

Researchers' Interest in Data Service in Bosnia and Herzegovina, Croatia, and Serbia

by Aleksandra Bradić-Martinović¹ and Aleksandar Zdravković²

Abstract

This paper presents findings on existing potentials for the establishment of social sciences digital data archives in Bosnia and Herzegovina, Croatia, and Serbia. Findings are based on a standardized survey that was conducted in all three countries on representative sample of social sciences researchers, with a 63% average rate of completed questionnaires. Results of the survey show that the potential for establishment of digital data archives in all considered countries are large, regarding the scope of data produced, as well as positive attitude of researchers toward data sharing and benefits of data archives. Also, results point to lack of knowledge in dealing with metadata as the main obstacle for data sharing, which further implies that existence of national data archives, with staff trained to underpin researchers' efforts in process of data documentation and preservation, should be very beneficial for the future development of social sciences in these countries.

Keywords: Data archive, data services, social science, Bosnia and Herzegovina, Croatia, Serbia

Introduction

Today, knowledge is one of the most important sources of global economy growth (WB, 2012) and a key driver of a company's value (Bock, Zmud, Kim & Lee, 2005). Transmission of information, which is widely expanded as a result of Internet development and the possibility of collection, storage and dissemination of that information, has given the new possibilities for the use of knowledge. This is particularly important having in mind that knowledge has the characteristic of growing in the process of dissemination (Arzberger et al., 2011). Research has indicated (UKDA, 2002;

Corti, et al., 2011) that there has been a sharp increase in collecting data that has been used in studies of economic, political and other social issues, over the last decades. Regarding the fact that the process of collecting primary data is the most expensive and time consuming phase in the research process; establishment of national digital data archives for research data in social sciences and their integration into the standardized system for data sharing on the international level is considered as a cost savings solution (Bradić-Martinović, Zdravković, 2012).

Data collection was always a part of scientific research process. Natural sciences use data from experiments, while social sciences use various methods of primary data collection, such as questionnaires, interviews, focus groups etc. The collected data are often necessary to keep in order to check the results or to be used for further research.

The phenomena of digital data have changed the way that data are collected and preserved. For that purpose, many countries established data centers or data archives. Some data archives have a very long history, like the UCLA Social Science Data Archive, US (formed in 1961), and the United Kingdom Data Archive, UKDA, UK (formed in 1967).

In the Western Balkans region, none of the countries have digital archives in social science. Thanks to the successful implementation of the FP7 SERSCIDA project (Support for Establishment of National/Regional Social Sciences Data Archives) Bosnia and Herzegovina, Croatia, and Serbia have an opportunity to establish data services in the social sciences. It is a strategic project, designed to support cooperation and knowledge exchange between those EU countries that

are members of the Council of European Social Sciences Data Archives (CESSDA) and the Western Balkan countries in the field of social science data archiving. The project addresses the existing potential for use of information-communication technologies for the benefit of scientific research and exchange of knowledge, as laid down in the call for proposals topic. It aims to produce tangible results and improve the capacities for exchange of knowledge and data collected through research in social sciences between the European countries and Western Balkan countries involved.³

The results presented in this paper are outcomes of Work Package 2 of SERSCIDA project - Analysis of existing potentials for the establishment of social sciences digital data archive, presented at the IASSIST 2013 Conference in Cologne, Germany.

Methodology and sample

The first step of analysis was to develop appropriate methodology by local partners with assistance of CESSDA partners. To fulfill that aim we designed an online questionnaire for researchers which considered both their experience of documentation, re-use, and disseminating of research data; and also which type of statistical/analytical software packages, methodology and data they used primarily in their research. Accession to the questionnaire did not imply any restriction (academic network users, for instance) and no registration were needed.

The survey has five parts which covered: 1) Characteristics of respondents; 2) Producing data; 3) Methods of data gathering; 4) Archiving practices and preferences; and 5) Use of data and secondary analysis.

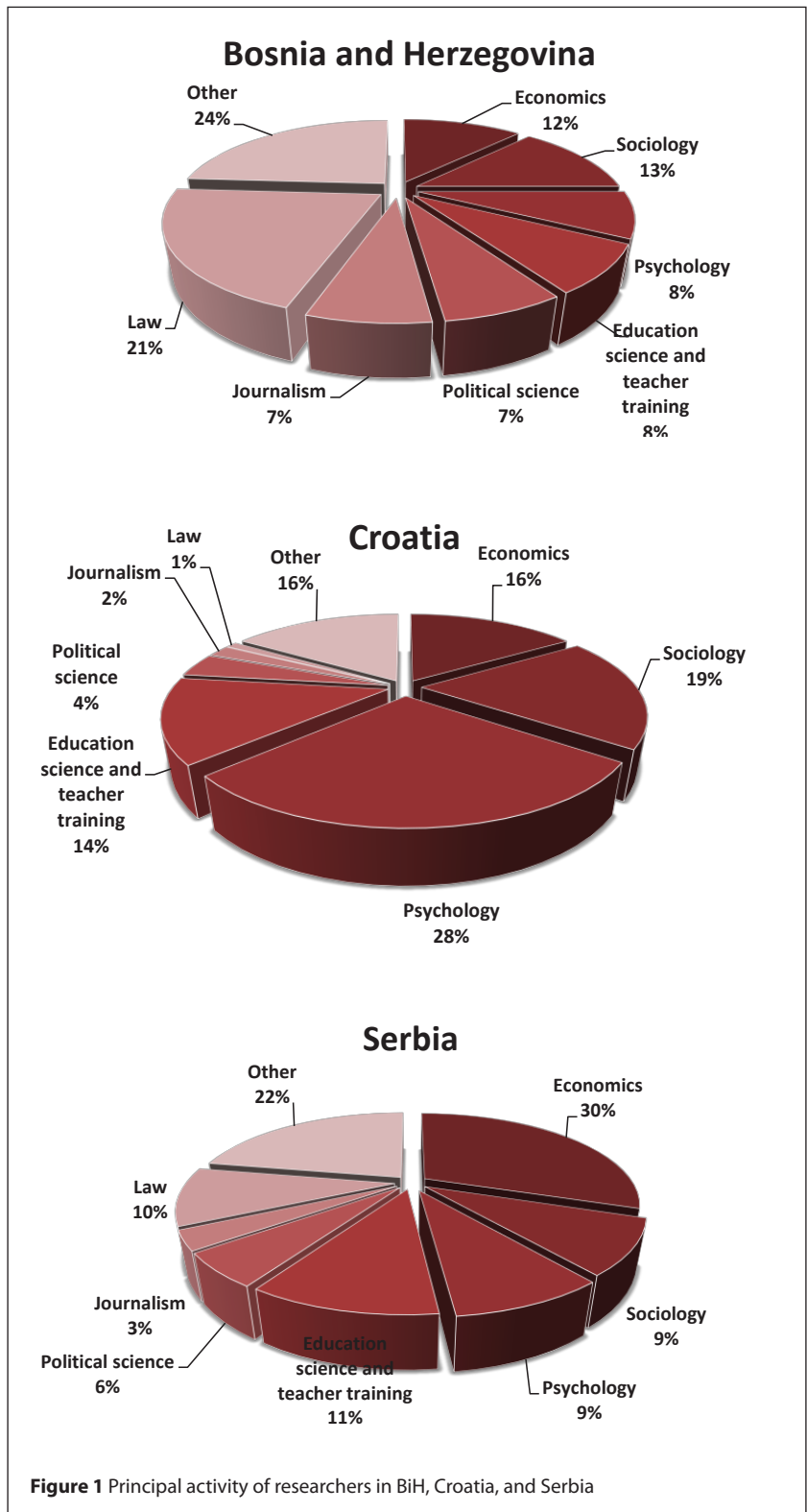
It was conducted during June and July, 2012. Bosnia and Herzegovina had 139 completed questionnaires out of 225; Croatia had 186 completed questionnaires out of 307 and Serbia had 322 completed questionnaires out of 493. The database of potential respondents was made on the basis of extensive gathering of researchers' contact addresses either from relevant government institutions or individual websites of research institutions involved in social sciences. The average rate of completed responses for all three countries was 63%.

Hereafter we selected the key questions that will shed light of the situation on this topic in Bosnia and Herzegovina, Croatia, and Serbia.

Results of the survey

Characteristics of respondents

The first few questions in our survey were aimed to make us more familiar with the basic characteristics of our respondents regarding their principal activity and research discipline. The results on principal activities are slightly different between countries. In Bosnia and Herzegovina (BiH) over a half of respondents were



undergraduate students, doctoral students, or teaching assistants and researchers or professors and in Croatia and Serbia about 80% of respondents were doctoral students or teaching assistants and researchers or professors. There are also some differences between countries regarding research discipline. According to the respondents' answers within the context of research discipline in BiH major researchers were in law science (21%), sociology (13%) and economic (12%), while other discipline like psychology, education science and teacher training, political science and journalism are below 10%. In Croatia most of the researchers are

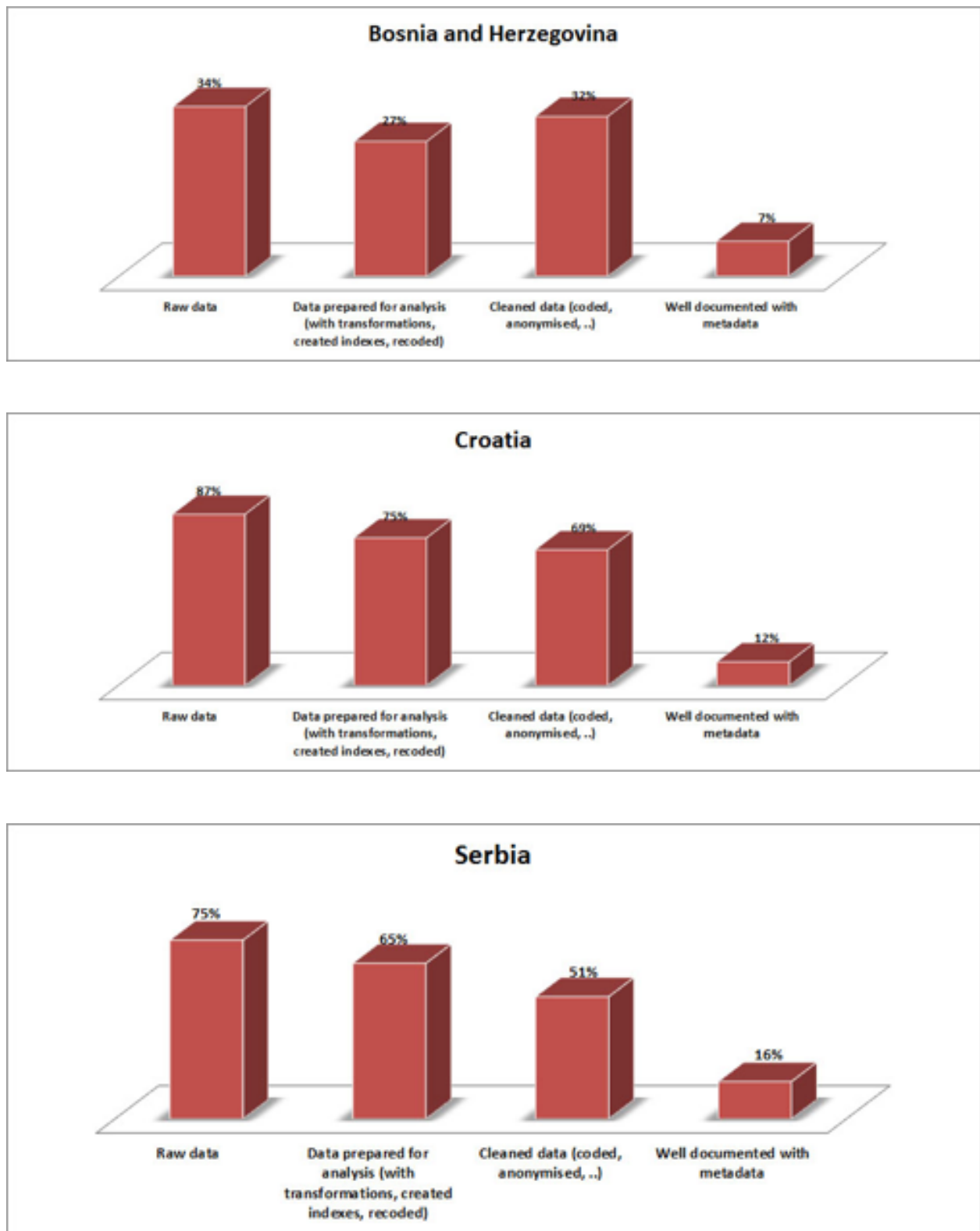


Figure 2 Type of stored data in BiH, Croatia, and Serbia

in the field of psychology (28%), sociology (19%), economics (16%) and education science and teacher training (14%). In Serbia most of the researchers are in the field of economics (30%), while other disciplines, education science and teacher training (11%), law (10%), sociology and psychology (9%) have a much smaller share. Differences in the structure of the respondents might be biased by the willingness of respondents to cooperate on collegial solidary

basis, in regard to the primary discipline in which institutions, that sent questionnaires to researchers, are engaged⁴. (see Figure 1)

Producing data

The second part of survey was dedicated to questions related to production of data and research activity within the past five years. The first question attempted to determine how many datasets were produced during that period. In each country over 50% of

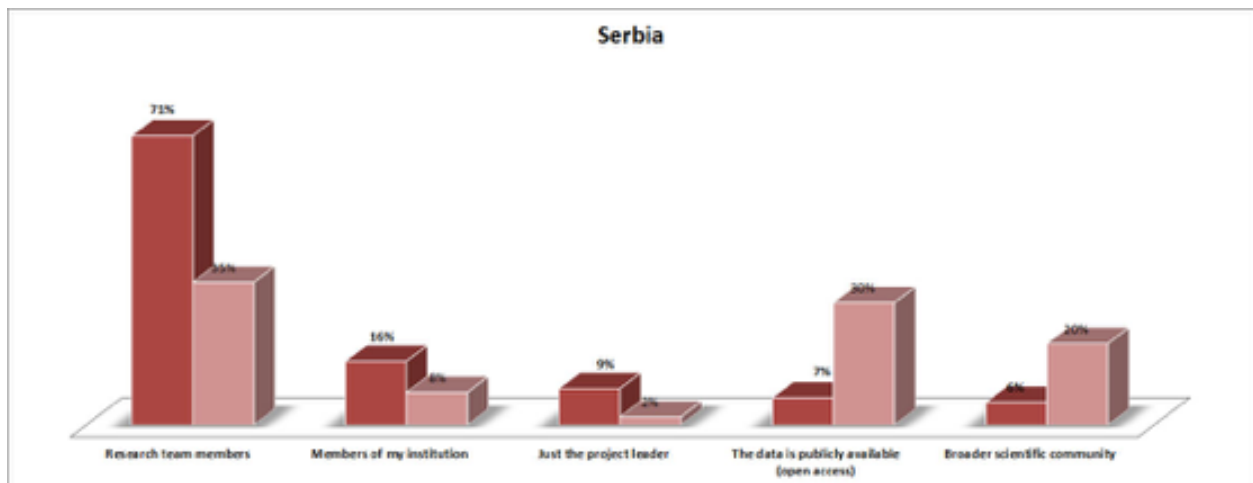
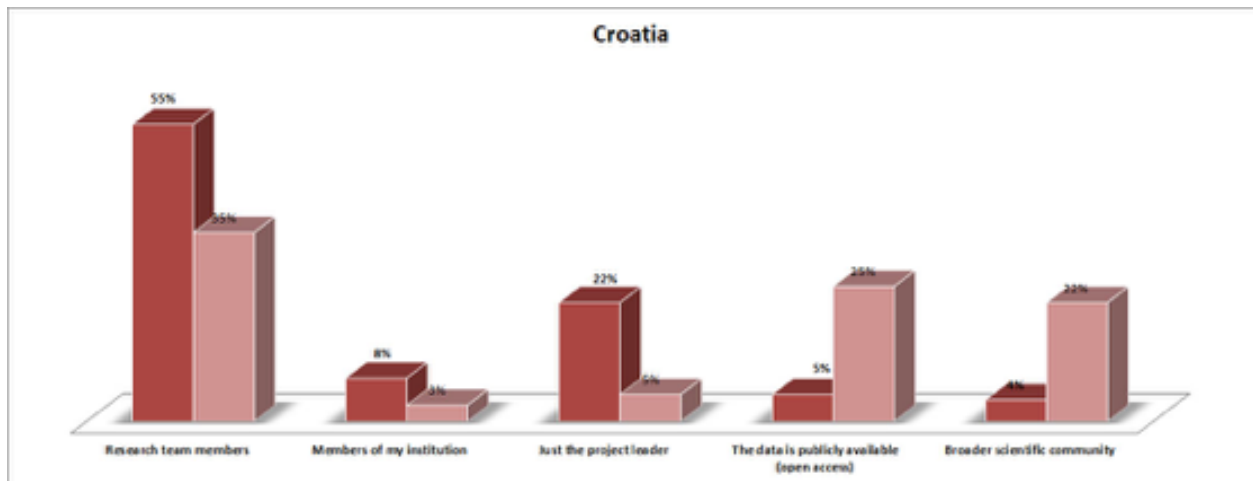
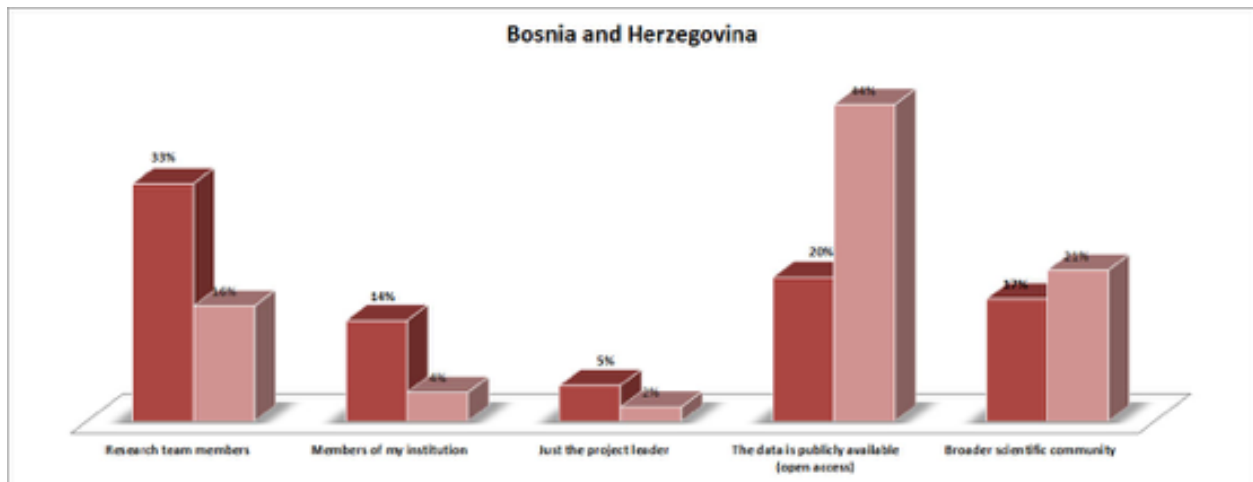


Figure 3 Current access to data vs. ideal level of access in BiH, Croatia, and Serbia (darker columns present current situation and lighter present opinions about ideal situation)

researchers confirmed that they produced five or more datasets during the past 5 years and based on this we can conclude that there is enough research potential in our countries, and that our researchers produce a substantial amount of datasets. In all countries the largest numbers of researchers have produced between 6 and 10 datasets (BiH 19%, Croatia and Serbia 24%).

But there is a slight difference between the countries in the case of a subsequent frequency; in BiH 13% of researchers produced 11-20 datasets and 12% produced 21 and more. In Croatia 19% of researchers produced 5 datasets, while only 8% produced 10-20. In Serbia the situation is similar to Croatia because 15% of researchers produced 5 datasets. Insight into the amount of research

conducted was very important because it showed that there is a sufficient number of datasets in all three countries and that the establishment of the data archive is justified, considered this criterion. Additionally, we determined that the number of datasets generated has a growing trend. The results of all three countries are almost identical. Within last five years researchers completed 45% of all datasets in 2012, 36% in 2011 and the rest in 2010 and earlier.

Methods of data gathering

Questions about applied data collection methods and the financial sources for projects were an essential part of the third segment of our survey. The question about applied data collection had open-ended answers with several offered examples (online questionnaire, structured interview, focus groups, experiment, etc.) so we received different results among the countries. Multiple answers were allowed and percentage share for each answer is defined with respect to total number of respondents (such approach is also applied in the archiving practices and preferences section). In BiH most researchers used either questionnaires (32%) or interviews (47%) in data collection in the last five years. In Croatia, the dominant method was surveys (53%) and quantitative (70%) or qualitative (33%) questionnaires (focus groups and interviews). Questionnaire (49%) was the dominant method of data collection in Serbia.

Results for the second question, about financing of research was very interesting to us. In Croatia and Serbia approximately 40% of all research was financed by public funding through national science funding bodies, while in BiH most of the research (40%) was financed through international funds/projects while only 7% had the support of national funding. International funding is also provided in the other two countries, but with a much smaller share compared to BiH. In Croatia, 17% of research was funded that way, and in Serbia 26%. The rest of the projects had been funded by institutions which conducted projects, publicly funded from other sources, private sector, and other.

The method of financing research in Croatia and Serbia is not the best possible, because the funds for the science are allocated from the budget of these countries. Therefore, these resources are often insufficient, particularly for the social sciences.

Archiving practices and preferences

The most important and interesting segment of the survey was the fourth part about existing archiving practices and preferences of researchers in three countries. Answers to the first question about type of stored data were very similar in all countries. Most of the researchers keep the data in the raw form, as data prepared for analysis (with transformations, created index, and recorded), or as cleaned data. But the main obstacle for further use is the absence of well-documented data with metadata. Detailed answers are presented in the Figure 2.

Answers to the question regarding where researcher keep data stored are also consistent between countries. The dominant number of researchers keep the data in their own computers (in average over 50%) or several copies in different computers (in average over 40%), and only a few of them keep the data in some form of institutional repository (approximately 3%). Our opinion is that the obtained result is not good for three reasons. First one is that researchers usually do not have proper procedures for backup, so the risk of losing the data is very large. Also, the other researchers in most cases do not have access to the data kept on personal computers and finally absence of well documented

data completely prevents the reuse of these data, because it is impossible to find them.

During the design of our survey we assumed the answer on the previous question and the related problems. Therefore, we allowed respondents to provide comparative answers to questions about their current level of access to data and what would be the ideal level of access, according to their opinion. In Figure 3 we present comparative values obtained on two questions. It is very interesting to see that in all three countries current access to the data is dominantly limited to the research team; however most researchers think that data should be publicly available (open access) or at least available to the broader scientific community. These results are very encouraging for us regarding establishment of digital data archives in our countries.

The last question in this segment was intended to provide insight on the willingness of researchers to provide research data to an archive, if the data would be safely preserved and access regulated. The great majority of researchers in all three countries want to provide research data to archive if the data would be safe with regulated access because 45% of them (in average) answered with Yes, certainly and 40% answered with Yes, probably. These responses are very encouraging, because they indicate a justification for the establishment of digital archives in BiH, Croatia, and Serbia, as well as a positive attitude within the research community about the future deposits and secondary use of data. Nevertheless, we are fully aware that the number of positive answers is probably higher than the real disposition of overall population of social scientists, primarily due to self-selection of survey participants based on their interest in the subject of research. Nonetheless, a data-archival institution would obviously address the reported existing needs and help in overcoming the current issues identified with respect to safe archiving and enabling of access to research data.

Use of data and secondary analysis

The last part of our survey was conducted to understand the practices of researchers in the field of secondary analysis and use of data. More than half of the respondents (BiH 75%, Croatia 51%, Serbia 64%), have stated that the sharing of research data is very important in their discipline and only 2% (on average) find it not very important.

The answers to the question "Would your scientific work benefit if you had better access to research data produced locally or internationally?" were expected. We offered two modalities for this question (for local and international research data). In Bosnia and Serbia the great majority of researchers stated "Yes, considerably" as the predominant answer for both types of research data, while Croatian researchers considered that their scientific work would benefit more considerably from better access to international than local research data. Detailed responses are presented in the Figure 4. We anticipated these answers due to the existing problems in funding of scientific work as the researchers are always faced with a lack of data.

The final question assured us of the great possibility for establishing data archives in BiH, Croatia, and Serbia, because over 50% of researchers consider it as very useful and only 1% as not useful at all. But regarding these answers we must be very careful. The researchers in these countries do not have appropriate knowledge about data archives, just general information about them or

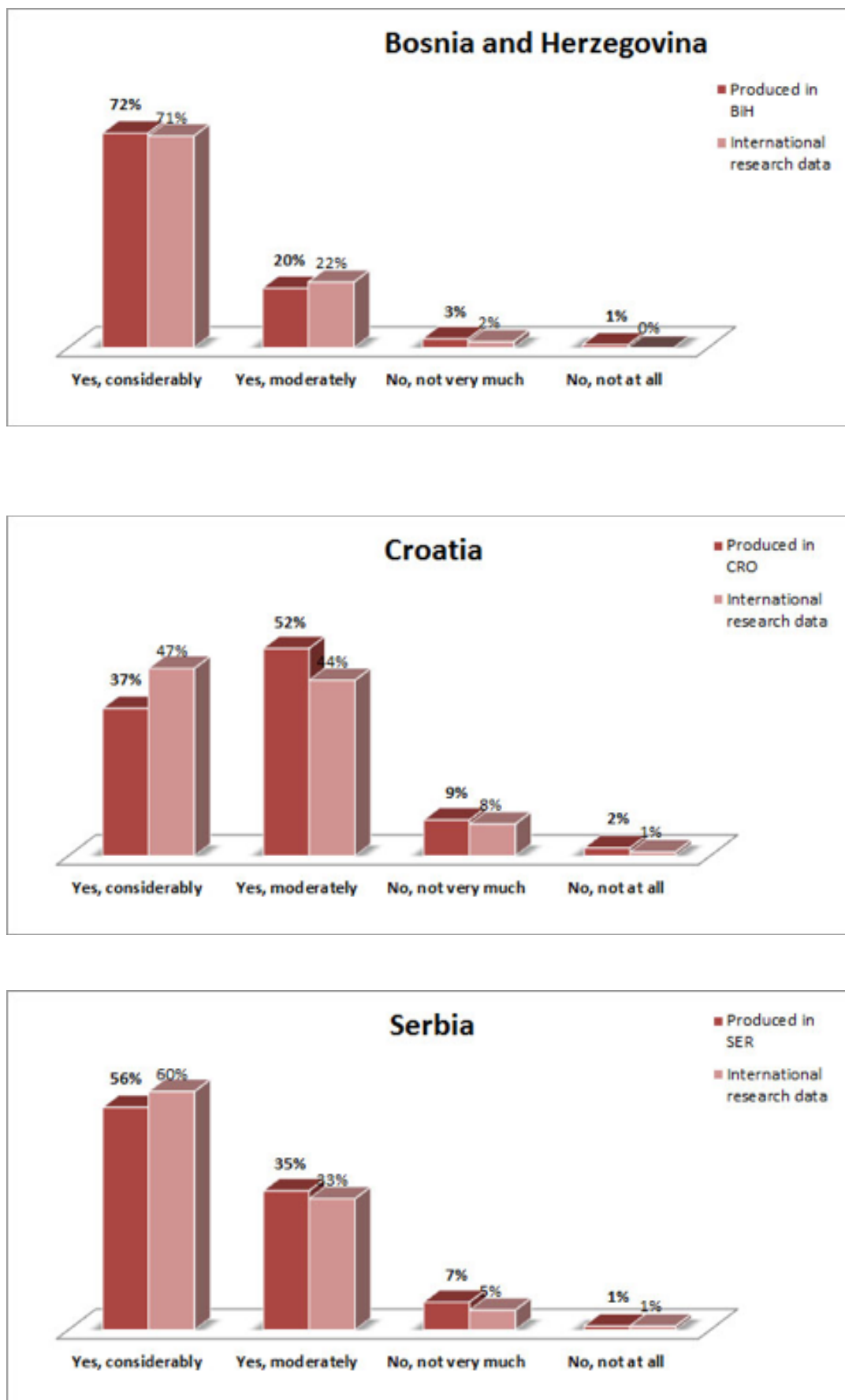


Figure 4 Better access to data as benefit for scientific work

absence of any knowledge. We assume that the situation will change by raising awareness of this issue in the near future.

Conclusion

In this paper we report findings of a survey conducted in Bosnia and Herzegovina, Croatia, and Serbia based on standardized questionnaires for all three countries on a representative sample, aimed to shed light on existing potentials for the establishment of social sciences digital data archives. We analyzed numerous single issues grouped into four general topics comprised by the survey—production of primary data, methods of data gathering, archiving practices and preferences and use of data (secondary analysis), with particular interests in scope and quality of data production, documentation, and dissemination. Similarity in the structure of responses among countries allowed us to easily generalize our conclusions across general topics, despite expected variations in responses on the level of particular issues.

Results of the survey related to issues of data production are quite promising, as researchers in social sciences in all three countries have been considerably active in recent years producing between 6 and 10 datasets, and what is more important, that scope of data production has a growing trend. However, the situation is not so bright when comes to the issues of data documentation and especially data dissemination. Most researchers keep data either in raw form or partially prepared for further analysis, but without appropriate documentation with metadata in accordance to international standards. The absence of appropriate documentation is identified as the main obstacle for further use of data in secondary analysis. Therefore, it is not surprising that researchers mostly keep data in their own computers (in average over

50%) or several copies in different computers with an approach limited to research teams. It is encouraging that researchers are willing to share their data and are also aware of the benefits that centralized data archive could bring about. Thus, we can conclude that potentials for establishment of digital data archives in considered countries are large, regarding the scope of data produced in social sciences and the positive attitude of researchers toward data sharing and benefits of data archives. In addition, existence of national data archives, with staff trained in assisting researchers in the process of data documentation and preservation according to international standards, will remove key obstacle for data sharing and further use of data in secondary analysis.

References

1. Arzberger, P., Schroeder, P., Beaulieu, A., Bowker, G., Casey, K., Laaksonen, L., Moorman, D., Uhlir, P., Wouters, P. (2004). Promoting Access to Public Research Data for Scientific, Economic, and Social Development. *Data Science Journal*, 3 (November), pp.135-152.
2. Bock, G., Zmund, R., Kim, Y., Lee, J. (2005). Behavioral intention formation Knowledge Sharing: Examining the Roles of Extrinsic Motivators, Social-Psychological Forces, and Organizational Climate. *Management Information Systems Quarterly*, 29(1), pp. 87-111.
3. Bradić-Martinović, A., Zdravković, A. (2012). Integration of Western Balkan Countries into the European System of Digital Data Archives in Social Sciences: Case of Serbia. *Review of Applied Socio-Economic Research*. vol. 4, issue 2, p.p. 32-41
4. Corti, L., Van den Eynden, V., Bishop, L., Morgan-Brett, B. (2011). Managing and sharing data. UK Data Archive, Colchester, Essex.
5. UKDA, (2002). Preserving and Sharing Statistical Material, The Royal Statistical Society & the UK Data Archive, University of Essex.
6. World Bank, Knowledge for Development – K4D, <http://web.worldbank.org/WBSITE/EXTERNAL/WBI/WBIPROGRAMS/KFDLP/0,,contentMDK:20269026~menuPK:461205~pagePK:64156158~piPK:64152884~theSitePK:461198,00.html> (last visit 05 May 2014)

Notes

1. Aleksandra Bradić-Martinović, PhD is Research Fellow in the Institute of Economic Sciences in Belgrade, Serbia in the fields of business information systems, e-banking and data management. She is also engaged in teaching as an associate professor in Belgrade banking academy in the same fields and can be reached by email: abmartinovic@ien.bg.ac.rs.
2. Aleksandar Zdravković, MA is Research Associate in the Institute of Economic Sciences in Belgrade, Serbia in the field of econometrics and macroeconomics. He is also a PhD student at Economic Faculty University of Ljubljana and can be reached by email: aleksandar.zdravkovic@ien.bg.ac.rs.
3. The more detailed information about SERSCIDA project can be found at www.serscida.eu.
4. Survey was conducted by three institutions, local participants in SERSCIDA project. In BiH it was Human Right Centre, University of Sarajevo. In Croatia it was Faculty of Humanities and Social Sciences, University of Zagreb and in Serbia it was Institute of Economic Sciences, Belgrade.