research **is the Kosovo National Research Programme.** This policy document, approved by the National Research Council in 2010, aims to identify research priorities, strengthen Kosovo's participation in international research programmes, and invest in relevant infrastructure. The National Programme outlines a clear picture of

[6] Available at: http://www.kuvendikosoves.org/common/docs/ligjet/Law%20on%20Scientific%20Research%20Activities.pdf (accessed on 28 September 2015).

[7] Kosovo National Research Programme of the Republic of Kosovo, January 2010, p.5.

[8] "The State of Implementation of the National Research Programme of the Republic of Kosovo", December 2014. Available at: http://www.cpc-ks.org/repository/docs/Briefing_Paper_National_Science_Program_356224.pdf (accessed on 18 December 2015).

the current situation of research and science in Kosovo and sets up explicit objectives and priorities supported by a detailed action plan and specified budget planning in a timeframe of five years (2010-2015). The five objectives formulated in this Programme, intended to be met by 2020, are as below:

- Development of human capacity for research activities;
- Development of research infrastructure;
- Internationalisation of scientific research activities;
- Strengthening of the links between science, society and economy for enhancing economic and social development;
- Excellence in research and scientific activity.[9]

Nevertheless, a recent briefing paper entitled "The state of implementation of the National Research Programme of the Republic of Kosovo" published by the Center for Political Courage (CPC) finds that none of the activities planned within the Research Programme has been fully implemented.[10]

Another noteworthy policy document is the **Strategy on Scientific/Artistic Research and Development Activities 2013-2016**, which has been developed by the University of Prishtina, in line with the objectives specified by the National Research Programme. The strategy identifies four main areas of scientific research in the University of Prishtina for Kosovo's public policy to focus on: human resources, research infrastructure, international cooperation, and the relationship between economy and society.

1. **The human resources area** lists 6 priorities: the development of doctoral programmes, the provision of funds for short-term visits abroad, incentives for staff research productivity, incorporation of diaspora's capacities in research activities, and regulation of research activities of UP staff through employment contracts.

2. **The infrastructure area** lists 5 priorities, including: establishing a database of infrastructural capacities of UP, establishing laboratories, establishing institutes, and providing access to electronic libraries.

3. **The internationalisation area** lists 4 priorities including: strategic partnerships between the UP and international institutions, technical assistance for writing project proposals, the provision of vocational programmes in compliance with labour market demands, and UP co-financing in international projects.

4. **The science and economy area** lists 4 priorities including: establishing a database of UP co-operation projects with the public and private sectors, providing vocational programmes,

[9] National Research Programme of the Republic of Kosovo, January 2010, p.16.
[10] "The State of Implementation of the National Research Programme of the Republic of Kosovo", December 2014. Available at: http://www.cpc-ks.org/repository/docs/Briefing_Paper_National_Science_Program_356224.pdf (accessed on 28 September 2015).

incorporating the business community in the process of drafting the academic programmes of UP and engaging the private sector in joint scientific research.[11]

Another important strategic document regarding scientific research in Kosovo is the **Strategy for Development of Higher Education in Kosovo 2005-2015,** compiled by MEST in 2004. In addition to the other objectives that aim to create inclusive and contemporary education policies as well as effective management of the higher education system, this strategy also intends to advance the capacity for research and scientific work. In this aspect, the strategy identifies the major deficiencies regarding the scientific research area, listed below:

- Lack of national policies and programs for scientific research;
- Incomplete legislation for scientific research;
- Scientific research does not serve solving societal problems;
- Lack of innovation and their evaluation;
- Lack of international cooperation in research;
- Lack of interdisciplinary approaches in the scientific research work;
- Lack of up-to-date information on current scientific results;
- Post-graduate studies are not coordinated with research priorities;
- Lack of standards for research work;
- Limited research experience;
- Lack of needs assessment for scientific research;
- Lack of institutional infrastructure for research work;
- Lack of incentives for faculty involved in scientific research;
- Low capacity for research work;
- Lack of public funds and other resources dedicated to scientific research;
- Lack of scientific projects supported by public funds and other resources;
- Lack of mechanisms for protection of intellectual property and industrial rights.[12]

[11] Strategy on Research/Artistic and Development Activities 2013-2016, p. 8. See: http://www.uni-pr.edu/getattachment/4fc95f03-b53f-4e7c-a2a0-c0056b9c43e4/Strategjia-per-veprimtari-kerkimore-shkencore-arti.aspx (accessed on 28 September 2015).

[12] The Strategy for Development of Higher Education in Kosovo 2005-2015, MEST, available at http://www.kryeministri-ks.net/repository/docs/Strategjia_e_Zhvillimit_te_Arsimit_te_Larte_ne_Kosove_2005-2015.pdf, p.13 (accessed on 28 September 2015).

3. Existing e-infrastructure in the public institutions

As part of the endeavours to initiate e–governance and to use updated, innovative information and communication technology, the Ministry of Public Administration established the Agency for Information Society in 2013. In general, the Agency "shall support the development of information technology, promote investment in the field of information society, development of training systems in information technology and shall do the coordination, management and supervision of processes and mechanisms of electronic governance".[13] Although the Agency does not have a scientific research character, it reflects the first institutional attempt to create innovative structures of information technology and to support the idea of opening access to administrative data.

4. Legal framework

The Law on Scientific Research Activities (Official Gazette, no. 04/L-135) regulates, defines and specifies all scientific research activities, rights, and obligations of scientists and researchers, as well as the financing sources of scientific research activity. In general, the law places great emphasis on creating and promoting an innovative and competitive research environment. The development of advanced research technologies and infrastructures, enhancement of cooperation between different stakeholders regarding scientific research activities, as well as the promotion of an effective dissemination of scientific results are seen as essential preconditions for an overall economic and social prosperity.

Article 5 of this law defines the funding of scientific research activities. Specifically, section 1.4 states, among others, that the financing means provided by the government of the Republic of Kosovo include "activities in systematical preservation and distribution of activities outcomes of basic research and those applied, such as scientific conferences, exhibitions, publications and other forms of documentation".

The Law on State Archives (Official Gazette, no. 04/L-088) regulates the organisation and operation of archival services in Kosovo's institutions. This law specifically defines archiving procedures and conditions of archived material and documentary material that are produced in public institutions. According to section 1.2 of article 3 of this law, archived material is defined as "all original and reproduced material by public institutions, legal and natural entities in their activities such as: manuscripts, printed records, drawings, photographs, stamps, audiovisual materials, electronic materials

[13] Law No. 04/L-145 On Information Society Government Bodies Article 6, available at: http://www.kuvendikosoves.org/common/docs/ligjet/Law%20on%20information%20society%20government%20bodies.pdf (accessed on 29 September).

and all other supporting tools that enable a clear understanding of the content of information and facilitate the use of them".[14]

This law also discusses the organisation of the State Archive Agency, as the main body responsible for receiving, identifying, maintaining, and publishing the archival material. In this sense, a great importance is given to a free, open and informed access of the interested general public to the material possessed by the Agency. Article 8 specifies that the Agency, "in order to improve opportunities for the use of archive material shall create, maintain and update the public archive registry". Furthermore, section 6 of this article states that the public archive registry should be managed by using advanced technology and should be available on the website of the Agency. Article 12 of the law, moreover, guarantees the right to access after the request is archived in the Agency without discrimination on any basis.

The Law on Official Statistics (Official Gazette, no. 04/L-036) is one of the most important laws regarding statistical data production, storing and dissemination in the Republic of Kosovo. It mainly defines the core principles of official statistics, the organisation of the official statistics system, data collection, processing, and storing of statistical data, as well as the dissemination and use of official statistics. According to article 4, section 1.5 of this law, one of the criteria that guarantees the quality of official statistics is that of accessibility and clarity, defined as "the conditions and modalities by which users can obtain, use and interpret data".

The law also specifies the responsibilities and tasks of the Kosovo Agency of Statistics (KAS) as the main coordinator of the statistical system in Kosovo. In this sense, among other responsibilities, KAS has to prepare the Programme of Official Statistics, which shall provide a detailed framework regarding data production and dissemination. With respect to statistical data storing, Article 26, Section 3 of this law states that "statistical material shall be stored in such a way as to prevent its destruction, misuse, alienation, misinterpretation, and its disclosure".[15] The law also obliges KAS and other authorities to disseminate all the official statistics in such a way that all users have the necessary support for ensuring quality on equal access.

The Law on Copyright and Related Rights (Official Gazette, no. 04/L-065) is also relevant to the aims and objectives of the SEEDS project. In general, the law regulates copyright that is based on intellectual property, which belongs to authors with respect to their works in the literary, scientific, and artistic domains. Among others, the law mentions the rights of database producers.

Regarding the right of distribution, Article 24 of this law states that it "is an exclusive right of the author to authorize or to prohibit the putting into circulation of the original or copies of his work, by sale or any other form of transfer of ownership, including their offering to the public for such purpose". In Article 33, on the other hand, the right of making the work available to the public is defined as "an exclusive

[14] Law No. 04/L-088 On State Archives, available at: http://www.kuvendikosoves.org/common/docs/ligjet/Law%20on%20State%20Archives.pdf (accessed on 5 December 2015).
[15] Law No. 04/L-036 On official statistics of Republic of Kosovo, available at: http://www.kuvendikosoves.org/common/docs/ligjet/Law%20on%20officiial%20statistics.pdf (accessed on 5 December 2015).

right of the author to authorise or prohibit that his work is made available to public through linear and nonlinear communication, in a way which enables access to individuals from place and time they choose".

The Law on Access to Public Documents (Official Gazette, no. 03/L –215) guarantees the right of access to official documents maintained, drawn or received by public institutions to every natural and legal person without discrimination on any grounds. Furthermore, it obliges all public institutions to assign units or officers who will be responsible for receiving and conducting an initial review of applications for access to documents. According to this law, public institutions are also obliged to publish in electronic form and through the publication in the Official Gazette of Republic of Kosovo all that can be public. Any applicant of a document has the right of access without the obligation to specify the reasons.

The Law on the Protection of Personal Data (Official Gazette, no. 03/L – 172) determines the rights, responsibilities, principles and measures with respect to the protection of personal data that have been processed by public and private bodies. The law regulates the legitimacy of data processing, rights of the data subject, as well as the transfer of personal data.

Specifically, Article 9 of this law defines the processing of personal data for historical, statistical, and scientific-research purposes. It states that personal data may be further processed for historical, statistical, and scientific research purposes regardless of the initial purpose of collection. In this case, personal data should be anonymised unless otherwise provided by law, or if the data subject has given his or her prior written consent. Also, the publication of the results of processing should be in anonymised form, unless the data subject has given his or her prior written consent for the publication in a non-anonymised form, or unless the prior written consent for such publication has been given by the legal heirs of the deceased individual.

The Law on Information Society Government Bodies (Official Gazette, no. 04/L-145) aims to determine the relevant institutions as well as their functions and responsibilities in the development and implementation of information technology in the Republic of Kosovo. Regarding the functions of the Agency for Information Society, Article 6 of this law, among others, states that this Agency is responsible for "accumulation, administration, dissemination and storage of data by creating the State Data Electronic Centre, security and protection of electronic communication infrastructure and data, facilitation of the access to public information in electronic form".

5. Interviews with research policy and funding institutions

In order to gain a better understanding of the needs, current policies, possible institutional solutions and financing means regarding the establishment of a national data archive, we conducted interviews with the main institutions responsible for research policy development and for financial support of scientific innovation. The interviews were conducted in the second half of October 2015. The institutions contacted were the National Scientific Council, MEST and its Scientific Council. These institutions were

contacted via email. They were introduced to the overall objectives of the SEEDS project and were invited to participate in an interview to further discuss the role that each of them could play on developing institutional solutions for the establishment of a functional and enduring national data archive of social sciences. While two interviews were conducted with the representatives of MEST, we were unable to meet with the representatives of the National Scientific Council and Scientific Council of MEST since they did not respond to our repeated invitations. We present below the main results of the discussions we had with the representatives of MEST.

A meeting with the director of the Department of Science and Technology in MEST was held with the purpose of discussing the current policies and activities regarding data preservation and archiving, as well as open data access initiatives. Possibilities for future funding of a national social sciences data archive were also conferred.

The director presented two major projects related to open access that are still in the ideation phase due to lack of inter-institutional coordination and financial constraints. One of the projects, ideated in cooperation with the private institution of higher education "University of Business and Technology", has to do with the designing of an online platform where scientific researchers could present and share their research work and scientific publications, while the other one aims to establish an e-library that could be openly accessed by the general public. Although none of these projects is directly linked with primary data archiving, we consider them to be important indicators of an institutional and cultural shift toward a more open scientific and academic environment.

Currently, the MEST does not operate with any specific policy regarding research data archiving. The funding of scientific research projects and activities is not accompanied with requirements for ensuring data open access and sharing. In this respect, the idea of the establishment of a national data archive was welcomed and considered as potentially very beneficial for the scientific community in Kosovo. The representative of the MEST expressed willingness to play an active role in supporting the process of setting up the archive. However, a stable and unceasing financing by the MEST was deemed to be very problematic and difficult due to recent budget cuts.

The director of the Department of Science and Technology in MEST was not able to provide clear and specific answers for many of the questions related with archiving policies. Currently, the MEST does not have any concrete plans regarding research data archiving, and the situation is justified with the lack of financial and human resources. When asked about the establishment of an institution for data archiving, the director emphasised that any of the possible solutions should come as a result of a broader debate among all institutions and stakeholders involved. He considered that MEST, and the Department of Science and Technology in particular, should have a leading role in all this process. He also indicated that alternative sources of funding need to be sought because of the very limited budged of the MEST for innovative project, such as SEEDS is.

An interview was also conducted with a representative of the Division for Science in MEST. Mainly, the application procedures for small scientific grants were discussed. Despite the fact that the SEEDS project does not fit completely into the categories of the research projects that are financially supported, due to

its specific innovative character we were encouraged to apply for this grant scheme. The professional decision-making body for grant allocation is the Scientific Council of MEST.

As was stated above, we were unable to meet with any members of this Council in order to discuss funding possibilities and to take their recommendations for a future functional model of the data archive.

Part II: Survey on production, preservation and use of research data among researchers

1. Survey

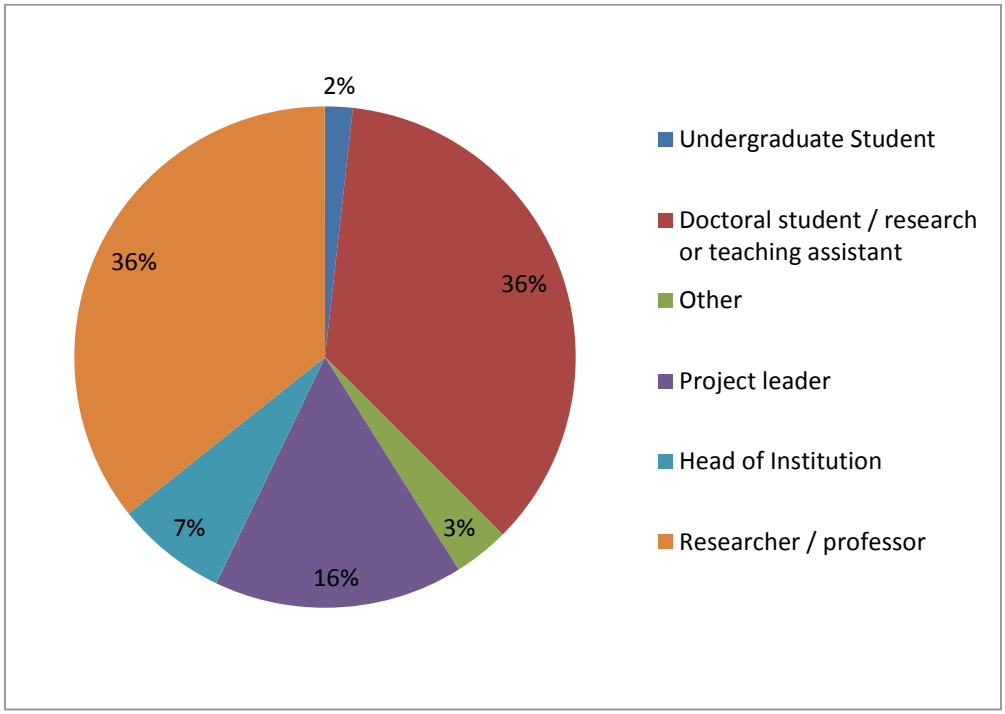
We e-mailed the invitation to participate in our survey in the second half of July 2015. The invitation was e-mailed to around 100 different units such as universities, departments, and institutes. We also targeted specifically 603 different individuals, whose contact information we retrieved from institutional websites. Our goal was to reach out to researchers in the fields of anthropology, economics, education, political science, psychology, sociology, and other related social science disciplines. In addition, we invited participants to forward the invitation to participate to their colleagues. The invitation to participate in the survey was subsequently resent two additional times. Because access to the survey was open and because we encouraged the invitees to forward our invitation to their colleagues, we cannot know with any degree of confidence what the participation rate was or how representative of the community of social science researchers in Kosovo the sample of participants that we eventually yielded is.

It is important to note, however, that we believe that we have reached a significant portion of the social science research community in Kosovo. While we do not have data on the universe of individuals conducting social science research in Kosovo, we do know, for example, from the last census conducted in 2011 that only 0.09 percent of Kosovo's total population above the age of 10, or 1336 individuals hold a PhD degree, which we take as a rough approximation of the research community in Kosovo. Given that many of the PhD holders are in non-social science disciplines, it is reasonable to believe that the number of social science researchers in Kosovo holding a PhD degree is much smaller than that. On the other hand, we also know that many researchers operating in Kosovo do not hold a PhD. That said, we believe that the number of researchers that we contacted in our survey constitutes a significant proportion of the entire population of Kosovo's social science researchers.

1.1 Survey Participants

Of the total of 117 respondents who accessed the Kosovo section of the questionnaire, only 57 of them ended up completing it. Data collected from these respondents, incomplete questionnaires, is excluded from the analysis. While we cannot know for sure why the number of respondents is so low, we believe that the low number, at least in part, reflects the low number of individuals conducting data-driven research in Kosovo. It is important to note, for example, that many university professors in Kosovo focus their activities on teaching and stop conducting research early on in their careers.

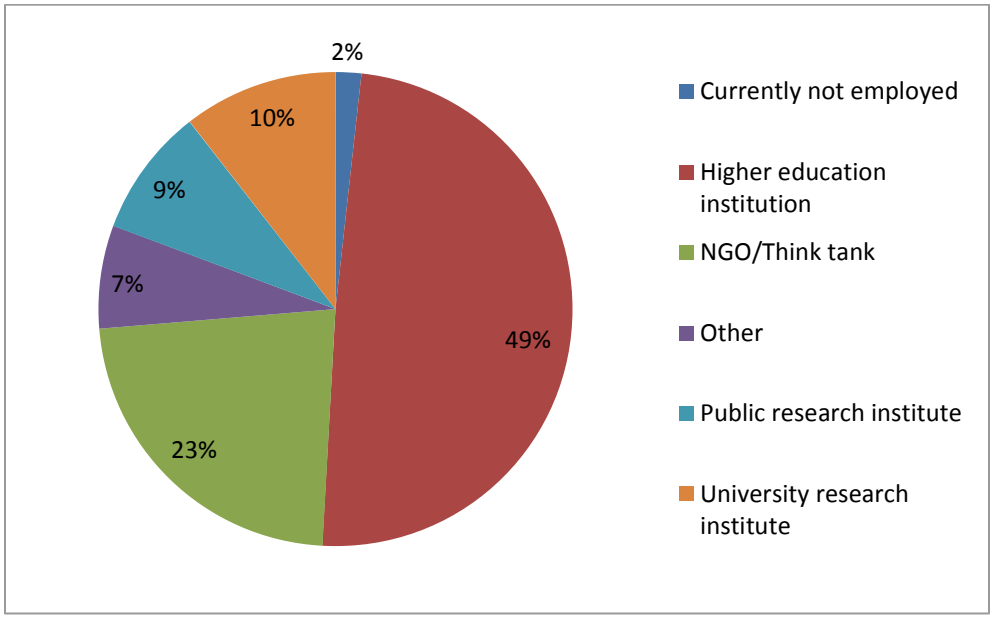
Figure 1: Distribution of respondents according to their principal activity



N=57.

As Figure 1 above shows, the survey participants included 20 doctoral students or research/teaching assistants and an equal number of respondents who declared that they were researchers or professors. 9 participants declared that they were project leaders whereas 4 others stated that they were heads of an institution. Among the two survey participants who selected the category of other, one stated that they were a researcher only and another that they were a monitoring and evaluation manager. Finally, one of the participants was an undergraduate student.

Figure 2: Distribution of respondents according to their institutional affiliation

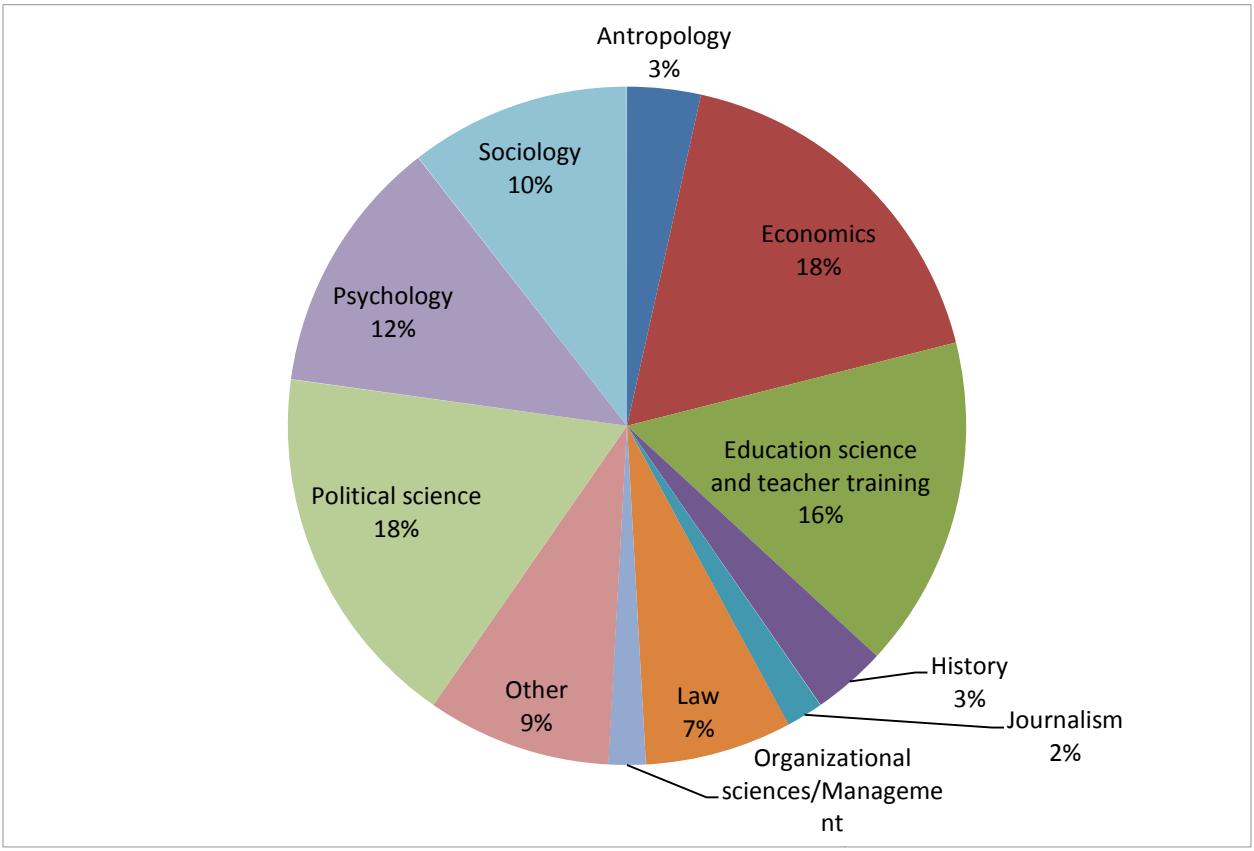


N=57

As expected, almost half of the survey participants were affiliated with a higher education institution. A significant portion, however, came from the NGO or think tank world, which is in line with our broader observations that think tanks and NGOs constitute a significant portion of social science research activity in Kosovo. One participant declared that they were unemployed, whereas the rest were divided roughly equally between university research institutes, public research institutes, or some other type of institution. 2 out of the 3 respondents who declared other as their institutional affiliation further specified that they belonged to a private research agency.

As can be seen from Figure 3 below, the largest portion of respondents consider political science or economics to be their primary discipline (18% each), followed by education science and teacher training (16%), psychology (12%) law (7%), anthropology (3%), organisational sciences and management (2%), history (2%), and journalism (2%).

Figure 3: Distribution of respondents according to their primary discipline



N=57

Three of the respondents declared that their primary field was not in social sciences (rather in molecular biology, hydromechanics, and computer science) and were subsequently dropped from the analysis.

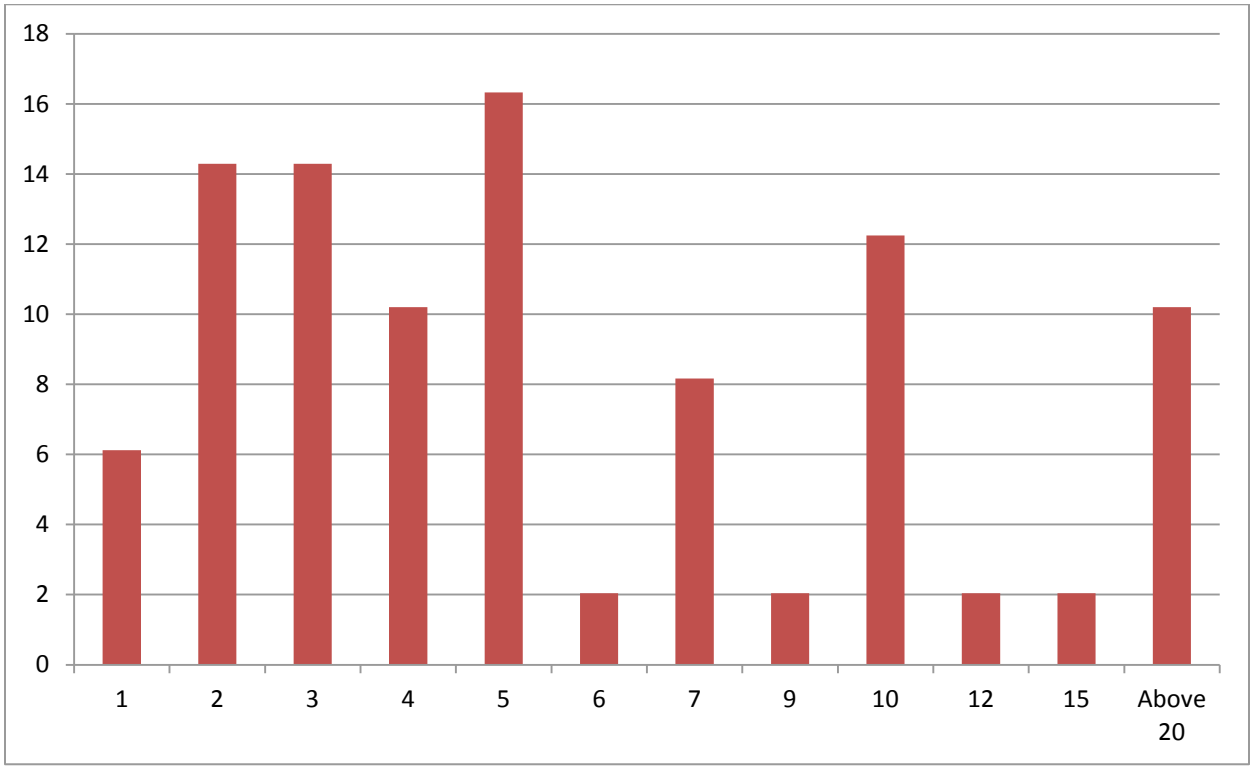
1.2. Production of data

Of the survey respondents who conduct social science research, only 6 out of 54 stated that they did not produce or help produce any research data. In other words, 88 percent of the eligible respondents stated that they had been involved in producing data in the last 5 years. This statistic most likely over-estimates data production among Kosovo's social scientists, since we suspect that there was significant selection bias in survey participation, with those who engage in data gathering being much more interested in participating.

As figure 4 below shows, the number of datasets produced differs substantially, from 1 dataset produced by 6 percent of respondents to above 20 produced by 10 percent of them. Analysing the data

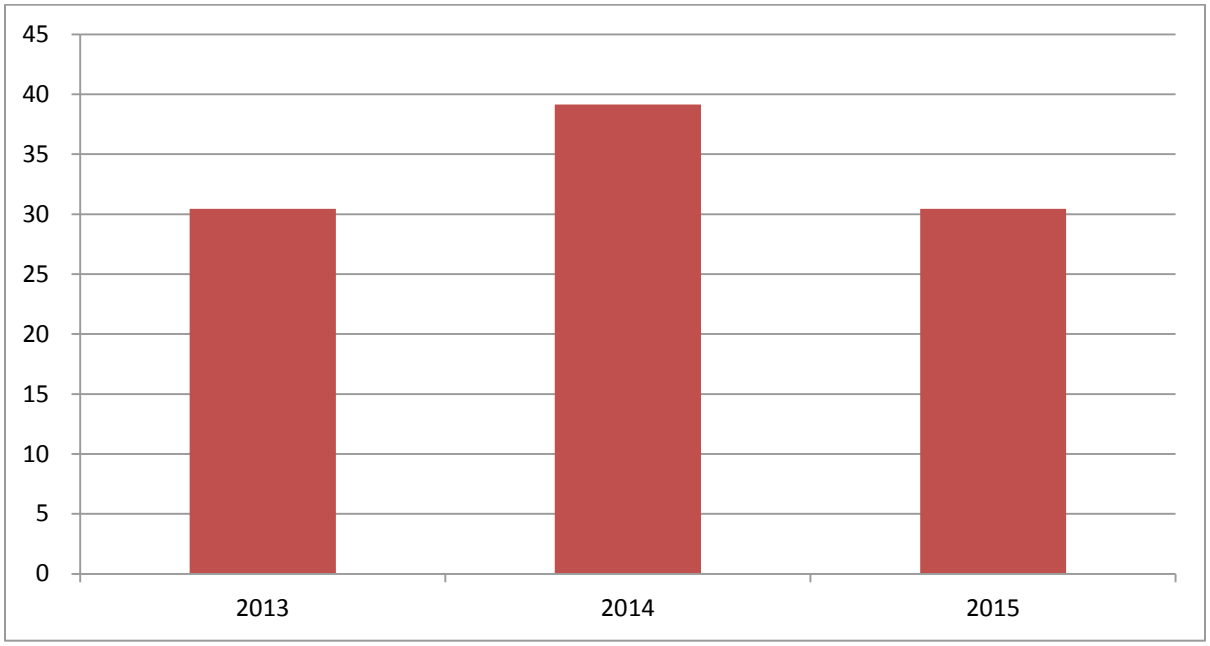
further did not reveal any pattern between data production and primary discipline. Those who indicated participation in more than 20 research projects included an anthropologist, a political scientist, an economist, a sociologist, and a historian. Likewise, those who indicated participating in 1 to 3 research projects in the last 5 years were from the fields of political science, anthropology, education science, law, and economics.

Figure 4: Percentage distribution of researchers according to the number of datasets they were involved in producing



N=49

Figure 5: Year most recent data collection completed (in percent)

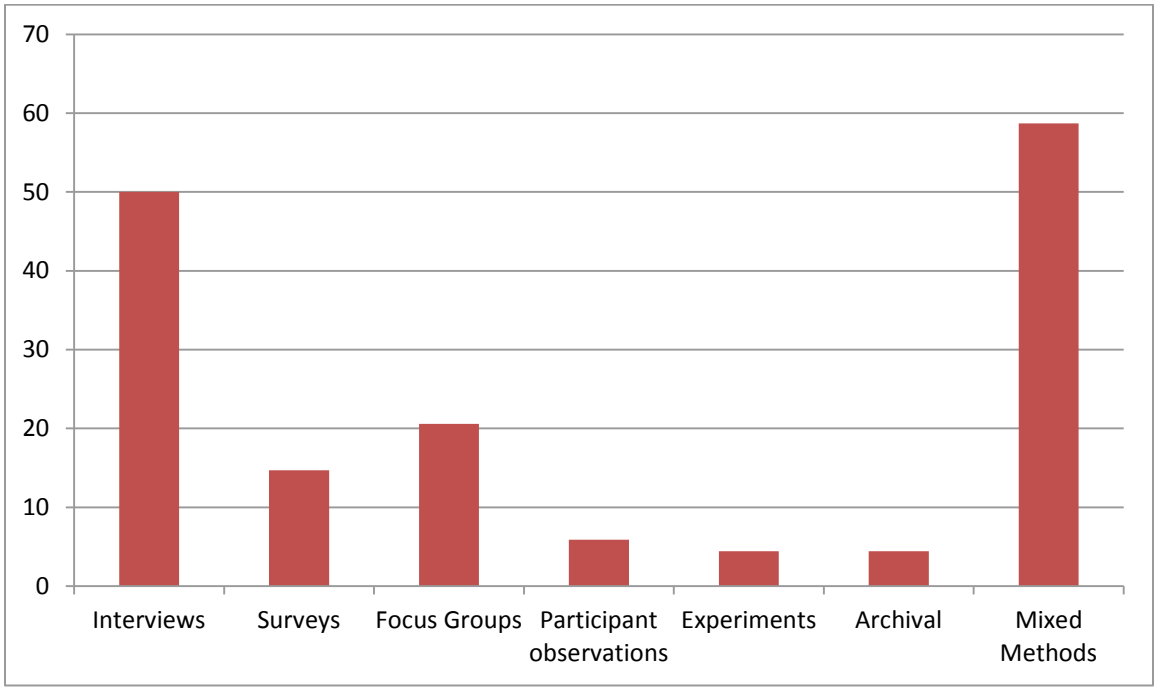


N=46

Among participants who completed fieldwork or data collection in the last 5 years, about 30 percent participated in data collection between January and August 2015, another 39 percent declared that the most recent year in which they participated in data collection was 2014, whereas around 30 percent participated in data collection on or before 2013.

Figure 6 below shows the methods that the survey respondents used in their last research project. 58 percent of the survey respondents reported using more than one data collection method. Among the specific data collection methods used, interviews (structured, semi-structured, or elite interviews) were by far the most common method of choice, followed by focus groups (used by 20 percent of the respondents). Another frequently used method was surveys - 15 percent of the respondents mentioned using survey methods, including large-scale surveys, online surveys, or on-the-spot surveys. Other data collection methods – participant observations, experiments, or archival data collections – were employed by about 5 percent of the survey respondents. Given that the majority of survey respondents indicated using qualitative data gathering methods, it is evident that for Kosovo qualitative social science data archives are particularly necessary.

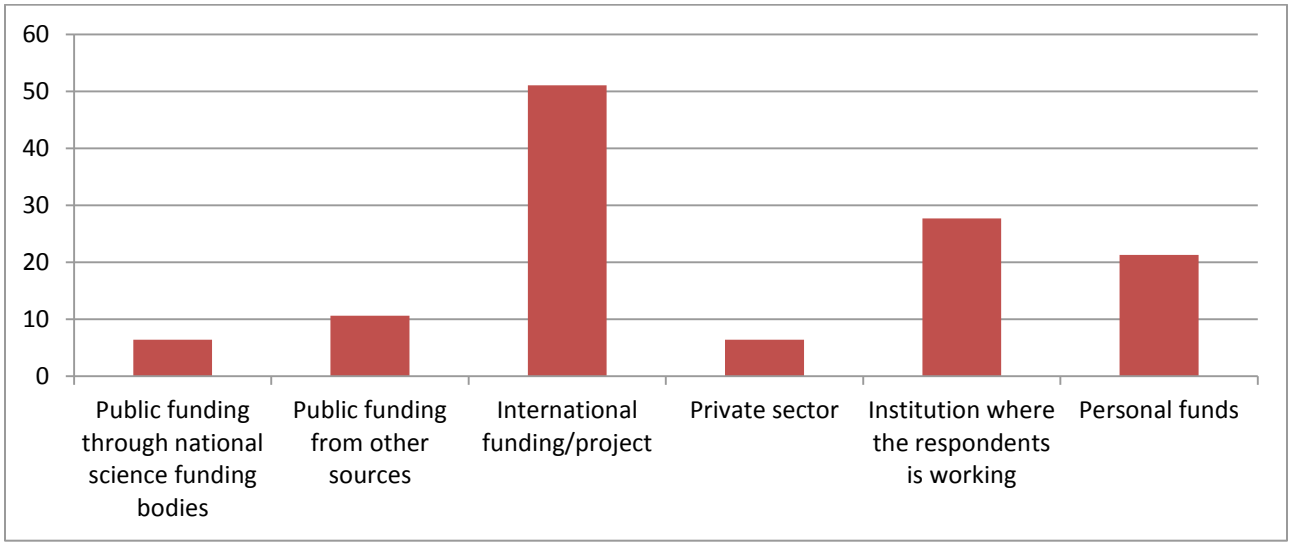
Figure 6: Methods used in the last research (percent)



N=46

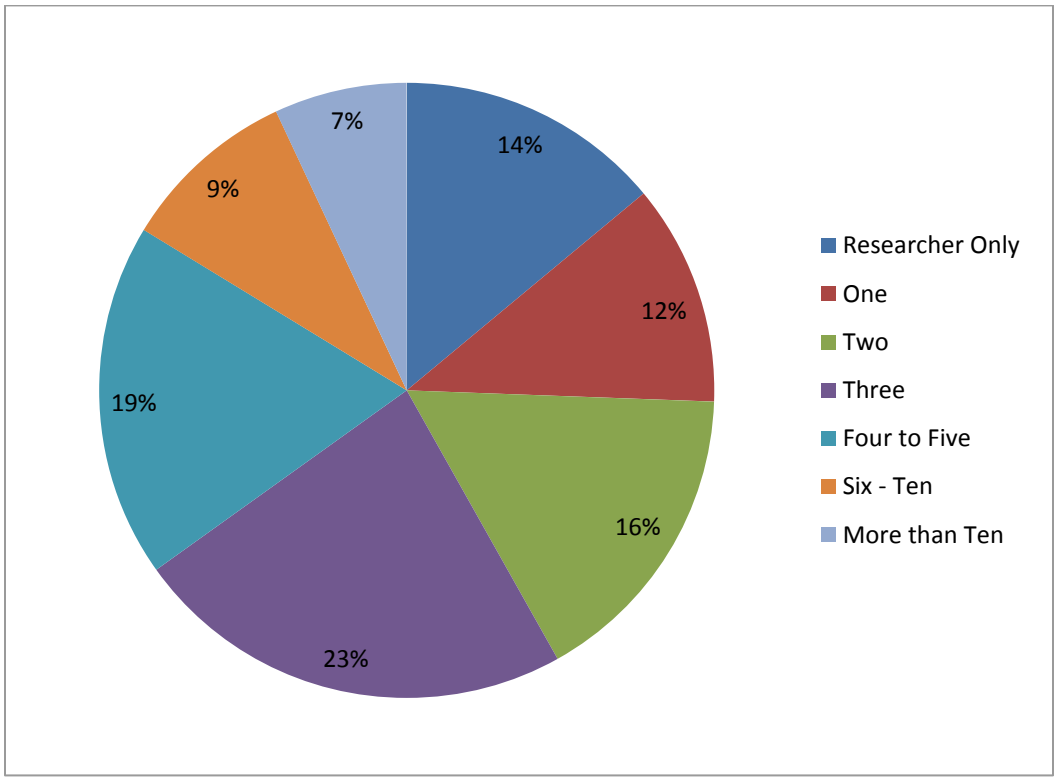
As far as funding sources for the last research project are concerned, figure 7 shows that a mere 6 percent of the respondents declared that they received public funding through national science funding bodies and another 11 percent of the respondents declared that they received public funding from other sources (such as ministries or national or local agencies). The majority, more than 50 percent, used international funds instead. The second most common source, mentioned by 27 percent of respondents, was the institution where the respondent is working. The third most common source, mentioned by 21 percent of respondents was personal funds.

Figure 7: Funding sources for last research project



N=47

Figure 8: Number of researchers involved in the most recent project



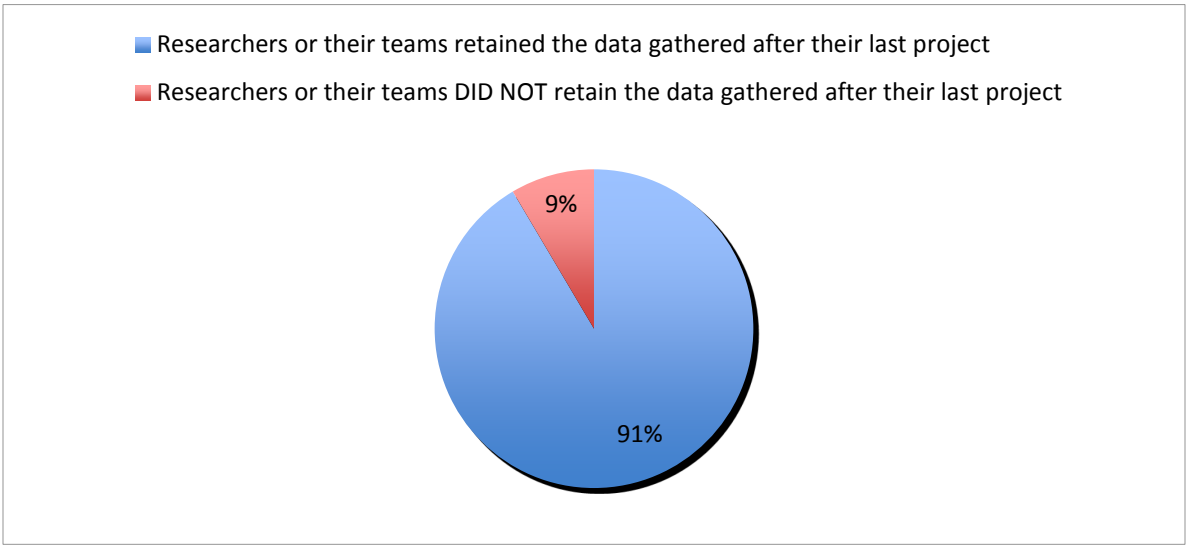
N=43

According to our survey results, the size of research teams involved in the last data gathering project varied substantially (figure 8). Whereas 14 percent of participants indicated that they were the only ones involved, 23 percent indicated that they had three other collaborators, and 7 percent indicated more than 10 researchers on their team. Nevertheless, most of the respondents indicated small teams, with more than 65 percent of respondents participating in teams of four or fewer.

1.3 Archiving practices and preferences

While much of the data is retained, it is kept primarily by the researchers themselves, and our survey results on data archiving confirm our preliminary conclusion that this is not done systematically. To begin with, 4 respondents (9 percent of the total) indicated that they or their research team did not keep/retain the data after completing their last project (figure 9). This is a surprisingly high number and indicates that perhaps more should be done in educating the community of social science scholars in Kosovo about the importance of retaining data. The necessity of educating the social science scholarly community about the importance of data preservation and dissemination was also highlighted by our survey participants in the question that invited them to leave any additional comments. For example, one of the researchers indicated that, "one should first raise the awareness among the community regarding the importance of data archiving and train the community in data archiving and research data sharing."

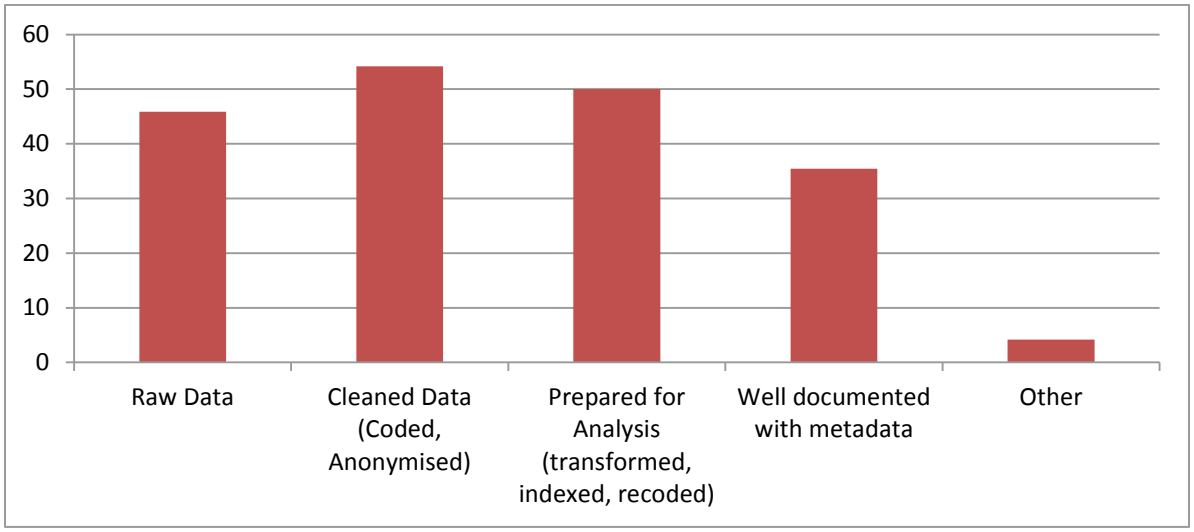
Figure 9: Distribution of researchers who retained or did not retain their data after their last data-gathering project.



N=54

Of those who did retain the data, on the other hand, there was no dominant pattern in archiving. Only 35 percent indicated that their data were kept well documented with metadata, but none of those used international standards for documenting and 13 respondents said that they had used an internal/institutional documentation standard (figure 10).

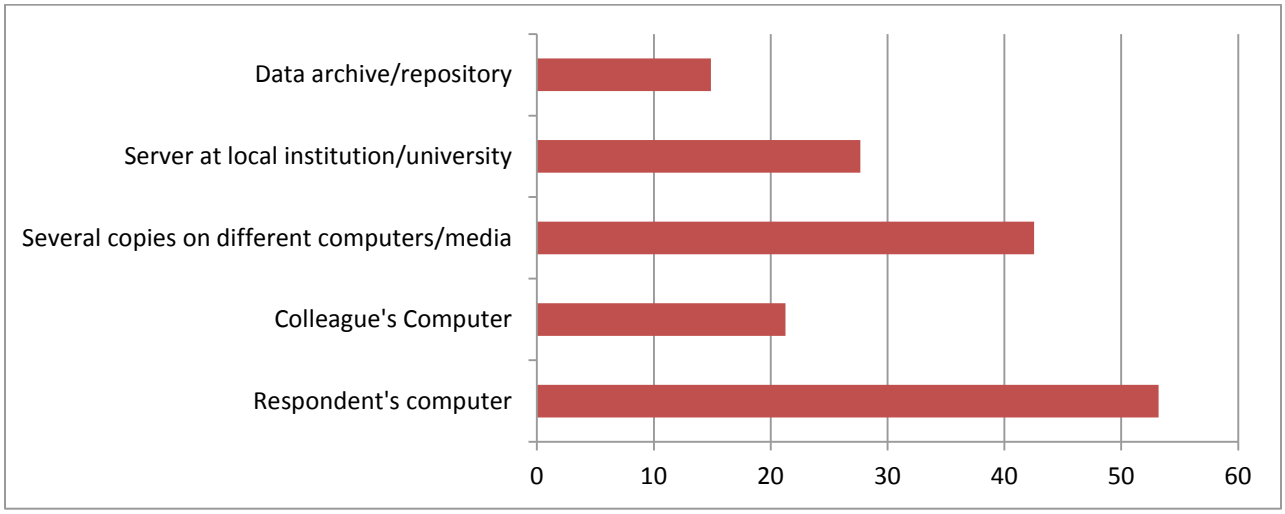
Figure 10: Distribution of the types of data retained from the last project (in percent)



N=44

Figure 11 below, which describes survey responses to the question of where the data are kept, provides further evidence for the necessity of a social science data repository in Kosovo. Only 15 percent of respondents indicated that their data had been stored in a data archive/repository, with the vast majority indicating that their data are stored on their own or their colleague's personal computers, and fewer than 30 percent indicated that their data were stored on a server at their local institution/university.

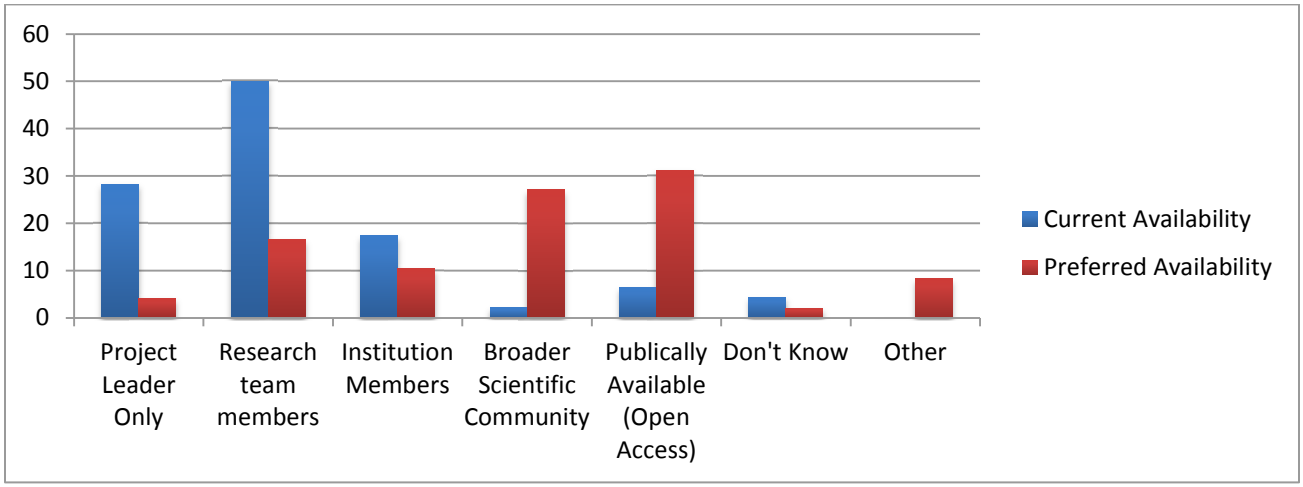
Figure 11: Distribution of storage locations of data from the last project (in percent)



N=44

1.4 Access to Data

Figure 12: Access to data from the most recent project (Current and Preferred Availability in percent)

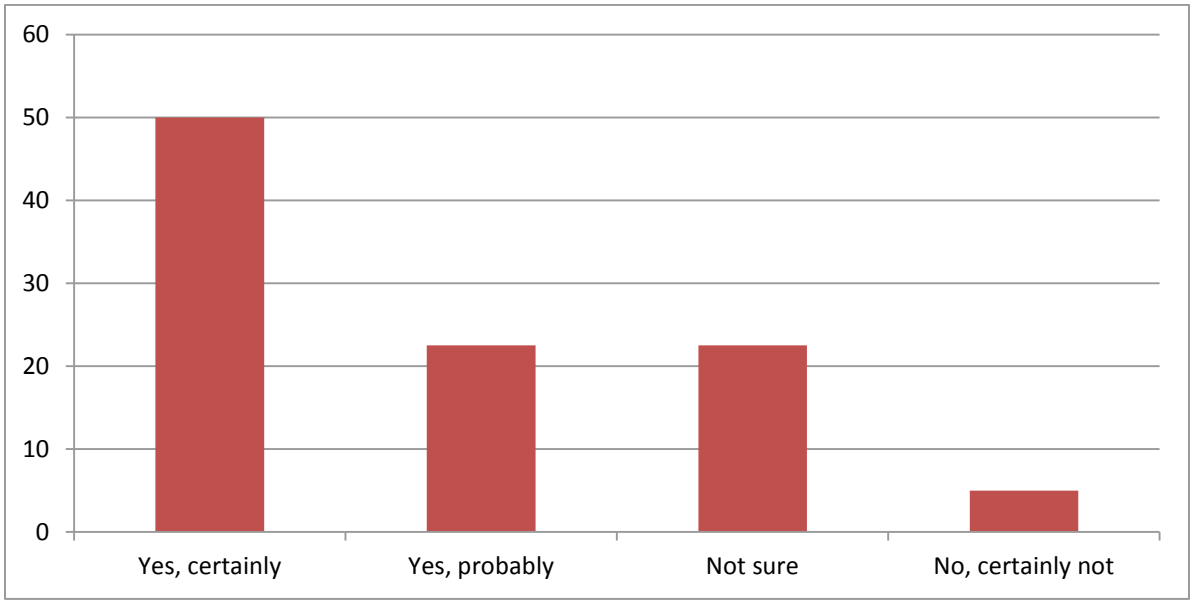


N=46

As for access to their data, the social science research community in Kosovo appears very reluctant to share their data. As shown in figure 13, almost 80 percent of the respondents currently make the data available only to the project leader (28%) or the research team members (50%). Only 6% indicated that the data from their most recent project is publicly available, and fewer than 20 percent have made the data available to their institution members. This situation differs from their stated ideal for access to

these data. When asked what the ideal of access to their research data should be, the majority (58 percent) indicated that their data should ideally be made available either through open access or to the broader scientific community.

Figure 13: Willingness to share data in a social science data archive, provided that data are preserved for the long term and safely secured



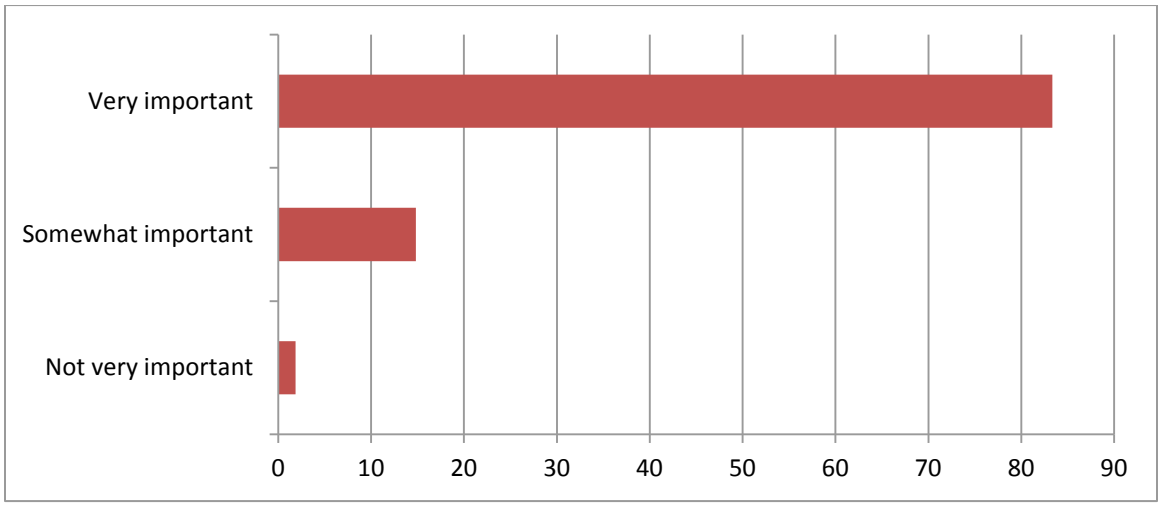
N=40

Provided that their data are preserved for the long term in a secured environment, however, more than 70 percent of our survey respondents stated that they would certainly or probably share them in a social science data archive. Five percent, on the other hand, noted that they would not, with 22 percent unsure. The small minority who appear reluctant to share their data noted that they would be worried about preserving confidentiality. In the comment section, some of the main concerns stated by the respondents had to do with “misinterpretation of data”, “personal and sensitive nature of the data” as well as “copyright misuse”. However, given that the absolute majority are very willing to share their data, it would seem that a well-run and secure data repository would be of great interest and potentially widely used by the research community.

1.5 Data Usage and Secondary Analysis

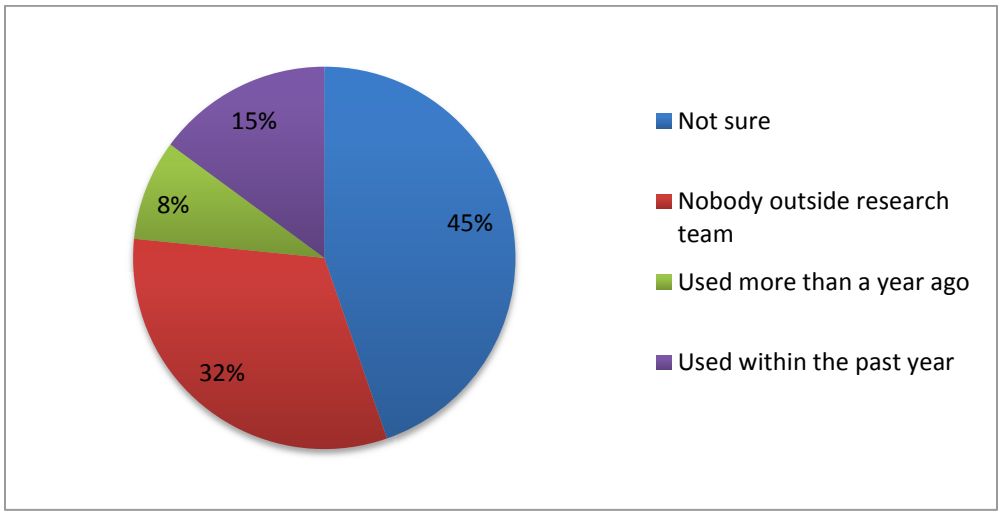
More than 80 percent of survey respondents agreed that sharing data within their discipline is very important (figure 14). However, as Figure 15 shows, the majority stated that they were either not sure (45%) or knew that nobody outside their research team (32%) had used their data for secondary analysis. Only 23 percent stated that they were aware that their data had been used for secondary analysis by someone other than their research team members.

Figure 14: Perceived Importance of sharing data within respondent's discipline (in percent)



N=54

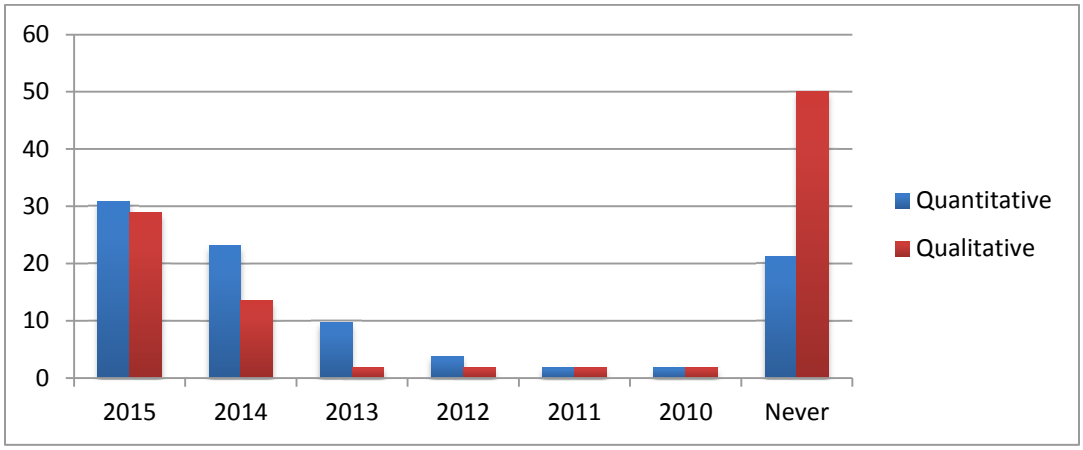
Figure 15: Respondent's knowledge of whether their data has been used for secondary analysis by someone other than their research team member (percent)



N=47

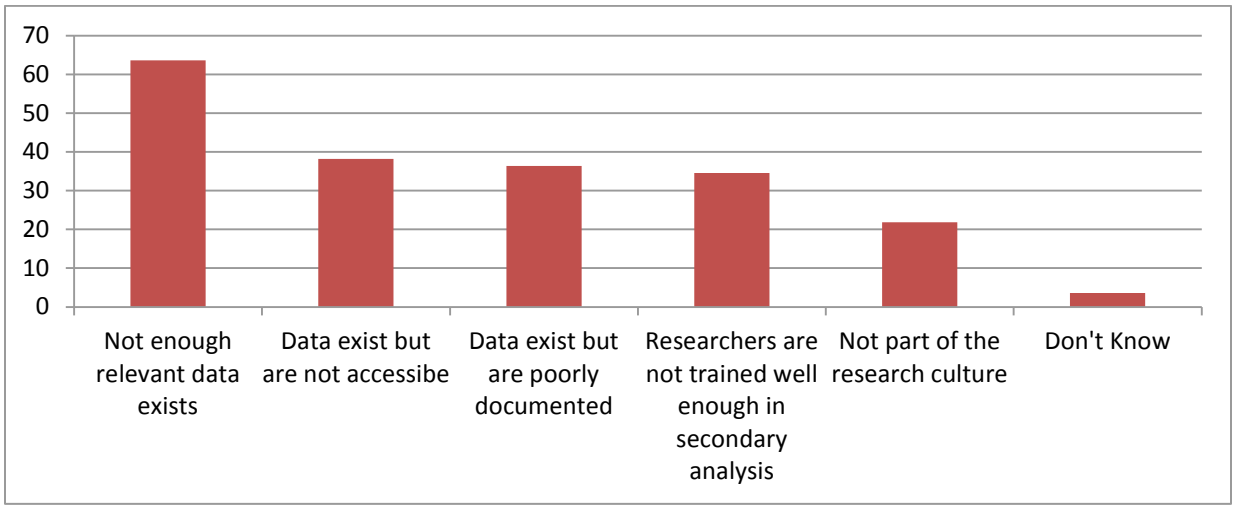
In addition, around 30 percent of the respondents noted that the last time they had used secondary data produced by others was in 2015, and 50% stated that they had never used qualitative data produced by others (Figure 16). Given that the majority of Kosovo's social science researchers conduct qualitative data gathering, however, it is important that a national data archive places particular emphasis on enabling the storage and dissemination of qualitative research data.

Figure 16: Last time respondent used secondary data produced by others (percent)



N=52

Figure 17: Perceived barriers to conducting secondary analysis in Kosovo (percent of survey respondents)

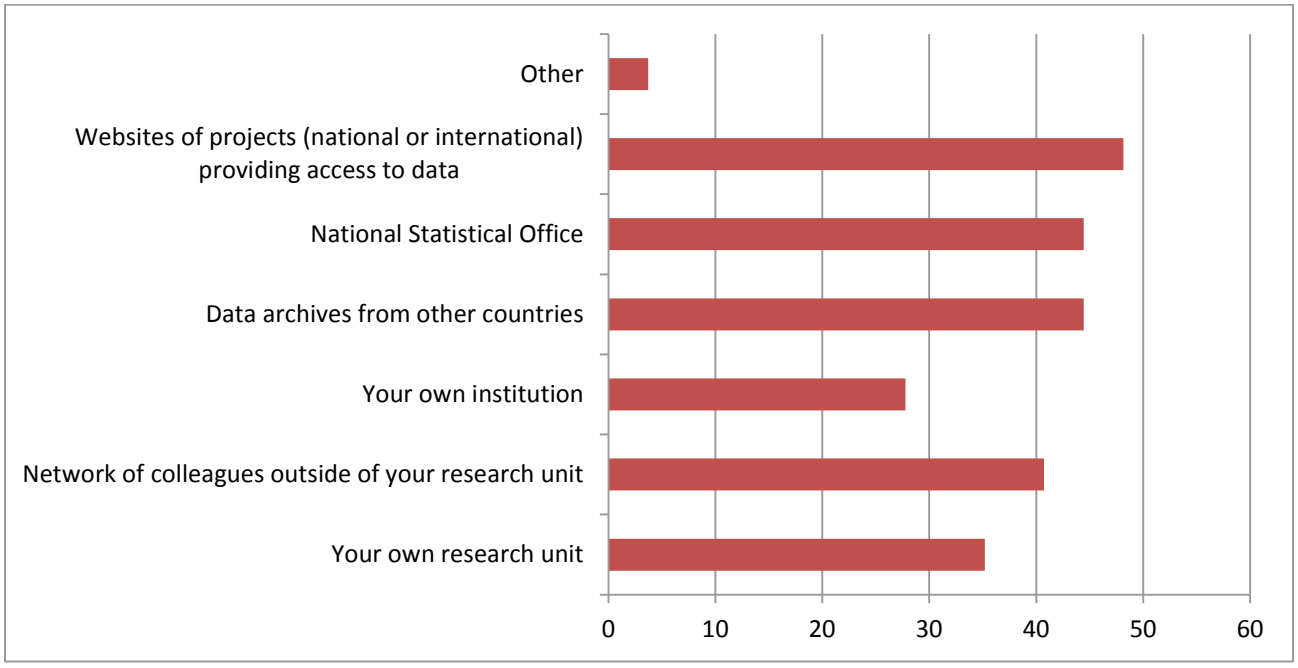


N=54

According to the survey results, the main barrier to conducting secondary analysis in Kosovo is perceived to be the lack of relevant data (64%) (figure 17). In addition to this, 38 percent of the respondents thought that the inaccessibility of the existing data had prevented them to use secondary analysis in their research. Approximately the same number of respondents shared the opinion that poor documentation of data (36%) and lack of training in secondary analysis (34%) contributed to the obstacles of conducting secondary analysis. Whereas around 22 percent of the respondents considered that secondary analysis was not part of research culture in Kosovo.

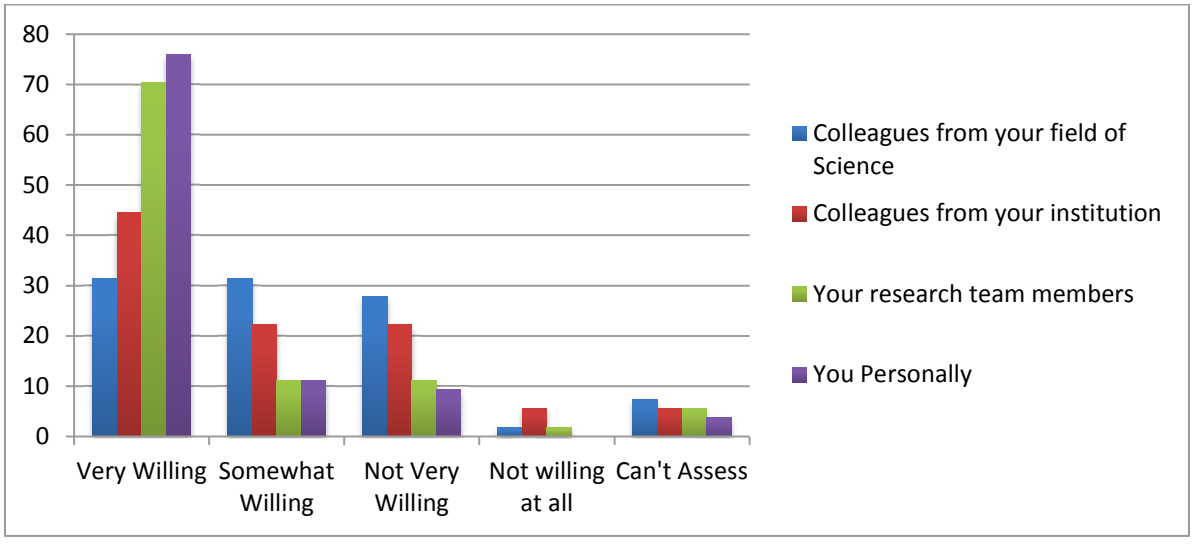
When asked about the sources used to obtain data, 48 percent of the respondents declared that they had obtained data from websites of national or international projects and 44 percent had used National Statistical Office and data archived from other countries (figure 18). Less widely used sources were the institutions (28%) and the research units (35%) of the researchers. Since the majority of the survey participants are affiliated with higher education institutions, it can be inferred that the lack of data documentation and systematization in these institutions makes them less preferable when it comes to data obtaining.

Figure 18: Sources used to successfully obtain research data (percent of respondents)



N=44

Figure 19: Perceived prevalent attitude with respect to sharing one's own research data (percent)



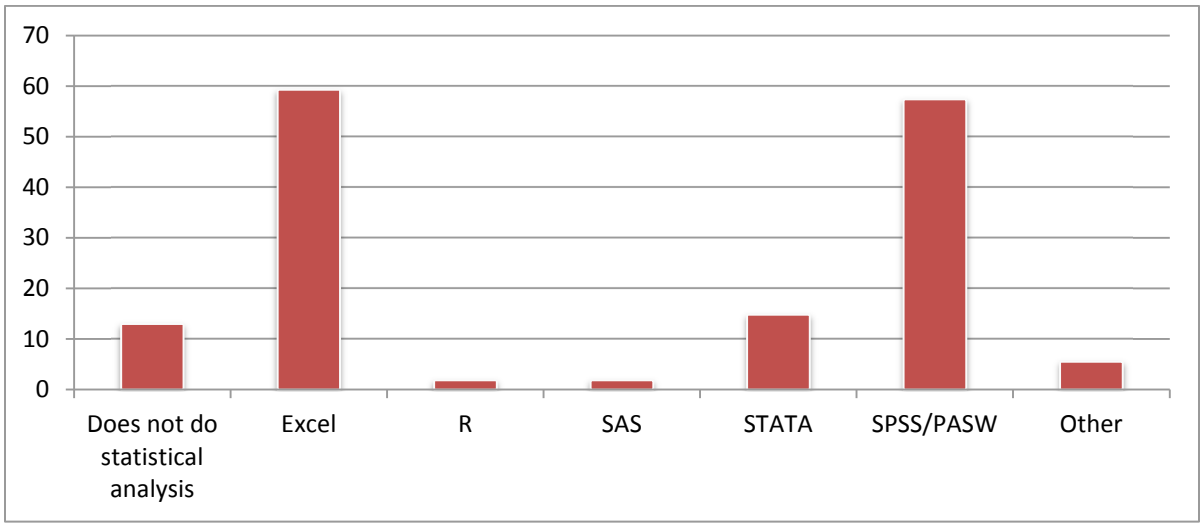
N=54

When asked about the prevalent attitude with respect to sharing one's own research data, most of the respondents (76%) declared to be very willing to share their data, while only 9 percent of them stated that they would not be very willing to share their data, mostly because of the fear of misinterpretation and misuse of them (figure 19). The majority of the respondents (70 %) also considered that their research team members are very willing to share their own data. On the other hand, 27 percent of respondents perceived that the colleagues from their field of study are not very willing to share their research data. 22 percent of the respondents also considered that colleagues from their institutions are not very willing to share their data. Moreover, 5.5 percent of the respondents declared that the colleagues from their institutions are not willing at all to share their data. This suggests that there is a lack of data sharing culture inside research institutions in Kosovo.

1.6 Data analysis

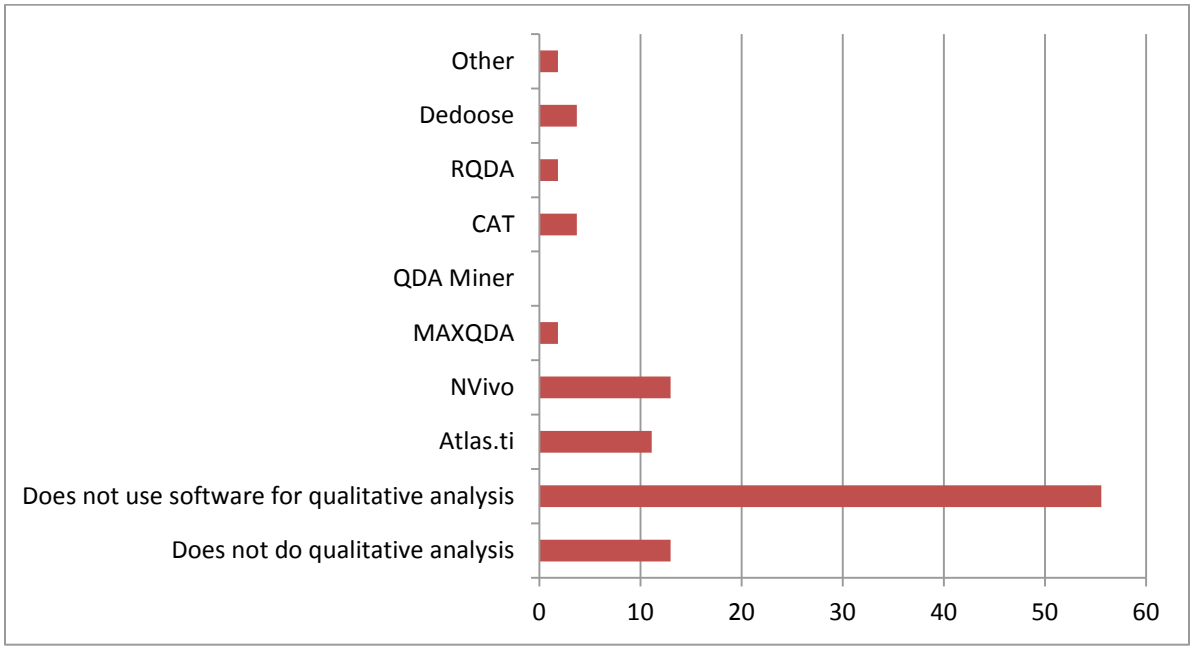
Our survey analysis indicates that the software of choice for conducting statistical analysis is Excel (59%), followed by SPSS (57%) and STATA (15%) (figure 20). On the other hand, the majority of those respondents who conduct qualitative analysis indicated that they do not use any software for qualitative analysis. Around 13 percent of the respondents declared that they are not engaged in qualitative analysis. One of the respondents even indicated in a comment that they had not been aware that such software exists. However, among those who do qualitative analyses, NVivo (13%) and Atlas.ti (11%) are the most used software programs (figure 21).

Figure 20: Software used to conduct statistical analysis (percent of respondents)



N=54
Note: Other include Limdep, eviews, mplus, Praar, CheckVocal

Figure 21: Software used to conduct qualitative analysis (percent of total survey respondents)

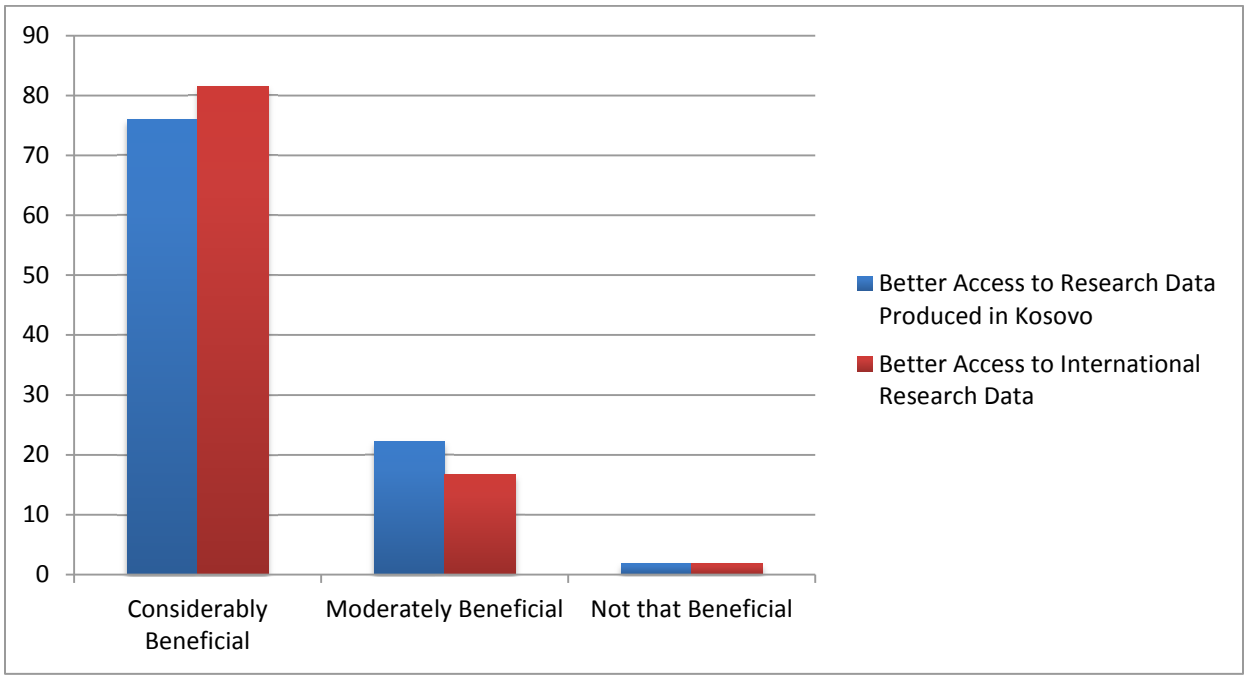


N=54
Note: Other includes Vensim

1.7 Perceived need for data archive

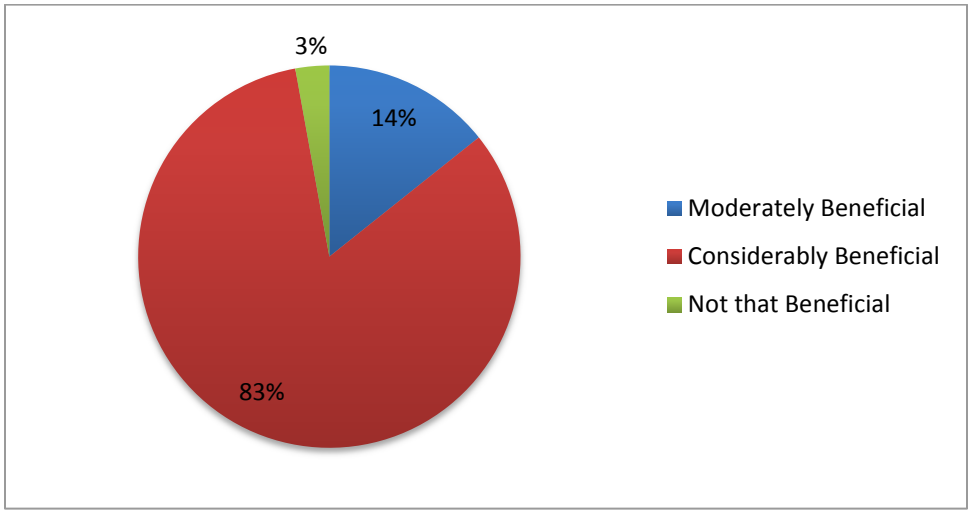
The vast majority of our respondents indicated that their teaching and research would benefit from better access to data produced in Kosovo or abroad (figure 22). Specifically, 81 percent of the respondents stated that their research would considerably benefit from better access to international research data, while 76 percent of them indicated that better access to research data produced in Kosovo would be considerably beneficial for their research. 83 percent of the respondents declared that better access to data produced in Kosovo or elsewhere would be beneficial for their teaching, 14 percent of them considered that better access would be considerably beneficial, whereas only 3 percent of the respondents perceived the better access to data produces in Kosovo or elsewhere as not that beneficial (figure 23).

Figure 22: Respondent's perceived benefits to their research from better access to data produced in Kosovo and elsewhere (percent of respondents)



N=54

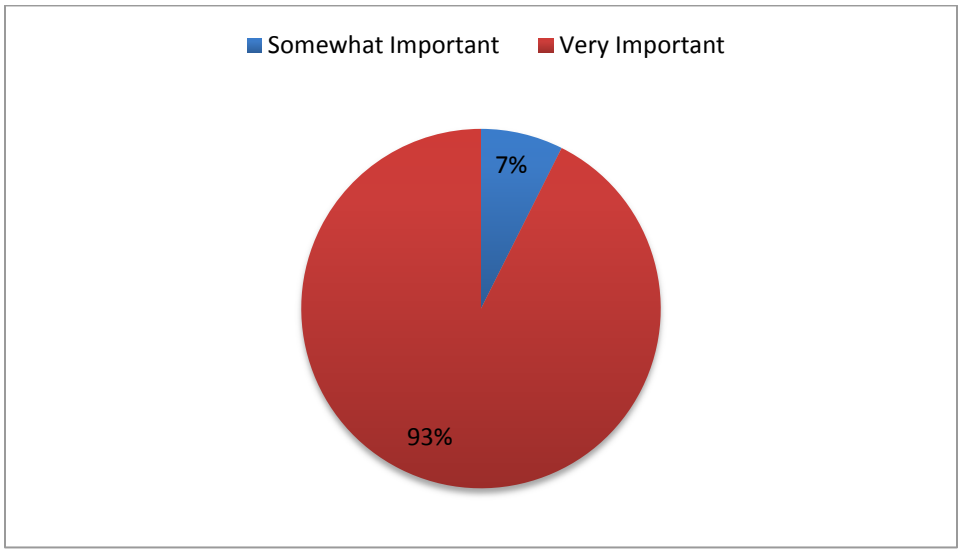
Figure 23: Respondent's perceived benefits to their teaching from better access to data produced in Kosovo and elsewhere (percent of respondents)



N=35

In addition, 93 percent of the survey respondents agreed that an institution that specialises in Data Archiving in Kosovo would be very important, and 7 percent of them considered that a specialised archiving institution would be somehow important (figure 24). This shows strong enthusiasm for such an initiative. In the additional comments section of the survey one of the respondents even suggested that such an initiative would be "revolutionary" for social science in Kosovo, although they and others also noted that it would probably be best if it were accompanied with a public campaign that sensitises the social science research community in Kosovo about the importance of data storing and dissemination.

Figure 24: Perceived usefulness of an institution that specialises in data archiving in Kosovo



N=54

Part III: Survey on existing infrastructure and interest in data archiving

In order to have a clear picture regarding the existing data infrastructure, current and planned data preservation practices, as well as the potential for establishing a national data service, during the period July-October 2015 a survey was conducted with the representatives of the main research infrastructure institutions in Kosovo as well as relevant NGOs involved in the field of social science research. The list of contacted institutions includes the following:

1. The National Library of Kosovo
2. State Agency of National Archive
3. Kosovo Agency of Statistics
4. University of Prishtina, Institute for Social and Humanistic Studies
5. Institute of History
6. Institute of Albanology
7. Academy of Sciences and Arts of Kosovo
8. Humanitarian Law Center Kosovo (NGO)
9. Open Data Kosovo (NGO)
10. Agency of Information Society/Ministry of Public Administration

These institutions were identified as potential hosts and users of a national archive and as future collaborators in establishing a data service. With respect to this, during face-to-face interviews, the representatives of these institutions were informed about the SEEDS project, with the intention of receiving helpful suggestions and recommendations and the need to map out their needs as well as possible obstacles during the process.

In general, none of the above-mentioned institutions except for the Kosovo Agency of Statistics uses any archiving software program. Most of them, in fact, do not possess a general database where all the data could be stored and classified. The data they preserve are mostly stored in physical form or on the personal computers of the researchers who belong to these institutions. Considering this, most of the questions in the questionnaire regarding data documentation and standards as well as technologies and tools used were not relevant during the interviews.

Below, we will present the most important information regarding data preservation and existing infrastructure obtained from the institutions that were contacted, including attitudes and recommendations for the possible establishment of a national data service and potential collaborators in this endeavour.

Kosovo Agency of Statistics (KAS) is the main national coordinator in Kosovo when it comes to collecting and preserving official statistics. The collected data from the Agency belong mainly to the field of economics, demography, and agriculture.

The possessed data are kept on special servers, so-called virtual servers and also in SQL 2012 databases. The data are considered to be safe due to the Back Up DPM, which contains recovery policies for any

eventual case of damage or lost. All the data in KAS are digitalised, deposited and convertible into different types of information.

Documentation standards used by KAS are part of Eurostat guidelines. The agency is also engaged in research projects of international character, such as MICS (Multiple Indicator Cluster Survey) as well as projects dealing with remittances in collaboration with UNDP. The agency annually compiles the Calendar of Publications, which lists in advance all the publications foreseen to be conducted within a year. Lately (2015), KAS has launched the Online Platform of data dissemination, named ASKDATA.[16]

KAS, as a governmental body that operates within the office of the prime minister, is primarily financed from the budget of the Republic of Kosovo. Nevertheless, the agency has also benefited from international donations on specific research projects. Intellectual property rights are transferred to the agency.

Three staff members are engaged in the sector of data preservation and dissemination: one in the office of IT and two others in the office of dissemination. They are familiar with software such as SPSS, STATA, R, Excel, CSPRO, SQL, Acess, and MapInfo. However, there is considerable need for additional capacity building as well as human resource development.

Access to the data possessed by KAS is defined according to the Law on Official Statistics of the Republic of Kosovo. Persons interested in accessing data are asked to submit a request where they should justify their motivation and their need for data access. In principle, all individuals without difference have equal right of data access. This said, the so-called Micro Data, which are collected at the local level, are not accessible due to possible infringement of confidentiality.

Kosovo State Archives Agency is mainly responsible for collecting, systemising, and processing archival documentation in order to make it usable for the general public. This agency supervises and coordinates six inter-communal archives.

Currently, the agency does not have any system of digitalisation due to lack of a software programme that would fit the specifics of the country and the lack of human resources.

The documents preserved in the agency are kept under specific physical conditions in archival boxes and protected shelves. The documents are preserved only in original copies. The representatives of the agency consider that under these conditions of preservation the possessed archival material is protected and accessible for the distant future. In 2005, the agency became member of the International Council of Archives and since then operates with these documentation standards: ISO (15489F) and ISAD (G).

In principle all archival materials are accessible to the general public. The agency does not have a catalogue of archived documents but it has hard copy guidelines. The main financial source is the budget of the Republic of Kosovo and it has not benefited from any international funds so far.

[16] https://askdata.rks-gov.net/PXWeb/pxweb/en/askdata/?rxid=0b4e087e-8b00-47ba-b7cf-1ea158040712/

The representative of the agency has listed two main challenges that, according to the representative, the agency faces regarding the preservation and dissemination of archival documents:

1. Lack of needed infrastructure, and
2. Lack of professional capacities

In total, the agency possesses 7300 meters in length documentation and 1200 archival files and has a staff of 90 workers.

Agency of Information Society (AIS) / Ministry of Public Administration is the main institution responsible for developing, maintaining and preserving the IT infrastructure and administrative data of public institutions. This Agency has been established within the Ministry of Public Administration as part of e-Governance initiatives that aim to reform the IT system in public institutions and increase data security and transparency.

Currently, the Agency administers and maintains all online governmental services, internet access of public institutions and governmental email. It also provides secure preservation of administrative data. The Electronic Data Center operates within the Agency, dealing mainly with administrative data systemizing and archiving. This Data centre possesses the Electronic Archive where all digitalised data are kept and preserved. The data are secured by special hosting and back up devices. However, no documentation standards or special standards for trusted digital repositories are used.

AIS has also invested in the digitalisation of archives of the academic units of University of Prishtina. Necessary hardware and software technology has been provided to the archive offices of each faculty. While some of the faculties use this technology regularly, others claim that the developed archiving program does not meet the needs and the characteristics of their daily work.

Because of its existing IT infrastructure, its professional human capacities and its national public character, the Agency of Information Society has been identified as one of the most likely cooperators of the SEEDS project. The representative of the Agency appreciated the idea of a national data service in the field of social sciences, claiming to be open to further communication and partnership. In this respect, AIS could serve the duty of providing technical support for the archive. Further crystallisation of the establishment plan of the data archive is needed to be done in order to specify the exact role and form of cooperation that AIS could have in this process.

The National Library does not possess any raw research data from the social sciences. However, it does contain preserved special collections, manuscripts, photographs, and video recordings, as well as copies of PhD theses. The material it possesses is also preserved in an electronic format. The representative of the National Library considers that the materials are kept in safe conditions and are not in danger of being lost.

The library has planned a research project titled “Memory of Kosovo”, which intends to collect data from fieldwork, photographs, narratives and video recordings. However, due to lack of financial resources, the implementation of the project has not begun yet.

The Institute of Albanology has a considerable number of files, mainly with qualitative data collected from fieldwork in the domains of folklore, ethnology, oral history, and archaeology. The data are preserved in a physical form in the Archive of Scientific Research Activity of the Institute. There have been continuous endeavours regarding the digitalisation of the data files and, in 2014, with a new and more qualitative device this process has been intensified.

The Institute has only one staff member responsible for data archiving, preservation and systematisation. This minimal number of staff is reflected in the slowdown of the data digitalisation process. It is considered to be in immediate need for additional staff members trained in accordance with international practices of documentation and archiving.

At this end in general, all the representatives of the institutions and organisations that were interviewed agree that data produced from research in the social sciences in Kosovo are not preserved properly for a long-term period and thus are in danger of being lost. However, they do not feel competent enough to declare to what extent the research data produced in Kosovo are lost because they are not stored in a safe setting for the long-term. They are critical regarding the data storing system in the institutions that they represent and declare that one of the reasons for this poor infrastructure of data preservation and dissemination is the lack of financial support and trained staff members.

The idea for a possible National Data Service for social sciences was unanimously supported by the interviewed representatives and was considered as an important and useful initiative toward a more consolidated data preservation system. The prevalent opinion was that all the stakeholders including researchers, academic institutions, research institutes, and think tank organisations would benefit from such a data service. No firm recommendations or possible models on how the service could function or be organised were given however. Nevertheless, some of the important steps that should precede the establishment of the national data service are below:

- a round table with the representatives of all stakeholders should be organised in order to discuss the need and the potential of each of them;
- a promotional campaign of the project should be led in all academic and research institutions and beneficial aspects of data sharing should be emphasised;
- a willingness for data sharing has to be expressed from all the stakeholders involved in the project;
- a sustainable financial means has to be ensured;
- a leading mechanism (e.g. a core group) of the project has to be established.

Among others**,** as an important possible collaborator we find Open Data Kosovo, which is one of the few initiatives in Kosovo that aims to promote and enable open access and data sharing. Its activity is directly linked with enhancement of good governance, government transparency, and accountability through development of innovative technological applications and programmes. Some of the applications that have been developed by Open Data include Kosovo Municipality Procurement Visualizer, Kosovo Election Monitoring Visualizer, Gender Corruption Survey Visualizer. All the data from

these surveys can be accessible online. Open Data Kosovo has also assisted the government of Kosovo with professional help to create the Portal of Open Data (http://data.rks-gov.net/en/) where administrative data from municipalities and ministries are posted online and made available for the general public. Open Data Kosovo sees the SEEDS project as very useful for the academic environment and general public space of Kosovo and is open for future possible cooperation with our project.

Annexes

Annex 1. List of interviews held

Institution	Person interviewed	Date
The National Library of Kosovo	Izet Avdyli, Secretary of the Library	10.08.2015
State Agency of National Archive	Ruzhdi Panxha, Director of the Agency	11.08.2015
Kosovo Agency of Statistics	Burim Limolli, Head of IT Office	12.08.2015
Institute of Albanology	Hysen Matoshi, Director of the Intitute of Albanonogy	13.08.2015
University of Prishtina, Institute for Social and Humanistic Studies	Shemsi Krasniqi, Professor of Sociology	17.08.2015
Institute of History	Arbër Kuçi, Secretary of the Institute	20.08.2015
Academy of Sciences and Arts of Kosovo	Hilmi Zogjani, Head of the Department of Library and Achieve	28.09.2015
Humanitarian Law Center Kosovo	Bekim Blakaj, Executive Director	29.09.2015
Open Data Kosovo	Georges Labreche, Chief Data Officer	09.10.2015
Electronic Data Center/ Agency of Information Society	Xheladin Graiqevci, IT expert	22.10.2015
Agency of information Society/Ministry of Public Administration	Selim Lulaj, Directorate for Development of e-Governance	26.10.2015
Faculty of Philosophy, UP	Agim Hyseni, Archivist	26.10.2015

Annex 2. Questionnaires

Annex 2. A.: Questionnaires Researchers' questionnaire

[SEEDS] Data use Survey

In this questionnaire, we are particularly interested in your research practices and needs related to collecting empirical data, their preservation and use for secondary analysis.

This survey forms part of an international project "South East European Data Services" - ***SEEDS***, coordinated by Swiss Centre of Expertise in the Social Sciences. The purpose of the project is to establish durable infrastructures for storage and secondary use of data generated in social science research. The project is a collaborative effort of eight countries, including partners who already have functional services for data archiving (Switzerland, Slovenia, Croatia, Serbia), together with regional partners, whose countries have yet to establish these services (Albania, Kosovo, FYR Macedonia, Montenegro).

Thank you for taking 10-15 minutes to complete the questionnaire. The information provided by you in this questionnaire will be used for research purposes only. It will not be used in any manner that would allow identification of your individual responses.

This questionnaire has 38 questions.

About you

First, we would like to get some information about you.

[] What is your current principal activity? *

If you choose 'Other:' please also specify your choice in the accompanying text field.

Please choose **only one** of the following:

- ○ Undergraduate student
- ○ Doctoral student / research or teaching assistant
- ○ Researcher / professor
- ○ Project leader
- ○ Head of institution
- ○ Other

Choose one answer only

[] With what type of institution are you currently principally affiliated? *

If you choose 'Other:' please also specify your choice in the accompanying text field.

Please choose **only one** of the following:

- ○ Higher education institution
- ○ University research institute
- ○ Public research institute
- ○ NGO/Think tank
- ○ Currently not employed
- ○ Other

Choose one answer only

[] Do you work in public, private or nongovernmental sector? *

Only answer this question if the following conditions are met:
Answer was NOT 'Public research institute' at question '2 [Q2]' (With what type of institution are you currently principally affiliated?)

Please choose **only one** of the following:

- ○ Public sector
- ○ Private sector
- ○ Nongovernmental sector

This question concerns your primary job.

[] What is your principal research discipline? *

If you choose 'Other:' please also specify your choice in the accompanying text field.

Please choose **only one** of the following:

- ○ Economics
- ○ Sociology
- ○ Psychology
- ○ Education science and teacher training
- ○ Library and information sciences
- ○ Political science
- ○ Journalism
- ○ Business and administration
- ○ Law
- ○ Organizational sciences/Management
- ○ Public administration
- ○ History
- ○ Anthropology
- ○ Other

Please select most appropriate of listed International Standard Classification of Education (ISCED) fields.

[] What country are you working in? *

Please choose **only one** of the following:

○ Albania

○ Kosovo

○ Macedonia

○ Montenegro

Your research

For this survey, the term "data" refers to **raw data**, that is, information collected in specific research projects that is recorded in machine readable format and used for analytic purposes. This could be survey data, interviews in sound files, video footage, notes, images, etc.

By "data", we do not mean analyses, descriptions, statistics, facts, or conclusions that appear in reports, papers, websites, or scientific publications.

[] **In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.**

Please choose **only one** of the following:

○ Yes

○ No

Please count only research where you were involved at a substantive level with planning or organization of research, fieldwork management, cleaning or coding of research data.

[] **Enter the number of datasets that you have produced or helped to produce during the past 5 years.**

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.)

Only an integer value may be entered in this field.

Please write your answer here:

For example, one survey or series of interviews are both to produce one dataset. However, one research project might produce more than one dataset if several methods (e.g. focus groups and questionnaire) or distinct data collection efforts took place (e.g. employer survey and employee survey).

Your most recent research

The following questions are about your **most recent** research effort which **involved data collection**.

[] In which year was the fieldwork (or data collection phase) completed?

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.)

Your answer must be between 1960 and 2015

Only an integer value may be entered in this field.

Please write your answer here:

[] **Which data collection method was applied in this research?**

(e.g. online questionnaire, structured interview, focus groups, experiment, ...)

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.)

Please write your answer here:

If research involved data collection through application of several methods, list all of them, separated by ***";"***

[] **What was the approximate scope of raw data collected in this research?**

(e.g.: 8000 respondents; or 15 focus groups; or 50 firms; or 700 case reports; or 500 newspaper articles; or 200 hrs of video footage)

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.)

Please write your answer here:

[] How was this research financed?

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.)

Please choose **all** that apply:

☐ Public funding through national science funding bodies (science ministry, science foundation...)

☐ Public funding from other sources (other ministries, state agencies, cities and municipalities...)

☐ International funding/project

☐ Private sector

☐ Own funding (institution you are working in paid from its own funds)

☐ Other:

Mark all that apply

[] **Apart from you, how many researchers were involved in this research project?**

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.)

Only an integer value may be entered in this field.

Please write your answer here:

Persons involved in field-roles during execution phase only (such as respondents in surveys) are not to be counted.

Data preservation

For this survey, the term "data" refers to **raw data**, that is, information collected in specific research projects that is recorded in matching readable format and used for analytic purposes. This could be survey data, interviews in sound files, video footage, notes, images, etc.

By "data", we do not mean analyses, descriptions, statistics, facts, or conclusions that appear in reports, papers, websites, or scientific publications.

[] **After you completed your last research project, did you or your research team members keep/retain the data?** *

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.)

Please choose **only one** of the following:

○ Yes

○ No

[] What kind of data was kept?

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.) *and* Answer was 'Yes' at question '13 [Q12]' (After you completed your last research project, did you or your research team members keep/retain the data?)

Please choose **all** that apply:

☐ Raw data

☐ Cleaned data (coded, anonymized, ...)

☐ Data prepared for analysis (with transformations, created indexes, recoded)

☐ Well documented with metadata

☐ Other:

Mark all that apply

[] Have you used any special documentation/metadata standard for description of your research data?

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative

data.) *and* Answer was 'Well documented with metadata' at question '14 [Q13]' (What kind of data was kept?)

Please choose **all** that apply:

☐ DDI

☐ DC

☐ ISO 11179

☐ Internal/institutional documentation standard

☐ Don't know

☐ Other:

Mark all that apply

[] Where is the data from your last project kept? *

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.) *and* Answer was 'Yes' at question '13 [Q12]' (After you completed your last research project, did you or your research team members keep/retain the data?)

Please choose **all** that apply:

☐ On my computer

☐ On my colleague's computer

☐ Several copies on different computers and/or different media

☐ Server at my local institution/university

☐ Data archive/repository

☐ Don't know

☐ Other:

Mark all that apply

[] Who may be granted access to the data from your last project for research use? *

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did

you produce or help to produce any research data? This could be quantitative and/or qualitative data.) *and* Answer was 'Yes' at question '13 [Q12]' (After you completed your last research project, did you or your research team members keep/retain the data?)

Please choose **all** that apply:

☐ Just the project leader

☐ Research team members

☐ Members of my institution

☐ Broader scientific community

☐ The data is publicly available (open access)

☐ Don't know

☐ Other:

Mark all that apply

[] In your opinion, what would be the ideal level of access to this research data? *

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.)

If you choose 'Other:' please also specify your choice in the accompanying text field.

Please choose **only one** of the following:

○ Just the project leader

○ Research team members

○ Members of my institution

○ Broader scientific community

○ The data should be publicly available (open access)

○ Don't know

○ Other

Mark one option that you consider most suitable

[] If you knew that your data would be preserved for the long-term in a secure environment, and shared only with accredited researchers, would you be willing to provide your data to a social science data archive? *

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.) *and* Answer was NOT 'Data archive/repository' at question '16 [Q15]' (Where is the data from your last project kept?)

Please choose **only one** of the following:

- ○ Yes, certainly
- ○ Yes, probably
- ○ Not sure
- ○ No, probably not
- ○ No, certainly not

Choose one answer only

Data sharing and secondary analysis

Now we would like to ask you several questions regarding the sharing of data and secondary analysis.

By **sharing** we mean practices where researchers access and use data that they themselves did not produce.

Secondary analysis is defined as analysis of data that were produced by others, where one was not involved in the original research.

[] With respect to your own discipline, how important is the sharing of research data? *

Please choose **only one** of the following:

- ○ Very important
- ○ Somewhat important
- ○ Not very important
- ○ Not important at all

Choose one answer only

[] Do you know if any other researcher outside your own team had used for secondary analysis any of the research data that you produced? *

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '6 [Q5]' (In the context of your research activity within the past 5 years, did you produce or help to produce any research data? This could be quantitative and/or qualitative data.)

Please choose **only one** of the following:

○ Nobody outside my team had ever used research data that I/we have produced

○ Yes, I know of an occasion when my/our data was used, but this was more than a year ago

○ Yes, I know that my/our research data was used for secondary analysis recently, the most recent occasion being within the past year

○ I am not sure if anybody used my/our data for secondary analysis

Choose one answer only

[] **When was the last time that you analyzed quantitative data that were not produced by yourself or your research team.**

Please write your answer(s) here:

Year

Month (1-12)

If you have never done such analysis, just enter **0** in the year field.

[] **When was the last time that you analyzed qualitative data that were not produced by yourself or your research team?**

Please write your answer(s) here:

Year

Month (1-12)

If you have never done such analysis, just enter **0** in the year field.

[] In general, what are the barriers to conducting secondary analysis in your country? *

Please choose **all** that apply:

☐ Not enough relevant data exist

☐ Data exist but are not accessible

☐ Data exist but are poorly documented and unusable

☐ Researchers are not trained well enough in secondary analysis

☐ It is not part of the research culture

☐ Don't know

☐ Other:

Mark all that apply

[] There are different ways to obtain research data produced by others. Please indicate all the sources that you ever used successfully to obtain such data. *

Only answer this question if the following conditions are met:

-------- Scenario 1 --------

Answer was NOT '0' at question '22 [r15q3]' (When was the last time that you analyzed quantitative data that were not produced by yourself or your research team. (Year))

-------- or Scenario 2 --------

Answer was NOT '0' at question '23 [Q22]' (When was the last time that you analyzed qualitative data that were not produced by yourself or your research team? (Year))

Please choose **all** that apply:

☐ Your own research unit

☐ Network of colleagues outside of your research unit

☐ Your own institution

☐ Data archives from other countries

☐ National Statistical Office

☐ Websites of projects (national or international) providing access to data

☐ Other:

Mark all that apply

[] What statistical software do you commonly use for your quantitative analyses? *

Please choose **all** that apply:

☐ I don't do quantitative analysis

☐ Excel

☐ R

☐ SAS

☐ Stata

☐ SPSS / PASW

☐ Other:

Mark all that apply

[] What software do you commonly use for your qualitative analyses? *

Please choose **all** that apply:

☐ I don't do qualitative analysis

☐ I don't use any software for qualitative analyses

☐ Atlas.ti

☐ NVivo

☐ MAXQDA

☐ QDA Miner

☐ CAT

☐ RQDA

☐ Dedoose

☐ Other:

Mark all that apply

[] Would your scientific work benefit if you had better access to *

Please choose the appropriate response for each item:

	Yes, considerably	Yes, moderately	No, not very much	No, not at all
research data produced in your country	○	○	○	○
international research data	○	○	○	○

[] **In your opinion, what is the prevalent attitude with respect to sharing one's own research data among...***

Please choose the appropriate response for each item:

	Very willing	Somewhat willing	Not very willing	Not willing at all	Can't assess
colleagues from your field of science	○	○	○	○	○
colleagues from your institution	○	○	○	○	○
your research team members	○	○	○	○	○
you personally	○	○	○	○	○

Mark one assessment for each group mentioned and for yourself.

[] What are the main reasons for which you are not very willing to share your own research data with others?

Only answer this question if the following conditions are met:
Answer was 'Somewhat willing' *or* 'Not very willing' *or* 'Not willing at all' at question '29 [Q29]' (In your opinion, what is the prevalent attitude with respect to sharing one's own research data among... (you personally))

Please write your answer here:

[] Does your professional activity include teaching responsibilities? *

Please choose **only one** of the following:

○ Yes

○ No

[] In the context of your teaching, how often do you analyze or discuss research data? *

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '31 [Q31]' (Does your professional activity include teaching responsibilities?)

Please choose **only one** of the following:

○ Regularly

○ Sometimes

○ Rarely

○ Never

Choose one answer only

[] Which type of data do you use in your teaching? *

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '31 [Q31]' (Does your professional activity include teaching responsibilities?) *and* Answer was 'Rarely' *or* 'Sometimes' *or* 'Regularly' at question '32 [Q32]' (In the context of your teaching, how often do you analyse or discuss research data?)

Please choose **all** that apply:

☐ Data collected by students through the courseworkstion

☐ Data from past projects in which you have participated

☐ Publicly available data and datasets (e.g. international surveys)

☐ Artificially generated data or datasets included with the textbooks/software

☐ Other

☐ Other:

Mark all that apply

[] Would your teaching benefit if you had greater access to more national or international data? *

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '31 [Q31]' (Does your professional activity include teaching responsibilities?)

Please choose **only one** of the following:

○ Yes, considerably

○ Yes, moderately

○ No, not very much

○ No, not at all

Choose one answer only

[] In your view, how useful could be an institution that specializes in data archiving in your country? *

Please choose **only one** of the following:

○ Very important

○ Somewhat important

○ Not very important

○ Not at all important

[] Please include any comments that you think would be helpful for understanding the social science research community in your country, or the likelihood of success of a national data infrastructure/archive at the service of researchers.

Please write your answer here:

[] If you are interested in archiving of and access to research data in your country, please leave us your contact information so that we can send you further information about our project.

Please write your answer(s) here:

Name

Institutional affiliation

E-mail address

Your contact data will be kept confidential and used for contact purposes only. ***Data entered here won't be linked with answers that you have provided in this survey without your*** **permission.**

[] Can we associate your email address with information about recent data collections in which you participated? Only questions from sections "Your research" and "Your most recent research" will be linked. * Please choose **only one** of the following:

○ Yes

○ No

Submit your survey.
Thank you for completing this survey.

Annex 2. B.: Questionnaires Guidelines for semi-structured interviews with government counterparts and potential

Research policy setting: interview with policy makers/funders

PREPARATION

Stakeholder identification

- Who is involved in making science policies?
- What are the main sources of financing scientific work?

Reading relevant documents, domestic and international relevant documents:

- Current documents

- Project funding criteria and rules

- Criteria for evaluation of scientific work

- identify everything concerning science infrastructure and data archiving in relevant documents
- are we one of the members of ESFRI? (Croatia and Serbia are associated members)

http://ec.europa.eu/research/infrastructures/pdf/esfri/membership/esfri_membership_july_2011.pdf#view=fit&pagemode=none

http://ec.europa.eu/research/infrastructures/pdf/esfri/home/esfri_inspiring_excellence.pdf#view=fit&pagemode=none

Good data source (not all information are up to date)

http://erawatch.jrc.ec.europa.eu/erawatch/opencms/information/country_pages/

Invitation letter

- Maximum one page

- introduce the project and yourself, explain what data archive is

- ask for meeting to talk about relevant issues on establishing data archive

- attach topics for discussion and questions (so that they can prepare themselves

TOPICS FOR DISCUSSION

A. Needs

B. Current policies and activities, plans for future

C. Possible institutional solutions

D. Financing

QUESTIONS

A. Have you met the need for archiving of research data in the social sciences?

Do you rely on the results of research in social sciences when proposing/designing scientific policies and for other activities? If so, can you give some examples? In this process was there a need for examining the raw data obtained in these studies?

Have you ever used the online available source data or the results of research (eg Eurostat)? What data sources did you use?

Have you met the requirements of the users / authors of the study (scientists and researchers) to archive research data in order to preserve and use them in future research?

B. Is there something in the scope of your current policy and actions with regard to data archiving in the social sciences? What are the plans for the future?

What is the current science policy in relation to the archiving of research data?

Are there any requirements related to research data when financing projects or other activities? (E.g. if funding publishing scientific journals, criteria could include and requirement related to research data archiving)

Is there a policy relating to open access to research results?

If there is nothing currently:

- Why is not there?

- Can this be changed and how?

(Assumption is that if there is need, and no current activities, there must be plans for the future)

C. What do you think how an institutional solution for data archiving should be established? What kind of institution should carry this task? (Scientific institute? faculty? university? library?) Centralized, distributed?

D. What are the possible sources of funding for activities related to the collecting, processing, use and dissemination and permanent preservation of research data? Available or prospective (EU)

Annex 2. C Questionnaires Data Services

SEEDS - South-Eastern European Data Services

"South East European Data Services" - **SEEDS**, coordinated by Swiss Centre of Expertise in the Social Sciences is a project that aims to establish durable infrastructures for storage and secondary use of data generated in social science research. The project is a collaborative effort of eight countries, including partners who already have functional services for data archiving (Switzerland, Slovenia, Croatia, Serbia),

together with regional partners, whose countries have yet to establish these services (Albania, Kosovo, FYR Macedonia, Montenegro).

We are particularly interested in existing work within [*name of country*] with respect to the long-term preservation of research data and the potential for a national social science data infrastructure.

*(For the interviewer: Please instruct and remind where relevant that the term "data" refers to **raw data**, that is, information collected in specific research projects that is recorded in machine readable format and used for analytic purposes. This could be survey data, interviews in sound files, video footage, notes, images, etc. By "data", we do **not** include analyses, descriptions, statistics, facts, or conclusions that appear in reports, papers, websites, or scientific publications.)*

About your institution

1) Name of institution

2) Type of institution

a) University
b) Public research institute
c) Private research institute
d) Library
e) National/Regional archive
f) National Statistical Institute
g) Other (specify)

3) What is your institution's principal research discipline? (*Only applicable if respondent answered b or c on Question 2*)

4) What is your institution's principal research methodology focus? (*Only applicable if respondent answer b or c on Question 2*)

a) Quantitative methodology
b) Qualitative methodology
c) Mixed methodology
d) Other (specify)
e) Not applicable

5) What is the scope (discipline) of your data collection? *(Only applicable if respondent answered d, e, or f on Question 2)*

Existing infrastructure and data capacities

6) Does your institution store or disseminate data for use by researchers?

a) Yes (*go to question 8*)
b) No (*go to question 7*)

7) What happens to the data produced at your institution? Are they at risk of being lost forever? (*Skip to question 34 after answering this question.*)

8) What discipline(s) are covered by the data you store or disseminate?

a) Social sciences
b) Humanities
c) Other (specify) _____

Data preservation

9) How does your institution store research data for the long-term? Could you briefly describe this process?

10) How safe are the research data that are preserved at your institution? That is, are the data kept on servers that are protected? Are there backup or formal preservation systems? (*Interviewer to explore which.)*

11) Are the data treated in a way that assures that they can be accessed and used again in 20-50 years? (e.g., kept, with necessary documentation, in a non-proprietary machine readable format)

Data documentation and standards

12) Does your institution use any documentation standards for the research data that it preserves? (*If yes:)* Which standards does it use, e.g., DDI, Dublin Core, other?

13) Does your institution follow a particular standard for trusted digital repositories, such as OAIS or the Data Seal of Approval?

14) Does your institution participate in any international survey projects that aim to harmonise datasets from different countries for comparative purposes?

15) Does your institution make use of any thesaurus to translate and/or index your data? *(If answer is yes, interviewer to ask for details).*

Tools and technologies

16) Does your institution have any experience with particular data service tools for the social sciences, such as NESSTAR, FEDORA or Dataverse? (*If yes:*) Could you briefly describe the purpose for which you use these and your experiences of using them?

Data discovery and dissemination

17) Does your institution allow access to the research data that it preserves?

(*If no, skip to next section*.)
18) Who is allowed to have access to the data? *(Explore if respondent answers researchers, whether this includes researchers in their own organization only or also in other organizations)*

19) What are the conditions that must be met to access these data?

20) Do you have some kind of authentication system that identifies who is eligible to access the data?

21) Are all of the data equally accessible? That is, are there some data that are more accessible than others?

22) Can the data be accessed from outside of the country? (*Interviewer: If yes, probe to see whether the access is for national researchers who are abroad, or whether there are also foreign researchers who can access the data.*)

23) By what means does your institution disseminate its research data? For example, are the data sent out on CDs, or are they available on screen, or can they be downloaded from a website?

24) Does your institution have a data catalogue that allows people to find the data that they are looking for? (*If yes:)* Is the catalogue visible outside of your institution, and what software is used to enable this?

Data policy and service funding

25) Does your institution have any policy or other documents regarding long-term preservation of research data? (*If yes:)* Could you briefly describe the policy or documents? (*Also, ask if they could send it to us by e-mail.*)

26) Does your institution use any legal agreements for storing, disseminating, and/or using research data? This might be in the form of deposit contracts or end-user licenses. (*If yes :*) Could you give a brief description of these legal agreements*? (Also, ask if they could send them to us by e-mail.)*

27) Do the intellectual property rights remain with the researcher/data producer, or are they transferred to your institution?

28) Does your institution receive any external funding for archiving/data service activities? (*If no:)* How is the work of data preservation and dissemination paid for at your institution? *(If yes, ask how much, whether the income is constant and whether it is dedicated to particular activities, e.g., preservation, dissemination, user support).*

29) In your view, would your institution do more to preserve and disseminate research data if it had more resources?

Staff capacities

30) Does your institution have dedicated staff for the preservation and dissemination of research data? (*If yes:)* How many?

31) Do these staff members have specific training in data preservation and dissemination? (*If yes:)* Please explain.

32) What sort of additional training do you think would be needed for your staff to acquire sufficient knowledge about data service policies and practices?

33) What kind of statistical software package experience do they have?

a) SAS
b) SPSS
c) STATA
d) R
e) MathLab
f) Excel
g) Other (specify)

About a possible National Data Service for the social sciences

34) In general, to what extent are *social science* research data preserved for the long-term in [*name of country*]?

35) Could you estimate roughly how much research data produced in [*name of country*] are lost because they are not stored in a safe setting for the long term (in percentage)?

36) Do you think it would be useful or important to establish a national data service for the social sciences in [*name of country*]? Please explain why or why not.

37) (*If yes to question* 35) we are interested in your views about what a national social science data service might look like in [*name of country*]. Can you say what key functions such an institution should serve? How might it be structured and what kinds of relationships should it have to other institutions?

38) Are there any existing national data service infrastructures for other disciplines in [*name of country*] (for example, in the humanities, medicine, climate and environment, natural sciences, or technology)?

a) Humanities

b) Medicine

c) Climate and environment

d) Natural Science

e) Technology

f) Other (specify)

g) No