The future of science – Open Science and Open Data

UDC 316.74:001]:004.738.5
DOI 10.18485/infotheca.2020.20.1_2.3

ABSTRACT: The development of ICT brings many changes in different fields of life –health care, education, economy, science. Opening of the resources that could be helpful to the society becomes a priority in the developed countries. With digital technologies and the Internet, Open Science is at its highest level. This term refers to opening of the scientific results, so the academic community can use it as a support in some further scientific research. The countries of the European Union are united in the process of adopting the legalisation about Open Science, as well as the non - European Union countries. Open Science is an umbrella term which includes Open Reviews, Open Access, Open Data etc. This paper presents definitions and benefits of Open Science, as well as the possibilities of access and use of Open Data, one of the most current representative of this movement.
KEYWORDS: Open Science,Open Data,Open Access,digital archives.

PAPER SUBMITTED: 17 June 2020
PAPER ACCEPTED: 04 November 2020

Jelena Banović
jelena.banovic@ien.bg.ac.rs

Institute of Economic Sciences
Data Center Serbia
for Social Sciences
Belgrade, Serbia

1 Introduction

The possibility of access to results of scientific work is one of the basic pillars of science. The use of ICT and digitization has made it possible to share scientific materials and information with the academic community. The concept of Open Science emerged with the first science journals in the late 16^{th} and at the beginning of the 17^{th} century (David, 2008). At that time, researchers recognized the necessity of accessing the discoveries of other researchers and the possibilities offered by providing access to their own scientific works, all with the aim of sharing scientific opinions and ideas.

The term "Open Science" can be defined as a transformation of science – from the closed and traditional study to opening, sharing, and digitization; also, Open Science is becoming an obligatory part and a necessity in scientific work. According to Maurer, Open Science emphasizes complete, accurate, and timely publishing of scientific results. Maurer also notes that there is no restriction of access and that the transparency of information increases (Maurer, 2003). This movement is encouraged by the changes in ICT that transformed society – its economy, health care, and science. The very idea of Open Science is actually a systematical change in the implementation of scientific research in the last 50 years – the transition from the regular practice of scientific publishing to sharing and using all available materials and knowledge (European Commission, 2016). According to Burgelman et al. (2019), an early and well-known example of Open Science, dating from the pre-Internet era, is the Human Genome Project from 1990. The data on the human genome were widely shared among the scientific community during the project while a moratorium on publishing kept encouraging optimal collaboration. Due to this open science practice, researchers were able to decode the human genome in less than 15 years.

2 Open Science

According to one definition, Open Science represents the making of the scientific process completely available to the scientific community, through free access to methodology, data, and results of scientific work, publications, and articles and all other information obtained during the research process (Foster, 2017). This concept refers to the sharing of materials, in order to improve scientific work through transparency. Also, Open Science defines possibilities of cooperation between researchers, and within it, the scientific processes are available to others - under conditions that enable reuse of data and information. OECD defines Open Science as a possibility to make publicly funded data obtained via primary scientific research accessible for free or with minimal restrictions.[1] Open Science provides unlimited open access to research results, thus making the scientific processes more efficient, transparent, and adaptable to different social challenges. One of its main goals is the transformation of science through the use of different digital tools, in order to make research more global, creative, and bring it closer to society (European Commission, 2019).

[1] OECD Science, Technology and Industry Scoreboard 2015., accessed 9.5.2020.

As early as 2012, European Commission encouraged all EU member states to ensure open access to the results of publicly funded research, in order to improve scientific research and strengthen the knowledge-based economy (European Commission, 2018). By forming the OpenAire initiative, which entails support for the open science practices as well as technical support for infrastructure based on availability, the EU provided the necessary support to the quality development of the availability of scientific works. The chief commitment of OpenAire is to make publicly funded research available. The vital part of this initiative is the National Open Access Desks (NOADs) network, which consists of local experts, aimed at adopting Open Science practices, and bringing the concept closer to the country's scientific community.[2]

Despite all benefits of open science practices, there are still some conflicting views on the distribution of scientific results. Namely, although many believe that the possibility of free access to scientific materials could improve science, others are convinced that sharing results without a fee is not recommendable. For example, many publishers charge access to scientific papers published in their journals. This became a very popular topic in the last couple of years, mostly because many universities, faculties, and institutions are canceling their annual subscriptions to the biggest publishers. For example, the University of California canceled its annual subscription to *Elsevier* in July last year, believing that the activities of the world's largest publishers have created a monopoly in science. Advocating the principles of Open Science, the University in California is not the only to have given up on paying large annual membership fees (Resnick and Belluz, 2019).

Anyhow, Open Science is frequently defined as an umbrella term that denotes various movements aiming to remove the barriers for sharing any kind of output, resources, methods or tools, at any stage of the research process (Foster, 2017). Today, it is a generally accepted fact that allowing access to research results contributes to innovation and more efficient scientific practice in both public and private sector, and at the same time, it provides great support to policy-makers. Document *Open Innovation, Open Science, Open to the world* (European Commission, 2016) foresees that by 2030, Open Science will become the base of scientific work offering a range of unlimited opportunities for research in various scientific fields around the world.

[2] Open AireNOAD, accessed 3.5.2020.

3 Open Science in Serbia

Recognizing the main benefits arising from open science practices, countries around the world are adopting legal norms to regulate this movement in different areas of operation. In July 2018, the Ministry of Education, Science and Technological Development of the Republic of Serbia adopted the Open Science Platform. This document is based on principles of Open Science defined by the European Commission, which were incorporated in the proposal of the action plan for the implementation of the Strategy of Scientific and Technological Development of the Republic of Serbia from 2016 to 2020.[3]

The Platform also defines Open Access to scientific publications as a person's right to use the content without paying a fee - while being obliged to cite the source of the information. In addition to the obligatory Open Access to scientific publications, the Platform provides proposals for Open Access to primary data. Also, the Platform states that the clauses from the document do not refer to data that the scientific organization marks as confidential; as well as data that are considered to be a business secret or part of intellectual property. The Ministry of Education, Science and Technological Development of the Republic of Serbia recommends Open Access to data immediately after the results are published, whenever there are no legal or ethical restrictions to do so.

4 Open Data

Most scientific disciplines used empirical research methods based on the collecting and processing of different types of data. In recent years, the data were mostly stored in the computers belonging to research team members or one of the portable devices. Due to inadequate storage, the data would usually be lost after some time. Encouraged by the Open Science movement, the scientific community recognized the necessity to open access to data collected during research.

Open data is defined as online, free of charge, accessible data that can be used, reused, and distributed, provided that the primary data source is adequately cited (Foster, 2017). Open data usually refers to the process of depositing collected empirical data in institutional or thematic repositories, in order to ensure their storage and reuse (Смедеревац и др., 2020). According to the Open Data Handbook, these data may be used and redistributed

[3] Open Science Platform (2018), accessed 13.5.2020.

by any user free of charge, while adhering to copyright principles. The creators are governments and public administrations, whereby accessibility and transparency are encouraged via open data access.

Another large contributor is the scientific community itself. According to OECD, open data represent the equal access to data of every member of the research community. Also, Open Access should be simple, timely, and available online.[4] In order for the research data to be considered open, there must be no barriers to using, sharing, and reusing data. Because of this, we could definitely say that open data contributes to all areas of modern society.

Research conducted within the field of social sciences and humanities sometimes involves a lengthy and costly process of surveys, interviews, focus groups, and data anonymization. With open data, scientific research becomes more transparent, the diversity of ideas and opinions is encouraged, and new research is promoted. According to (Смедеревац и др., 2020), the process of data collecting had been an invisible activity for a long time, therefore, responsibility toward data was not part of academic skills and activities that were passed down from one generation of scientists to another.

The answer to the question „why do we need open data access“ is very simple – by ensuring data accessibility that does not jeopardize the privacy of individuals, the valuable resources that reduce costs and time are enabled. With it, scientific research rises to a higher level and provides a wide range of opportunities – data support new scientific research, improve the economy, etc. If science is considered as a public good, then it is clear why Open Science and open data are a necessity.

Despite all the benefits of open data, many scientists remain very skeptical. There are various attitudes among them – that their data are not good enough to be open and shared, that others would not be interested in their data. There is also fear of possible misuse, but also a lack of information on how the data can be processed and openly accessed. It is encouraging that exhaustive work is being done to popularize the open data concept and maximize protection – datasets can have an identification number (usually DOI), data can be adequately cited and protected by licenses; some journals only publish datasets.

Fienberg (1994) listed the benefits of open access to data:

– Scientific research becomes available, with an effective self-correcting factor

[4] OECD (2007): Principles and Guidelines for Access to Research Data from Public Funding.

- Diversity of ideas and opinions are encouraged because researchers have access to the same data and they have an opportunity to analyze each other's data
- Promotes new research and allows testing of new or alternative methods; there are numerous examples of data being used in a way that the original researchers did not anticipate
- Provides good materials in the teaching process; secondary data are extremely valuable for teachers who have access to quality data through the digital archive which they can use in their work with students
- The costs are reduced by avoiding duplication during data collecting. According to Bradić-Martinović et al. (2018), some datasets, such as General Social Survey and the National Election Studies, were used as a resource in thousands of scientific papers whereby the authors did not have to set aside time and money to collect data. The advantage of these datasets essentially comes down to the fact that collecting the same data twice becomes needless.

Funders of research projects insist on open access to data collected during research. One of them is *Horizon 2020* which insists on all research funded by them being made publicly available. European Commission believes that the information and data financed from public funds should be publicly available and provide benefits for the European scientific community, but also to all stakeholders in the process of global progress. *Horizon 2020* has a legal framework for Open Access as well as a set of instructions the project team must follow during research.

Also, it is important to note that not all data can be open data. Some data contain very sensitive information, and even with anonymization, the surveyed individuals or entities would not be fully protected. These data are called sensitive data and they are usually subject to special processing and storage. Access to these data is most often made possible by signing a contract with the institution that owns the data, through which the individual undertakes to use data exclusively for scientific purposes. In most cases, access to these data is possible from special computers found in the secured rooms of institutions that are under constant surveillance.

5 Where should Open Data be stored?

The answer to the question "where should open data be stored" imposed itself a few years ago. Numerous online services, such as Cloud systems, provide

the possibility to easily store datasets online. Although they are easy to use, the lack of IT support and the possibility of abuse show that these services are not the best choice. Digital data archives have been established as a permanent software solution, which facilitates the storing of these formats. Digital data archives are considered very cost-effective due to their self-sustainability. With data sharing and open data, digital archives became the key places to keep data secure and accessible. The connection between archives and data is explained by the fact that archives would not fulfill their basic function without data, and on the other hand, open data would not be sustainable without archives. Digital data archives support open data, open access, and open science, and at the same time, protect sensitive and personal data. To ensure this, many institutions use *Creative Commons Licenses* for data stored in digital archives. With these free licenses, copyright holders are clearly defined, which ensures legal access, and the archives precisely record to whom the data belong to.

6 Where should data be stored – state in the world and Serbia

Many developed countries throughout Europe have had national data centers for many years now, such as Social Science Data Archive from Slovenia, or UK Data Archive. Experts in different areas in the process of collecting, storing data, and open data processes are at service to the researchers in the field of social sciences and humanities, where they provide answers to various questions, but also promote open science and open data.

Researchers in Serbia are still not aware of the possibilities of open data access, even though project backers insist on open data. Experts from our country are working intensively on promoting this movement by organizing workshops and seminars so that the academic community can recognize the benefits of open science and open data.

Established in 2014, Data Center Serbia for Social Sciences is at the disposal of all researchers in Serbia. In November 2018, the Ministry of Education, Science and Technological Development of the Republic of Serbia provided necessary support to the Center, turning it into a national service provider for scientific data archiving, and Serbia thus became an equal member of the Consortium of European Social Science Data Archives – CESSDA

Eric.[5] Center promotes Open Science and open data while also respecting all legal provisions.

In the upcoming period, the academic community of Serbia will be forced to fully comply with the concept of Open Science and open data in segments where it has not been done so far. This will be one of many conditions for participating in projects funded by the European Union, but also for the purpose of harmonization of scientific work with Europe.

7 Conclusion

The inclination toward open data and the availability of science has existed for several centuries, therefore, the need to fully adopt and adapt this practice in the era of modern technologies is not surprising. Open science pushes all boundaries in the scientific work and unites researchers from the same or similar fields of research. There are still some doubts that open science is not necessarily a good thing, and there is some resistance to it – many researchers fear that the results of their work will be misused and that someone else may take credit for it. On the other hand, large publishers are not in favor of this movement because it would decrease their financial gain. However, software solutions that provide a high level of security, the use of licenses that provide protection, and legislation that defines copyright holders, refute these fears and theories of scientists. Although a novelty, open data has been gaining momentum in recent years. Europe is intensively working on introducing legislation that would remove all financial obstacles for publishing the results of scientific work. Although lagging behind Europe, Serbia is still on the right track, as proved by the Platform for Open Science from 2018, which encouraged many scientific organizations to become involved in the open data process. It is expected that the upcoming period will actually be a time of great changes in this field, which will mostly refer to the regulation of legal frameworks and requirements when it comes to the issue of open science in the Serbian academic community.

References

Bradić-Martinović, Aleksandra, Jelena Banović i Aleksandar Zdravković. "Repozitorijumi: digitalni resursi savremenog obrazovanja". *XXIV skup Trendovi razvoja" Digitalizacija visokog obrazovanja"*, 138–141. 2018

[5] Ministry of Education, Science and Technological Development of the RS, 2019

Burgelman, Jean-Claude, Corina Pascu, Katarzyna Szkuta, Rene Von Schomberg, Athanasios Karalopoulos et al.. "Open Science, Open Data, and Open Scholarship: European Policies to Make Science Fit for the Twenty-First Century". *Frontiers in Big Data* Vol. 2 (2019): 43

David, Paul A. "The Historical Origins of "Open Science": An Essay on Patronage, Reputation and Common Agency Contracting in the Scientific Revolution. Capitalism and Society". *Capitalism and Society* Vol. 3, no. 2 (2008)

European Commission. "Open Innovation, Open Science, Open to the World – a vision for Europe", 2016. Accessed May 11, 2020. https://ec.europa.eu/digital-single-market/en/news/open-innovation-open-science-open-world-vision-europe

European Commission. "Open Science (Open Access)", 2018. Accessed May 15, 2020. https://ec.europa.eu/programmes/horizon2020/en/h2020-section/open-science-open-access#Article

European Commission. "Open Science Open Science – Shaping Europe's Digital Future", 2019. Accessed May 15, 2020. https://ec.europa.eu/digital-single-market/en/open-science

Fienberg, Stephen E. "Sharing Statistical Data in the Biomedical and Health Sciences – Ethical, Institutional, Legal and Professional Dimensions". *Annual Review of Public Health* Vol. 15, no. 1 (1994): 1–18

Foster. "Open Data Definition", 2017. Accessed May 23, 2020. https://www.fosteropenscience.eu/taxonomy/term/110

Maurer, Stephen M. "New Institutions for Doing Science: From Databases to Open Source Biology". In *European Policy for Intellectual Property Conference on Copyright and database protection, patents and research tools, and other challenges to the intellectual property system*, 2003

Resnick, Brian and Julia Belluz. "The War To Free Science – How Librarians, Pirates, And Funders Are Liberting The World'S Academic Research From Paywalls", 2019. Accessed May 22, 2020. https://www.vox.com/the-highlight/2019/6/3/18271538/open-access-elsevier-california-sci-hub-academic-paywalls

Смедеревац, Снежана, Дејан Пајић, Сања Радовановић, Силвиа Гилезан, Петар Чоловић и др. "Отворена наука: пракса и перспективе", 2020