STUDY ON THE MAPPING OF RESEARCH DATA IN THE REPUBLIC OF MOLDOVA

IN THE CONTEXT OF OPEN SCIENCE

Nelly Turcan[1], Andrei Rusu[2], Rodica Cujba[3],

[1] *Univ. Prof., Information Society Development Institute, Moldovan State University; Moldova*

[2]*Assoc. Prof., Institute of Mathematics and Computer Science „Vladimir Andrunachievici", Moldova*

[3] *Sc. Res., Information Society Development Institute, Technical University of Moldova, Moldova*

Abstract

The Open Science concept represents a new approach to the way in which scientific research based on cooperation and new ways of knowledge dissemination is carried out and organized, using new digital technologies, new tools for collaboration, and R&D infrastructure to ensure open access to research data.

This study uses data collected in May - July 2018 within a survey that aimed at investigating the scientific data ecosystem in the Republic of Moldova. Findings show that, although there are some concerns about the loss of property rights and copyright infringement in case of sharing and open access to research data, Moldovan academia is ready to provide access to research data. The research has highlighted that a new challenge is needed to solve scientific data issues by creating new type of infrastructure to ensure data retention, broad access to research results for the purpose of their dissemination and use, and creating new research opportunities based on research data.

Keywords: open science, open research data, e-infrastructure, Republic of Moldova

---

1 Introduction

The speed of progress in science has always been dependent on how effectively scientists can communicate their results to colleagues but also to those who want to implement these results in new technologies and practices.

Currently, we are witnessing an important change in what we call science in terms of organizing, conducting, evaluating, using and disseminating research results. This change that reflects the term "Open Science" is determined by the development of new technologies, increasing social role of scientific

research, the current political and institutional context (Cuciureanu, 2018, p. 15).

According to European Commission Open Science is about the way research is carried out, disseminated, deployed and transformed by digital tools, networks and media (European Commission, 2018). The Organization for Economic Co-operation and Development mentions that Open science commonly "refers to efforts by researchers, governments, research funding agencies or the scientific community itself to make the primary outputs of publicly funded research results – publications and the research data – publicly accessible in digital format with no or minimal restriction as a means for accelerating research; these efforts are in the interest of enhancing transparency and collaboration, and fostering innovation" (OECD, 2015, p. 7).

Open Science brings social, economic, cultural, political and technological change, based on openness and connectivity, on how research is designed, performed, used, assessed, and preserved. Open access platforms, open infrastructures, open data tools, open educational resources, open evaluation methods, open collaboration, or open citizen science activities are irreversible trends that are impacting all scientific actors and have the potential to accelerate the research cycle (Vicente-Saez and Martinez-Fuentes, 2018). By increasing access to publications and data, universities, research institutions, firms and individuals may use and re-use scientific outputs to produce new products and services.

One of the key elements of open science is open access to research data. Open research data is research data that combines the characteristics of open data and the types of research data (Onyancha, 2016).

As World Bank mentioned in his report (World Bank, 2018, ch.1) data is considered the new gold, or the new oil, and like oil, unprocessed data has relatively little value and needs to be mined, refined, stored, and sold on to create value.

In a research context, there is a growing opinion that most research data should be open, particularly data from publicly funded projects. This point of view is driven especially by research funder requirements for sharing and re-use data, upon principles regarding research data such as to be findable, accessible, interoperable and re-usable (FAIR principles). The research data are made open for two purposes: to provide evidence that the research was conducted properly and to provide data for reuse and the generation of further findings and outputs (Childs et al., 2014).

The updated Directive on Open Data and Public Sector Information obliges Member States to "support the availability of research data by adopting national policies and relevant actions aiming at making publicly funded research data openly available ('open access policies') following the principle of open by default and compatible with FAIR principles" (European Parliament, 2019).

Scientists are particularly interested in data collection, and the success of each experiment is determined by the new data generated, which can contribute to

advancing scientific knowledge. Any scientific research involves performing an observation, generating a hypothesis, running an experiment, and collecting data. Traditionally, for any research, the amount of data collected by scientists was not very extensive, and its analysis did not require the use of technology. Previously, scientists used technology in a very limited way, and data evaluation was not done using algorithms or software. However, significant changes have taken place over the past two decades, and changes in software and tools have made data acquisition and analysis a very important part of research.

2 Opening science in the Republic of Moldova

At present, in many countries, there are various actions to promote the transparency of governance; promoting dialogue between governance and citizens; the use of new technologies in governance and in dialogue with citizens. Opening and sharing public data and information is an essential policy in building a more open, responsible and efficient government. Open, democratic and transparent science could be a major factor for any government in the progress of a country.

Opening science requires a new systemic approach, especially in nationally and internationally agreed strategies and policies.

Special strategies and policies dedicated to Open Science are not yet developed in the Republic of Moldova. According to the national legislation and the number of open access policies approved at national and institutional level, the Republic of Moldova cannot be called a politically open territory. However, there are some encouraging examples, as well as a positive forecast for the future to support open access and open science in the Republic of Moldova.

Promoting and implementing of Open Science policies at national level is necessary for at least three reasons (Gh. Cuciureanu et al., 2018, p. 16):

a) the transition to Open Science is an official policy of the European Union and the Republic of Moldova declared its intention to integrate into the European Research Area;

b) the Open Science concept radically changes the way to do research, and its lack of implementation in the Republic of Moldova will make non-competitive native science;

c) the legislation of the Republic of Moldova provides for certain elements of Open Science (even if not consolidated) that have to be implemented.

In the Republic of Moldova, the state policy in the field of science and innovation is carried out under the Code on Science and Innovation of the Republic of Moldova (Parliament of the Republic of Moldova, 2004). Several articles of the Code assure guaranteed access to scientific information. In accordance with the current legislation, the State guarantees: support through access to information, through its dissemination; information assurance of the

topics from the scientific and innovation field; free and non-discriminatory access to scientific-technological information resources. However, at state level open access to scientific research is not confirmed in the Republic of Moldova, primarily that financed from the public money, and besides this, the mechanisms for ensuring free and open access to the scientific and technological information resources and open access to research data are not specified.

In order to support open access and open research data there have been launched several projects in the Republic of Moldova. This project focuses on improving the quality of academic studies which also focus on open and free access to scientific information and data.

3 Mapping of research data in the Republic of Moldova

In order to map the situation regarding generation, gathering, use, sharing and preservation of research data obtained within research projects carried out in the Republic of Moldova, Information Society Development Institute conducted a survey in May-July 2018.

The survey on the mapping of the research data ecosystem in the Republic of Moldova was carried out within the framework of the project *Development of the conceptual and methodological framework for data e-Infrastructure in the field of research, development and innovation in the Republic of Moldova* (e-IDSM) https://idsi.md/en/e-idsm). Unlike the previous survey this one was focused exclusively on research data. The main goal of this survey was to identify the needs of the RDI community in the Republic of Moldova on the management of scientific data over their lifecycle (creation, processing / analysis, storage / preservation, sharing / access and use). The specific objectives of the survey were:

- to identify the types / formats and sources of research data;
- to find out the modes of storing and preservation of research data;
- to discover the ways research data are processed and analyzed;
- to learn the procedures of research data management;
- to determine methods of sharing, access and use of research data.

3.1 Methodology

The survey covered questions regarding the entire lifecycle of research data. It included five sections:

- Creation and storage of research data;
- Storage and preservation of research data;

- Data processing and analysis;

- Data sharing;

- Research data management.

Responses were collected from 48 RDI institutions (92% success rate), including 13 higher education institutions. Respondents with various positions within these institutions participated in the survey, including: 25 heads of RDI institutions (12.3%); 42 project managers (20.7%), 65 laboratory / research group managers (32%), 34 scientists (16.7%), 23 university teachers (11.3%), 4 PhD students (2%), other positions (5%).

The authors were members of the research projects' teams and have participated in the design of the surveys, collection and aggregation of the results.

3.2 Results and Discussions

The data obtained in this survey allows us to find out that Moldovan academia produces and generates different kinds of research data Most respondents of the survey mentioned that they produce or generate the following types of data for research: text (86.2%), images (66%), numerical data (62.6%), tabular data (55.7%) (Figure 1). Only 13.8% of respondents do not produce or generate any type of research data. Other types of data were also mentioned, including: DICOM images, archive quotes, technical drawings, protocols, algorithms, programs, maps, etc.

Concerning the format of the data generated or collected, the survey participants indicated that they use different formats for scientific data (Figure 2). Respondents have specified that they use all categories of formats specified in the questionnaire. However, research data is mainly generated and collected in the following formats: text (93.1%), presentations (83.2%), graphics (67%), calculation sheets (46.8%), data bases (43.3%) and statistical software files (25.6%).

Taking into account data formats used by researchers, as well as the needs described by them in the survey, Moldovan researchers can be divided into two categories:

- Researchers using relatively widespread software tools in the academic and research environment, such as: Microsoft Office, SPSS, Adobe FineReader, and others.

- Researchers using research-specific software solutions such as: ArGIS, Geoportal, Mathematica, FoxPro, Endnote, 1C, EViews, GAMESS, Gaussian09 and others.

Scientific data can be obtained or generated as a result of research activities as well as from various sources. Survey participants noted in particular the following ways of production and generation research data:

- results of the experiments (69.3%),

- results of the observations (72.3%),

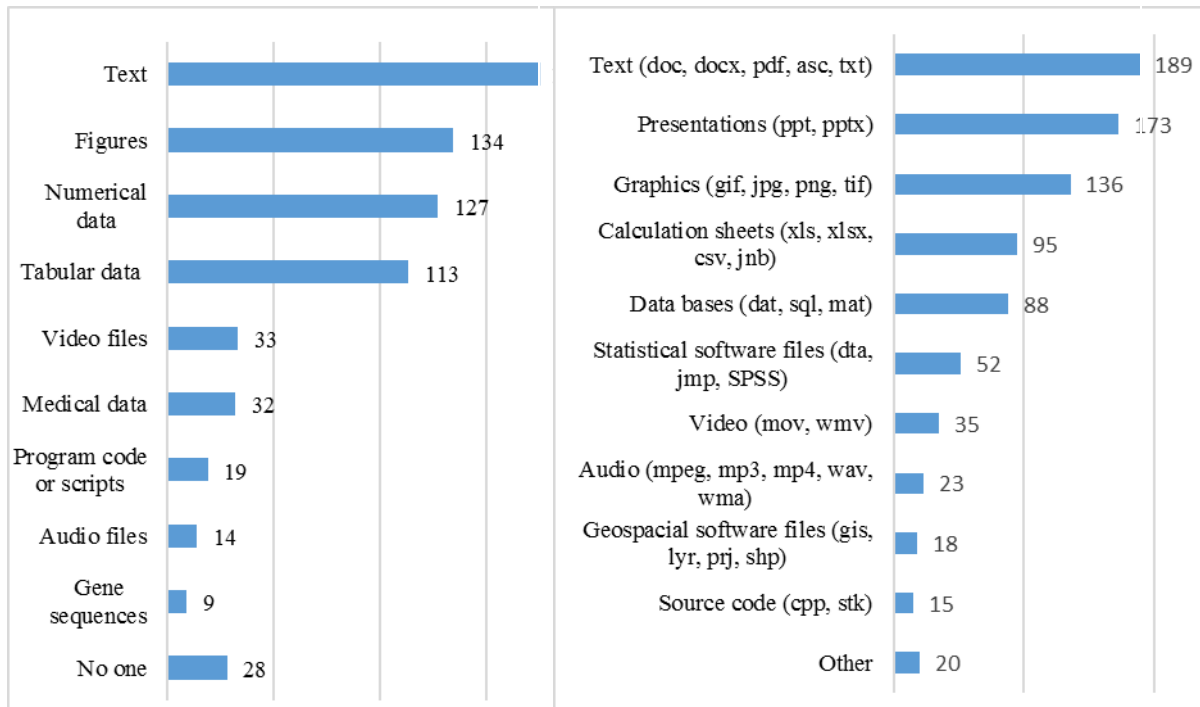- scientific publications (72.8%), statistical sources (52%),

- survey results (33.7%).

| Figure 1. Types of digital data produced / generated for research | Figure 2. Types of generated / collected data formats |

For some areas of research, data can be obtained from social media, sensor data, archive files, demographic forecasts, bilateral and international projects, media, weather station data, medical images, etc. (Figure 3).

Sources of input data used in the research process largely coincide with the mode in which scientific data are produced and generated (Figure 4). Thus, experiment data (74.3%) is the main source for the research data collection. Also, important sources of extraction, collection and input of research data are data from studies and surveys (66.3%), official national statistics (48.5%), international statistics (45%), public data sets (35.1%). At the same time, respondents noted analytical / theoretical results, archive data, social media, mass media, clinical data as inputs of data used in the research process etc.

Procedures for storing and archiving research data are very important not only for long-term preservation of research results but also for ensuring the integrity of these data. Thus, 102 (50.2%) of the respondents mentioned that they take steps to preserve research data, 61 (30%) of respondents said they did not take measures to preserve the data, and 40 (19.7%) of respondents

said that they do not know if data preservation measures are under way in the respective research projects. Among the specified methods of storing research data, the researchers indicated two large categories (Figure 5):

- locally, on the personal computer (96.6%) and / or on physical support, i.e. on paper (72.9%), CD, USB, or external HDD (63.5%) etc.;

- online in databases, specialized institutional repositories or research laboratories (24.6%), the institutional computer network (32%) and / or internationally, Dropbox, Google Drive (30.5%), etc.
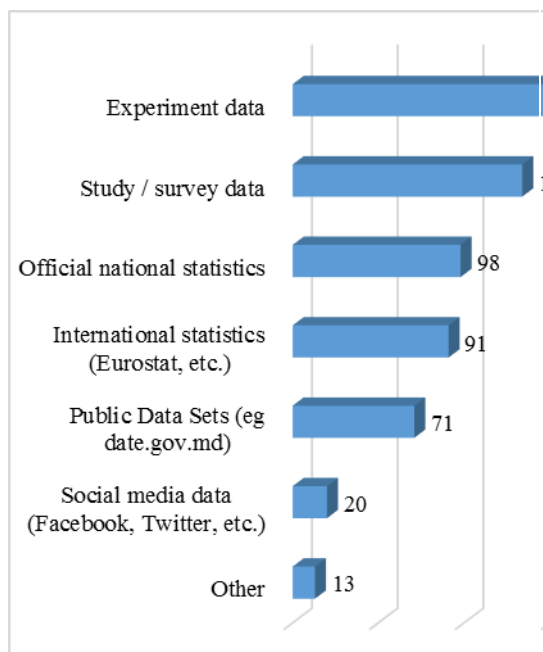
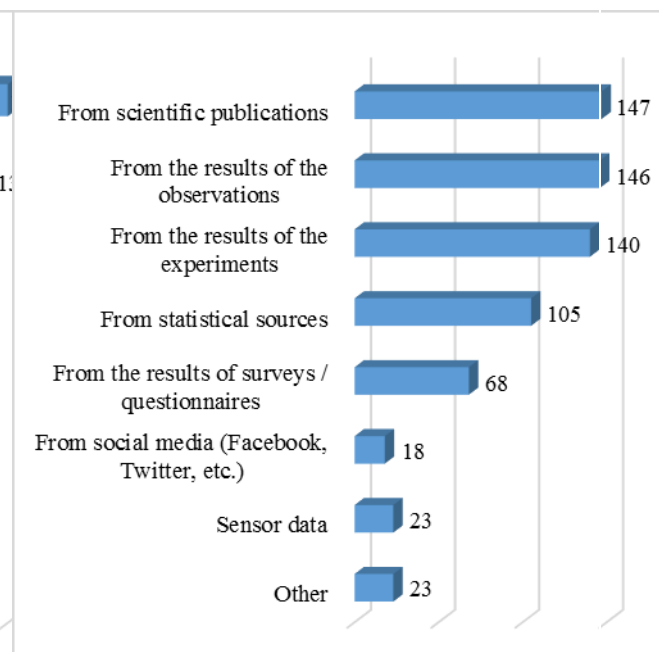Figure 3. Sources of input data used in the research process

Figure 4. The ways the scientific data is obtained (produced / generated)

It should be noted that many respondents have specified several ways of preserving research data, which contributes to the safety of data retention. However, 30% of respondents do not take measures to preserve research data, and 19.7% do not know if the institution or laboratory is taking such measures.

Respondents noted that they protect research data by limited access to research data (66%), password (63.1%), backups (34%), data anonymization (22.2%), data coding (11.8%), etc. More than 10% of respondents do not protect research data, and 3% of them destroy data after use.
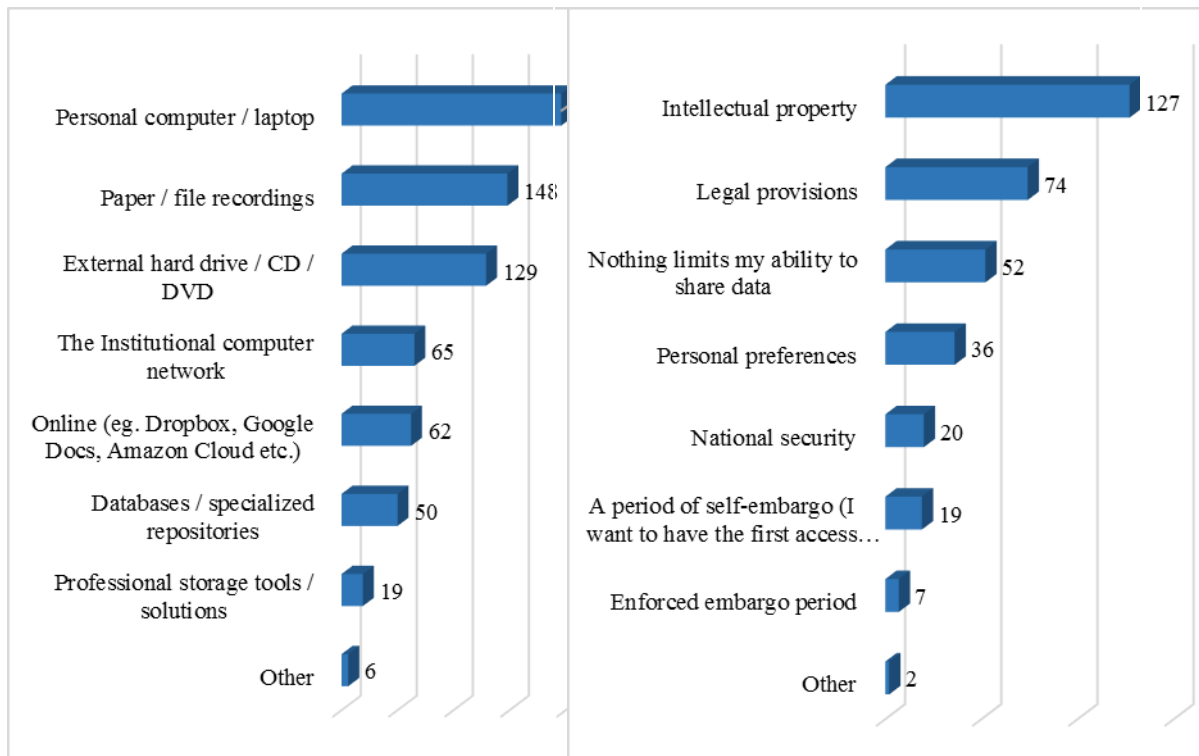
Figure 5. Methods of storing research data

Figure 6. Restrictions that limit the ability to share data

Currently, when scientific research is highly interdisciplinary and collaborative, it is necessary to share and use data interoperability procedures. Thus, 61 (31%) of the respondents replied that they share research data, while 122 (60.1%) share them according to the research project policy. Researchers also noted that there are some restrictions limiting the sharing of research data (Figure 6), such as intellectual property (62.6%), legal provisions (36.5%), national security (9.9%), embargo or self-embargo (12.8%). About one-fourth of respondents consider that nothing limits data sharing (25.6%), and 17.7% of those surveyed noted that data sharing limits depend on personal preferences.

Research data is usually documented using metadata. 76 respondents (37.4%) noted that they document or record certain metadata for scientific data or data sets, while the majority of respondents (62.6%) do not. Only 35 respondents (21.1%) use metadata standards when recording or documenting research data. Among the metadata standards used were mentioned CERIF, Diagnostic Classifier CIM10/HL7, Archival Standards, BibTeX, standards of National Bureau of Statistics, ISO 28258, ISO 11074 and ISO 15903, GenBank, CNAS, EUROSTAT, EIOPA, DICOM, etc.

It is important to have Data Management Plans for institutional data management policies or procedures as well as for research projects. The results of the survey revealed that 107 respondents, which make up more than 50% of the survey participants, do not know or believe that there are no institutional policies and procedures regarding the management of research

data (Figure 7). With regard to the development of the data management plan for research projects, only 21 respondents (10.3%) stated that funding agencies had requested such a plan (Figure 8).

Only 96 respondents (47.3%) know about the existence of institutional data management policies and procedures. They noted that there are various policies and procedures in place to protect, store, archive, share research data, among which: privacy policy, data storage policy, institutional policy on open access, institutional policy on intellectual property and technology transfer, primary data verification policy, old data removal policy, strict journaling of records, experiments and tests registries, contracts with organizations, non-disclosure agreements, internal networks specifying data access rights, etc.
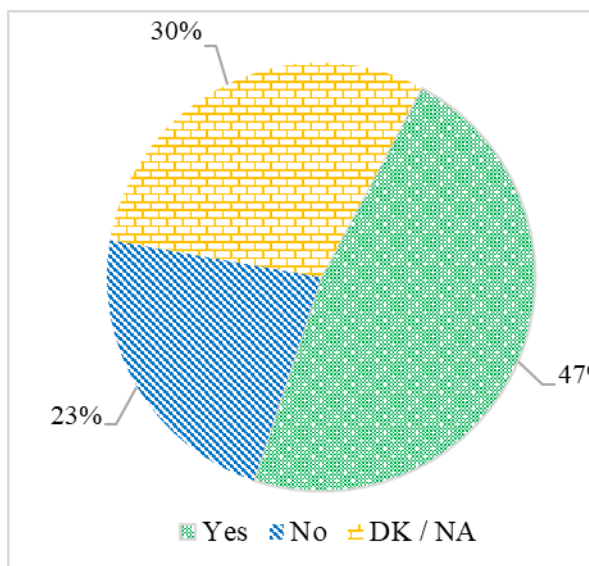

Figure 7. Existence of institutional management data policies and procedures
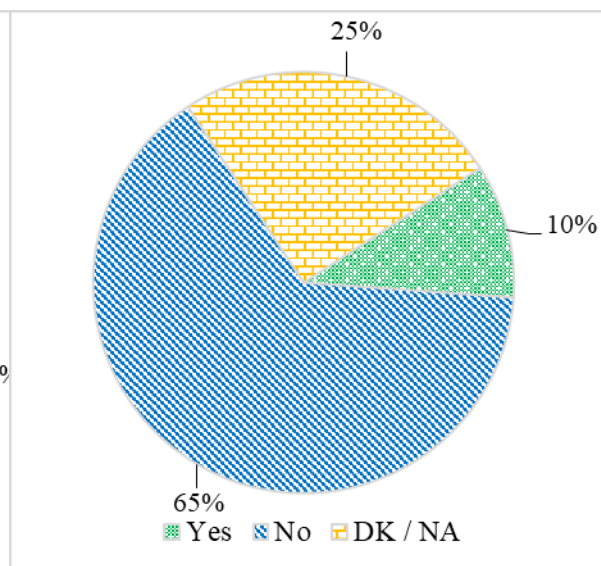
Figure 8. Request from financial agencies of Data Management Plan

Respondents noted that there are a number of issues they face in managing research data. The most common problems are:

- non-standard file formats with processing difficulties (34,5%),

- no financial resources are available to ensure data management (33%),

- difficulties to find files that have been developed themselves or by their colleagues, i.e. file versions, map structure, etc. (28,6%),

- searching for or accessing the scientific data of former colleagues, i.e. former PhD student, ex-employee (26,6%),

- identifying the storage location of data files, i.e. external hard drive, USB, DVD / CD, network (19,7%),

- there is no institution-based data management tools or solutions (19,2%), security and data file protection (15,8%),

- problems with setting up the data owner (11,3%),

- legal issues as a result of international data transfer (7,4%) etc.

However, it should be noted that the majority of survey participants (170 respondents – 83.7%) believes that training on research data management is needed (Figure 9). Respondents emphasized the necessity of training researchers from different fields on research data management technologies.

The survey highlighted the need to improve the circulation of knowledge and access to research data. Thus, 31% of survey participants noted that they unconditionally share research data, 60.1% said they share the research data according to the conditions specified in the research project, only 14.3% do not share their research data.

Survey participants were asked what they would choose in case they would share or plan to share research data. Most respondents noted that they will present data to journal as a support for the publication (64%). Also, other data sharing options have been identified, such as sharing the research data to colleagues on request or informally (39,9%), the data will be available online on the project or institution site (34,5%), storing data in a specialized database or repository (33,5%) etc.

This survey provided questions on open access to research data resulting from public funding (Figure 10). The majority of researchers (177 respondents) believe it is necessary to open the data resulting from public money-funded research. 57 respondents (28.2%) opted for unconditional open access, 29 respondents (14.4) were granted open access after an embargo, and access under certain conditions was supported by 91 participants' survey (45%). Only 25 researchers (12.4%) do not support the opening of data from state funded research.

At the same time, the survey participants have specified that research data must be accessible to colleagues, scientific community, PhD students, decision-makers, educational institutions and other users, and one of the primary conditions for using research data is to cite the source. Also, it was mentioned that there is no mechanism for managing and coordinating international projects in the Republic of Moldova, some data banks are not accessible to the public, and researchers do not have sufficient skills in managing research data.
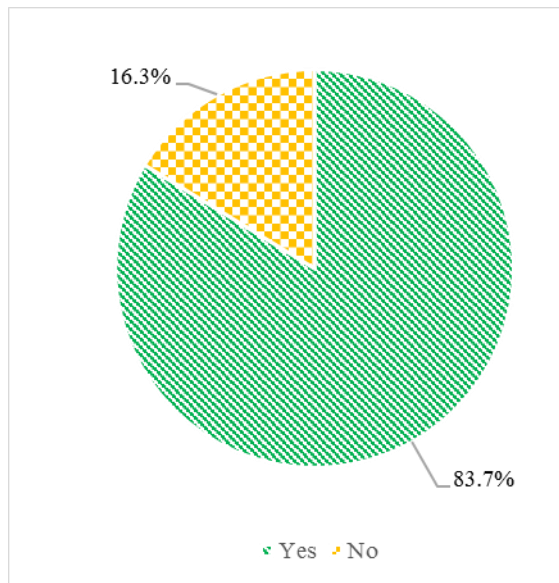
Figure 9. Moldovan academia's opinion on the necessity of the training on research data management
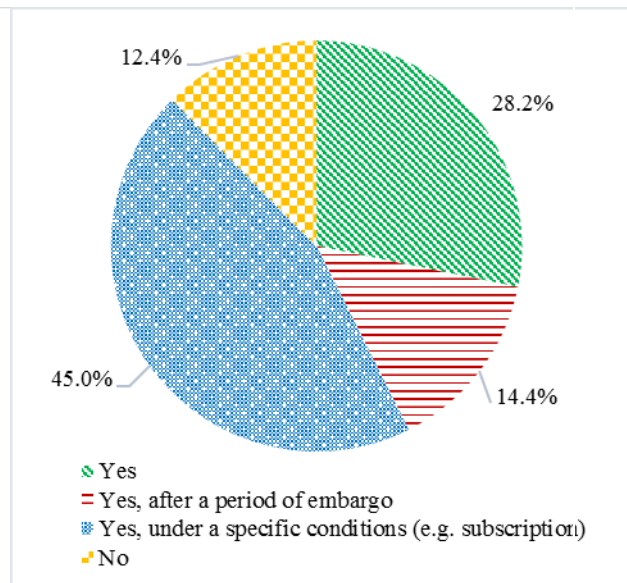
Figure 10. Open Access to research data resulting from public funding

4 Conclusions and Recommendations

In conclusion, we can mention that Moldovan academia is ready to provide access to research data. Most of researchers use digital media to access research data, but prefer to use traditional data storage formats (Word, Excel, PDF, etc.). Only some researchers use modern and innovative tools to process, access, store and archive research data. One of the main issues that discourages research data sharing is the issue of copyright protection. There are some concerns about the loss of property rights and copyright infringement in case of sharing and open access to research data. At the same time, in RDI institutions of the Republic of Moldova, the management of research data management is not implemented. There are problems related to long-term preservation, storage, sharing and access to research data.

Concluding results of these studies, the following recommendations can be made:

- Continuous analysis of the tendencies in research data management.

- Analysis of the international experience in the field of research data management.

- Establishment and approval of policies on research data management in research projects and / or research institutions.

- Training and familiarization of Moldovan academia in different fields of research data management.

- Training researchers in the Republic of Moldova on open source solutions that could be used in the research process as well as in the special case of research data management.

- Establishing rules / procedures / customs for research data management to be known to researchers, to be adopted by researchers and addressing all stages of research data management.

- Organization of round tables or other activities for presentation of solutions used in the field, as well as for exchange of views and experience in the field of research data management.

Acknowlegments

References

[1] European Commission. *Open Science*, Last update: 24 April 2018. Available at: https://ec.europa.eu/digital-single-market/en/open-science

[2] European Parliament. *European Parliament legislative resolution of 4 April 2019 on the proposal for a directive of the European Parliament and of the Council on the re-use of public sector information*, Brussels, 4 April, 2019. Available at: www.europarl.europa.eu/doceo/document/TA-8-2019-0352_EN.pdf

[3] Gh. Cuciureanu et al. *Open Science in the Republic of Moldova: Study [Ştiinţa Deschisă în Republica Moldova: Studiu] (in Romanian)*. Institutul de Dezvoltarea Societăţii Informaţionale, Chişinău, 2018.

[4] O.B. Onyancha. "Open Research Data in Sub-Saharan Africa: A Bibliometric Study Using the Data Citation Index". *Publishing Research Quarterly*, 32(3): 227-246, 2016. Available at: https://doi.org/10.1007/s12109-016-9463-6.

[5] OECD. "Making Open Science a Reality. OECD Science". *Technology and Industry Policy Papers*, 25. Publishing, Paris, 2015. Available at: http://dx.doi.org/10.1787/23074957.

[6] Parliament of the Republic of Moldova. *The Code on Science and Innovation of the Republic of Moldova [Codul cu privire la ştiinţă şi inovare al Republicii Moldova] (in Romanian)*. Nr. 259, 15.07.2004. Available at: http://lex.justice.md/md/286236/.

[7] R. Vicente-Saez and C. Martinez-Fuentes. "Open Science now: A systematic literature review for an integrated definition". *Journal of Business Research*, 88: 428-436, 2018. Available at: https://doi.org/10.1016/j.jbusres.2017.12.043.

[8] S. Childs, J. McLeod CHILDS, E. Lomas, G. Cook. "Opening research data: issues and opportunities". *Records Management Journal*, 24(2): 142-162, 2014. Available at: https://doi.org/10.1108/RMJ-01-2014-0005.

[9] World Bank. Information and Communications for Development 2018: Data-Driven Development, 2018. Available at: https://doi.org/10.1596/978-1-4648-1325-2.