The State of Dissemination of Open Research Data in Ukraine and the World: Bibliometric Analysis

Boiko, Anton; Kramarenko, Olha; Shabanov, Sardar

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:
Boiko, A., Kramarenko, O., & Shabanov, S. (2021). The State of Dissemination of Open Research Data in Ukraine and the World: Bibliometric Analysis. *European Journal of Management Issues*, *29*(4), 209-217. https://doi.org/10.15421/192120

European Journal of Management Issues

Volume 29(4), 2021, pp. 209-217

DOI: 10.15421/192120

Received: 05 November 2021
Revised: 28 November 2021
Accepted: 03 December 2021
Published: 08 December 2021

UDC classification: 005.3:001.103.2:330.34

JEL Classification: I23, I25, I28, O1

The State of Dissemination of Open Research Data in Ukraine and the World: Bibliometric Analysis

A. Boiko [‡]***,***
O. Kramarenko [##]***,***
S. Shabanov [###]

Purpose: To determine the current state of development of open science in the paradigm of open research data in Ukraine and the world, as well as to analyze the representation of Ukraine in the world research space, in terms of research data exchange.

Design / Method / Research Approach: Methods of synthesis, logical and comparative analysis used to determine the dynamics of the number of research data journals and data files in the world, as well as to quantify the share of research data repositories in Ukraine and the world. Trend and bibliometric analysis were used to determine the share of publications with their open primary data; analysis of their thematic structures; identification of the main scientific clusters of such publications; research of geographic indicators and share of publications by research institutions.

Findings: The study found a tendency to increase both the number of data logs and data files in Dryad (open data repository). The results of the analysis of the share of data repositories indexed in re3data (register of research data repositories) show that 51% of the total number are repositories of data from European countries, with Germany leading with 460 repositories, followed by the United Kingdom (302 repositories) and France (116 repositories). Ukraine has only 2 data repositories indexed in re3data. The trend of relevance of data exchange is confirmed by the increase of publications with datasets for the last 10 years (2011-2020) in 5 times. Research institutions and universities are the main sources of research data, which are mainly focused on the fields of knowledge in chemistry (23.3%); biochemistry, genetics and molecular biology (13.8%); medicine (12.9%). An analysis of the latest thematic groups formed on the basis of publications with datasets shows that there is a significant correlation between publications with open source data and COVID-19 studies. More than 50% of publications with datasets both in Ukraine and around the world are aimed at achieving the goal of SDG 3 Good Health.

Theoretical Implications: It is substantiated that in Ukraine there is a need to implement specific tactical and strategic plans for open science and open access to research data.

Practical Implications: The results of the study can be used to support decision-making in the management of research data at the macro and micro levels.

Future Research: It should be noted that the righteous bibliometric analysis of the state of the dissemination of data underlying the research results did not include the assessment of quality indicators and compliance with the FAIR principles, because accessibility and reusability are fundamental components of open science, which may be an area for further research. Moreover, it is advisable to investigate the degree of influence of the disclosure of the data underlying the research result on economic indicators, as well as indicators of ratings of higher education, etc.

Research Limitations: Since publications with datasets in Scopus-indexed journals became the information base of the analysis for our study, it can be assumed that the dataset did not include publications with datasets published in editions that the Scopus bibliographic database does not cover.

Paper type: Theoretical

[‡] *Anton Boiko,*
Doctor of Economic Sciences, Associate Professor,
Sumy State University, Ukraine,
e-mail: a.boiko@uabs.sumdu.edu.ua,
https://orcid.org/0000-0002-1784-9364

[##] *Olha Kramarenko,*
PhD Student, Sumy State University, Ukraine,
e-mail: o.kramarenko@ssu.edu.ua,
https://orcid.org/0000-0002-2616-0147

[###] *Sardar Shabanov,*
PhD, Associate Professor,
Azerbaijan State University of Economics (UNEC), Azerbaijan,
e-mail: sardar.shabanov@unec.edu.az,
https://orcid.org/0000-0001-8020-6954

Keywords: open science, research data management, data sharing, research data repositories, publications with datasets.

Reference to this paper should be made as follows:
Boiko, A., Kramarenko, O., & Shabanov, S. (2021). The State of Dissemination of Open Research Data in Ukraine and the World: Bibliometric Analysis. *European Journal of Management Issues*, 29(4), 209-217. doi:*10.15421/192120*.

Стан поширення відкритих даних наукових досліджень в Україні та світі: бібліометричний аналіз

Антон Бойко [ǂ]
Ольга Крамаренко [ǂ]
Сардар Шабанов [#]

[ǂ] *Сумський державний університет, Україна*
[#] *Азербайджанський державний університет економіки, Азербайджан*

Мета дослідження: Дослідження має на меті визначити сучасний стан розвитку відкритої науки в парадигмі відкритих даних досліджень в Україні та світі, а також проаналізувати представлення України у світовому дослідницькому просторі, в частині обміну даними досліджень.

Дизайн / метод / підхід дослідження: Методи синтезу, логічного та порівняльного аналізу, використані з метою визначення динаміки кількості журналів даних досліджень та файлів даних в світі, а також для здійснення кількісної оцінки частки репозитаріїв даних досліджень в Україні та світі. Трендовий та бібліометричний аналіз використано для визначення частки публікацій з їх відкритими первинними даними; аналізу їх тематичних структур; визначення основних наукових кластерів таких публікацій; дослідження географічних показників та частки публікацій за дослідницькими установами.

Результати дослідження: Дослідження виявило тенденцію до зростання як кількості журналів даних, так і файлів даних в Dryad (репозитарій відкритих даних досліджень). Результати аналізу частки репозитаріїв даних індексованих в re3data (реєстр сховищ даних досліджень) показують, що 51% від загальної кількості складають репозитарії даних країн Європи, причому Німеччина лідирує з 460 репозитаріями, за якою йдуть Великобританія (302 репозитарії) та Франція (116 репозитаріїв). Україна має лише 2 репозитарії даних, індексованих в re3data. Тенденція актуальності обміну даними підтверджується збільшенням публікацій з наборами даних за останні 10 років (2011-2020) у 5 разів. Науково-дослідні установи та університети є основними джерелами даних досліджень, які переважно зосереджені на галузях знань з хімії (23.3%); біохімії, генетики та молекулярної біології (13.8%); медицини (12.9%). Аналіз найновіших тематичних груп, що були сформовані на основі публікацій з наборами даних показує, що існує значна кореляція між публікаціями з відкритими первинними даними та дослідженнями COVID-19. Більше 50% публікацій з наборами даних як в Україні, так і світі спрямовано на забезпечення цілі SDG 3 «Міцне здоров'я».

Теоретичне значення дослідження: Обґрунтовано, що в Україні виникає необхідність впровадження конкретних тактичних й стратегічних планів щодо відкритої науки та відкритого доступу до даних досліджень.

Практичне значення дослідження: Результати дослідження можуть бути використанні для підтримки прийняття рішень в управлінні даними досліджень на макро та мікрорівні.

Перспективи подальших досліджень: Варто зазначити, що праведний бібліометричний аналіз стану поширення даних, які лежать в основі результатів досліджень, не включав оцінювання показників якості та відповідність FAIR принципам, адже доступність та можливість повторного використання є основоположними складовими відкритої науки, що може бути напрямком подальших досліджень. Крім того, також доцільно дослідити ступінь впливу оприлюднення даних, які лежать в основі результату дослідження на економічні показники, а також показники рейтингів закладів вищої освіти тощо.

Обмеження дослідження: Оскільки інформаційною базою аналізу для нашого дослідження стали публікації з наборами даних в журналах, що індексуються базою Scopus, можна передбачити, що у вибірку даних не увійшли публікації з наборами даних, що були опубліковані у виданнях, які бібліографічна база Scopus не охоплює.

Тип статті: Теоретичний

Ключові слова: відкрита наука, управління даними досліджень, обмін даними, репозитарії даних досліджень, публікації з наборами даних.

Состояние распространения открытых данных научных исследований в Украине и мире: библиометрический анализ

Антон Бойко [ǂ]
Ольга Крамаренко [ǂ]
Сардар Шабанов [#]

[ǂ] *Сумской государственный университет, Украина*
[#] *Азербайджанский государственный университет экономики, Азербайджан*

Цель исследования: Целью исследования является определить современное состояние развития открытой науки в парадигме открытых данных исследований в Украине и мире, а также проанализировать представление Украины в мировом исследовательском пространстве, в части обмена данными исследований.

Дизайн / метод/ подход исследования: Методы синтеза, логического и сравнительного анализа, использованные для определения динамики количества журналов данных исследований и файлов данных в мире, а также для количественной оценки доли репозитариев данных исследований в Украине и мире. Трендовый и библиометрический анализ использован для определения части публикаций с их открытыми первичными данными; анализа их тематических структур; определение основных научных кластеров таких публикаций; исследование географических показателей и доли публикаций по исследовательским учреждениям.

Результаты исследования: Исследование выявило тенденцию роста как количества журналов данных, так и файлов данных в Dryad (репозитарий открытых данных исследований). Результаты анализа доли репозитариев данных индексированных в re3data (реестр хранилищ данных исследований) показывают, что 51% от общего количества составляют репозитарии данных стран Европы, причем Германия лидирует с 460 репозитариями, за которыми следуют Великобритания (302 репозитарии) и Франция (116 репозитариев). Украина имеет только 2 репозитария данных, индексированных в re3data. Тенденция актуальности обмена данными подтверждается увеличением публикаций с наборами данных за последние 10 лет (2011-2020) в 5 раз. Научно-исследовательские учреждения и университеты являются основными источниками данных исследований, которые в основном сосредоточены в областях знаний по химии (23.3%); биохимии, генетики и молекулярной биологии (13.8%); медицины (12.9%). Анализ новейших тематических групп, сформированных на основе публикаций с наборами данных, показывает, что существует значительная корреляция между публикациями с открытыми первичными данными и исследованиями COVID-19. Более 50% публикаций с наборами данных как в Украине, так и в мире направлено на обеспечение цели SDG 3 «Крепкое здоровье».

Теоретическое значение исследования: Обосновано, что в Украине возникает необходимость внедрения конкретных тактических и стратегических планов открытой науки и открытого доступа к данным исследований.

Практическое значение исследования: Результаты исследования могут использоваться для поддержки принятия решений в управлении данными исследований на макро и микроуровне.

Перспективы дальнейших исследований: Следует отметить, что праведный библиометрический анализ состояния распространения данных, лежащих в основе результатов исследований, не включал оценки показателей качества и соответствие принципам FAIR, ведь доступность и возможность повторного использования являются основополагающими составляющими открытой науки, что может быть направлением дальнейших исследований. Кроме того, целесообразно исследовать степень влияния обнародования данных, лежащих в основе результата исследования на экономические показатели, а также показатели рейтингов учреждений высшего образования и т.д.

Ограничения исследования: Поскольку информационной базой анализа для нашего исследования стали публикации с наборами данных в журналах, индексируемых базой Scopus, можно предположить, что в подборку данных не вошли публикации с наборами данных, опубликованными в изданиях, которые библиографическая база Scopus не охватывает.

Тип статьи: Теоретический

Ключевые слова: открытая наука, управление данными исследований, обмен данными, репозитарии данных исследований, публикации с наборами данных.

1. Introduction

The active annual spread of new and worsening existing economic, social, environmental, and other problems requires the scientific community to publish their research results and disseminate the data that formed the basis of the research in the open access. The dissemination of data in open access plays an essential role in the implementation of sustainable development goals, on the one hand, ensuring equal access to scientific knowledge regardless of nationality, race, gender, income level, socio-economic status, career stage, religion, disability, ethnicity, and on the other, creating opportunities for faster research with the ability to achieve synergistic effects in solving global and national social problems.

The institutional development of the system for the dissemination of open data of scientific research is laid down in such documents as the Budapest Open Access Initiative (Chan et all., 2001), the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (Redalyc, Clase, & In-Com Uab, *2003*), and in Open Innovation, Open Science, Open to the World of the European Commission (*Open Innovation, 2016*). The above documents served as an impetus for developing open science and open data research around the world. In addition, they formed the basic components of open science: open access to publications, open research data, and open peer review.

The development of open exchange of medical research data is underpinned by a statement from the National Institutes of Health (NIH) published in 2002, which stipulated that research data supported by the NIH should be available for further analysis and use (NIH, 2002); launched in 2011 by the Yale University Open Data Access (YODA) Project with a mission of responsible clinical research data sharing, open science and research transparency (*YODA, 2011*); promulgated in 2014 by the Open Data Policy of the Public Library of Science (PLoS), which is based on the requirement for the availability of data underlying research results published in the public domain (*Bloom, Ganley & Winker, 2014*).

The accelerated progress of global scientific research leads to an increase in research data. However, the raw data obtained as a result of research and not placed on long-term access platforms can be lost forever. Thus, the modern research ecosystem emphasizes the need to focus on the management of research data. It is the development of an integrated data management system that will ensure the safety of information and ensure its effective use for future scientific research. The organization of research data management requires each country to have its strategies, plans, and data exchange programs. At the same time, it is fair to say that the greatest attention in the organization of research data management and support for researchers is focused on research institutions and higher education institutions. The construction of an effective data management ecosystem, the infrastructure of which will be based on government and commercial scientific and educational institutions, will ensure the proactive development of all spheres of the public life of citizens.

All of the above has determined the relevance of the study, its purpose, objectives, and content.

2. Theoretical background

Turning directly to the analysis of the current situation with the dissemination of open data of scientific research in Ukraine and the world, it should be noted that a growing number of leading journals and publishing houses of the world, such as Elsevier, Springer Nature, PNAS, are developing their data exchange policies (*Hrynaszkiewicz, 2019*).

With the advancement of open science and open data globally, the organizational and descriptive aspects of data that encourage data exchange and reuse are becoming increasingly important. Thus, a growing number of global financial donors such as the Arts and Humanities Research Council (AHRC), Biotechnology and Biological Sciences Research Council (BBSRC), Engineering and Physical Sciences Research Council (EPSRC), European Commission (EC), National Science Foundation (NSF)), The United States Geological Survey (USGS), the Australian Research Council (ARC) require grant applicants to submit a Research Data Management Plan at the start of the project (Nature Editorial, 2018) and disseminate research data after the completion of the study. One of the conditions for fully ensuring the implementation of the plan for sustainable data exchange is the availability of data for both humans and machines. Considering the above, in 2016, the FORCE11 community members identified key requirements for ensuring the quality of research data, the fundamental requirements of which were Findable, Accessible, Interoperable and Reusable – the FAIR principles (*Wilkinson et al, 2016*).

The development of the movement for the open exchange of research data poses the question of finding ways to distribute and provide long-term storage of research data. In this regard, Research Data Repositories (RDRs) greatly facilitate the sharing and reuse of research data.

Thus, on the initiative of a group of leading world journals and scientific societies, Dryad was organized as one of the most powerful repositories of open data research (*Dryad, n.d.*). As of October 5, 2021, there is an upward trend in both the number of data logs and data files in Dryad (*Fig. 1*). This trend evidences the support by world scientific publishers of the open science strategy in terms of data exchange by switching to the "data journal" model, for most of which open access to primary research data in appropriate repositories is potentially significant.

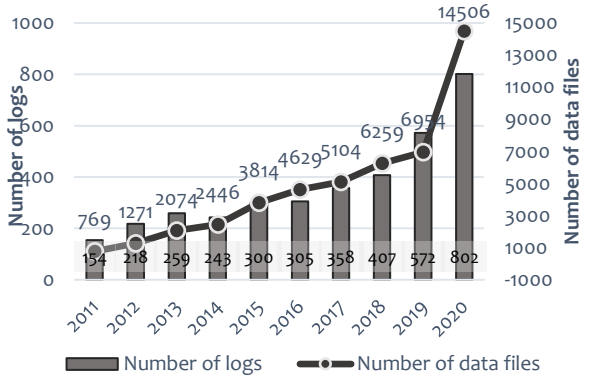

Figure 1: Number of logs and data files in Dryad (2011-2020)

Source: developed by the authors based on *Dryad (n.d.) data*

It should be noted (*Fig. 1*) that the progressive trend of growth in the number of data files for 2011-2019 was replaced by a rapid increase in this indicator in 2020. Thus, the number of data files increased in 2020 by 109%, which is 7,552 units more than in 2019. In turn, the outstripping growth in the number of journals has been going on for the last two years (2020 and 2019), as evidenced by their 40% annual growth. The corresponding trend is primarily due to the announcement in September 2018 of Plan S by Science Europe, which states that after January 1, 2020, all results of scientific research funded by state grants provided by national and European research councils and financial organizations and the data underlying such research should be published exclusively in open access journals and related open platforms (*Plan S, 2018*). The next milestone was the 40th session of the UNESCO General Conference, which took place in September 2019, at which 193 Member States instructed the Organization to develop an international standard instrument for open science in the form of the UNESCO Recommendation. It is envisaged that the Recommendation will define the general values and principles of open science, specific activities on open access, and open research data, with further proposals to promote the dissemination of scientific knowledge worldwide (*UNESDOC Digital Library, 2019*).

Since the dissemination of data and their long-term storage has become an essential component of an open science and data management process, the issue of organizing a unified register of data repositories has become relevant, allowing for the selection of a repository in accordance with the needs of a researcher, the conditions of a funding organization or publishing house. Thus, in 2013, the re3data (n.d.), a global registry of research data repositories, was launched. As of October 2021, this registry has over 2,749 data repositories worldwide. It is worth noting that the United States is a leader in the number of research data repositories (1,137 repositories). In turn, Europe is represented by 36 countries with 1,390 repositories, which is 51% of the total number in re3data (n.d.). Thus, the top three European countries in terms of the number of data repositories indexed in re3data are Germany (460), Great Britain (302), and France (116) (*Fig. 2*).

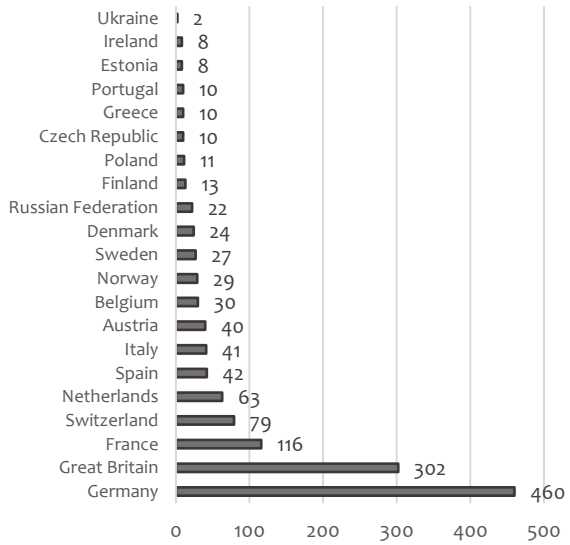

Figure 2: Number of European research data repositories indexed by re3data

Source: developed by the authors according to *re3data* (n.d.) *data*

Determining the state of development of data repositories in Ukraine, it should be noted that their share in re3data (n.d.) is insignificant and includes 2 repositories (*Fig. 2*).

National initiatives have been introduced at the state level by many countries worldwide to regulate open science, the exchange of data, and their reuse. The United States Office of Science and Technology Policy (OSTP) has issued a memorandum committing the public to expand access to taxpayer-funded research products (*Holdren, 2013*). The world's best open access initiatives are the Amsterdam Call for Action for Open Science (*2016*), which recommends that each EU member state develop its own National Open Science Plan; France's National Plan for Open Science, which focuses on regulating access to research data (*Second French Plan, 2021*), and the Netherlands' National Plan for Open Science (*van Wezenbeek, Touwen, Versteeg, & van Wesenbeeck, 2017*).

Turning to Ukraine, we note that for our state, September 2017 can be considered the beginning of the introduction of open science and open data research at the level of public administration, when the Association Agreement between Ukraine, on the one hand, and the European Union, the European Atomic Energy Community, and their member states, on the other hand, finally entered into force. According to clause 2 of Art. 375 of the said agreement, "cooperation between the Parties is aimed at promoting the involvement of Ukraine in the European Research Area" (*Association Agreement*).

Continuing to ensure the integration of Ukraine into the European Research Area, the Order of the Ministry of Education and Science of Ukraine No. 167 dated February 10, 2021, approved a roadmap (*MONU, 2021*). Thus, the goal of Priority 5b Open Science and Digital Innovation is to apply the open science policy at all stages of scientific research, including the introduction of accessibility, reusability, findability for all types of scientific data.

3. Literature review

Ukrainian scientists (*Durman & Tochtarova, 2017; Tarasiuk, 2017; Karpenko & Ryzhenko, 2018; Kovtun & Khriakova, 2018*) have conducted a number of studies on open data, the state of their development and distribution. However, these works are devoted to studying open government data with slightly different specifics than open research data.

Due to the lack of a universal approach to the formation of public policy of open science and research data in Ukraine, attention should be paid to the scientific works by Vasylenko, which deal with the analysis of existing approaches to the development of national visions of the Kingdom of the Netherlands (*Vasylenko, 2019b*), France (*Vasylenko, 2019a*), and the Baltic countries (*Vasylenko, 2020*). Of particular note is the scientific heritage of a team of authors from Ukraine (Serhii Nazarovets, Nataliia Kaliuzhna) and Germany, who in their work consider criteria for facilitating the search, accessibility, interaction, and reuse of research information in open infrastructures (*Hauschke et al., 2021*).

Foreign researchers have carried out a range of scientific works dealing with the research of open scientific data. In the study (*Hajduk et al., 2019*), scientists emphasize that data exchange should be simple, feasible, and accessible. Research by scientists demonstrating the benefits of applying in practice the principles of open science and the dissemination of research data should also be considered (*Pienta, Alter & Lyle, 2010; Henneken, & Accomazzi, 2011; Dorch, 2012; Piwowar & Vision, 2013; McKiernan et al., 2016; Zhang, Ma, 2021*).

4. Problem statement

The objective of the research is to determine the current state of development of open science and research data in Ukraine and the world and analyze the representation of Ukraine in the global research space in terms of the exchange of research data.

5. Research methodology

A bibliometric study of the state of data dissemination was analyzed using the SciVal tool. The research area "Publications with datasets indexed in Data Monitor," based on the data on the number of publications in Scopus with datasets, was selected as the information base for the analysis. The analysis was carried out in the range of the last 10 years (2011-2020). The Scopus database was used to obtain data on the total number of publications of scientists in Scopus-indexed journals.

6. Results

It is proposed to start the analysis of the current state of dissemination of research data in Ukraine and the world with an analysis of the total number of publications with datasets in Scopus (n.d.) over the past 10 years (2011-2020). According to SciVal (n.d.), the trend in the relevance of disseminating research data is confirmed by their annual increase by an average of 20%. In general, during 2011-2020, the growth in the number of publications in Scopus with datasets was 396% (or 61,091 units). For comparison, the average annual growth rate of the number of publications in Scopus is 3% (*Fig. 3*). This trend indicates the outstripping growth of Scopus-indexed publications with datasets over the last 10 years and, accordingly, confirmation of the hypothesis about the demand for publications, the quality of which is confirmed by the datasets underlying the study.

It is worth noting that the top 10 countries in terms of the number of publications with datasets in Scopus are the United States (97,239), China (93,602), Germany (42,393), Great Britain (42,031), Brazil (25,654), India (24,092), Japan (22,972), France (22,866), Canada (20,101), Australia (18,628). Ukraine ranks 45th out of 100 in the overall rating with 1,704 publications (*Fig. 4*).

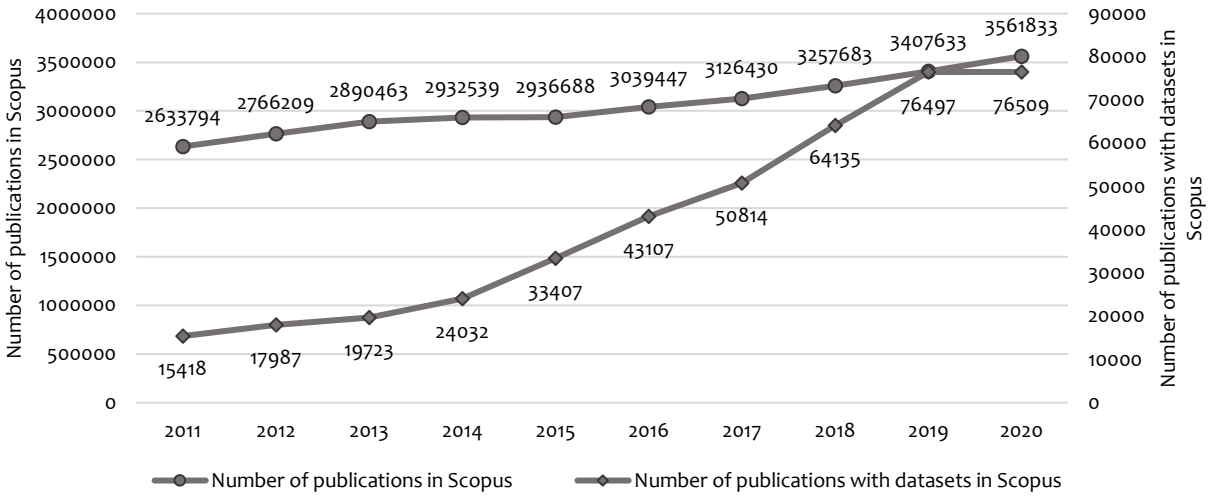

Figure3: Number of publications with and without datasets in Scopus database (2011-2020)

Source: developed by the authors based on *Scopus* (n.d.) *and SciVal* (n.d.) *data* (accessed on October 05, 2021)

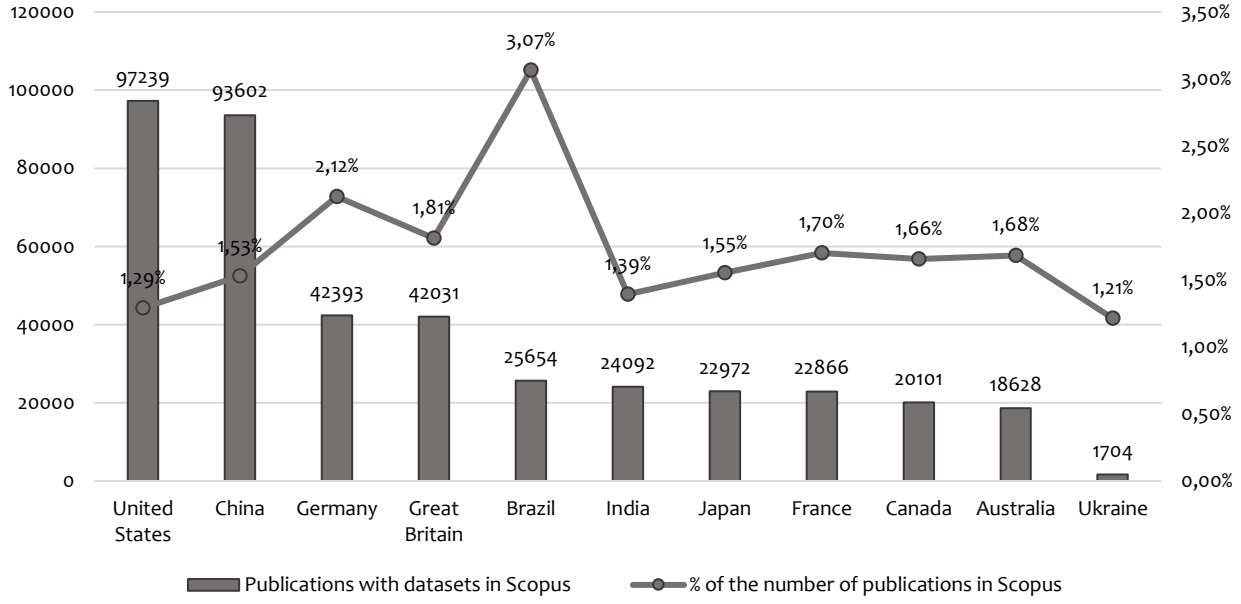

Figure 4: Top 10 countries by number of publications with and without datasets in Scopus (2011-2020)

Source: developed by the authors based on *Scopus* (n.d.), *SciVal* (n.d.) *data* (accessed on October 05, 2021)

Continuing the study of dozens of countries with the largest number of publications with datasets in Scopus, it should be noted that the ranking positions shift depending on the total number of publications in Scopus (n.d.). Thus, the percentage of the share of publications with datasets concerning the total number of publications in Scopus in Brazil is the highest and amounts to 3.07%, followed by Germany (2.12%), Great Britain (1.81%), France (1.70%), Australia (1.68%), the United States (1.29%). Thus, in Brazil and Germany, researchers pay much more attention to disseminating original research data (*Fig. 4*).

After analyzing the affiliation of researchers to the institution, we found that the largest number of publications in Scopus with datasets belong to scientists from the Chinese Academy of Sciences (14,575), the National Center for Scientific Research (CNRS) (13,892), the Ministry of Education, China (10,677), Harvard University (5,386), University of the Chinese Academy of Sciences (4,893) (*Tab. 1*).

Thus, the top 10 organizations by the number of publications in terms of the research field are formed by the government organizations of China (3 organizations) (*Tab. 1*). This trend confirms the impact of the domestic national initiative on the development of Open Science in the paradigm of publishing research data (*Normile, 2018*). In 2018, the General Administration of the State Council of China approved the Measures for the Management of Scientific Data (*Enago, 2018*).

Table 1: Top 10 organizations by the number of publications with datasets in Scopus (2011-2020)

Affiliation	Country	Number of publications
Chinese Academy of Sciences	China	14,575
French National Centre for Scientific Research (CNRS)	France	13,892
Ministry of Education, China	China	10,677
Harvard University	United States	5,386
University of the Chinese Academy of Sciences	China	4,893
University of Sao Paulo	Brazil	4,731
Oxford University	United Kingdom	4,563
Spanish National Research Council (CSIC)	Spain	4,532
Russian Academy of Sciences	Russian Federation	4,108
University of Toronto	Canada	3,992

Source: developed by the authors based on *SciVal* (n.d.) *data* (accessed on October 05, 2021)

In the context of our research, it is interesting to consider the main areas of knowledge in which materials with open data were published. The analysis was carried out according to the ASJC (All Science Journal Classification) classification used in Scopus (n.d.). Among the publications of scientists with datasets, research in the field of chemistry prevails, accounting for 23% of all thematic groups; biochemistry, genetics, and molecular biology is 13%; medicine covers 12% (*Fig.* 5).

Regarding the field-weighted citation index (FWCI), the thematic structures of economics and econometrics, decision-making, arts and humanities, management, and business have the greatest impact (*Fig.* 5). Thus, publications with datasets cover all subject areas. The greatest focus on the chemical, biochemical and medical industries is driven by the data sharing policies of medical journal publishers and medical research funding organizations (*National Institutes of Health, 2002; YODA, 2011; Bloom et al., 2014*).

Researchers worldwide have made significant contributions to 22 thematic groups, which have seen substantial growth in recently published research and have attracted the most funding, according to SciVal. We have studied groups of topics that contain 10 or more publications.

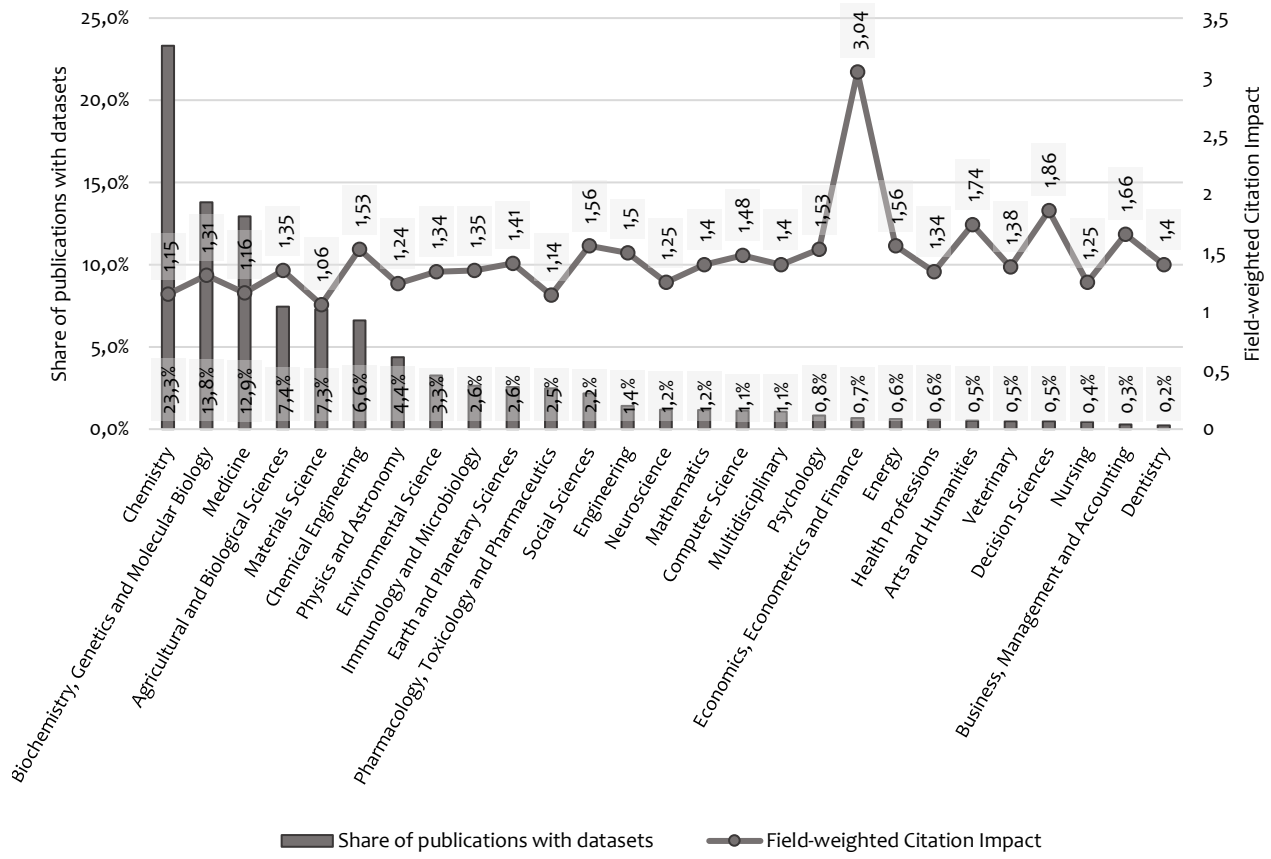

Figure 5: Structure of publications with datasets by subject area in Scopus (2011-2020)

Source: developed by the authors based on *SciVal* (n.d.) *data* (accessed on October 05, 2021)

Thus, the thematic groups by the number of publications were distributed as follows: Protective effect; Magnetic properties; Coordination polymer (219 publications); X-ray images; Clinical features; COVID-19 (200 publications), Nasopharyngeal swabs; Serological tests; COVID-19 (73 publications), ARIMA; Mathematical modeling; COVID-19 (60 publications) (*Tab.* 2).

The connection of publications with datasets with research related to COVID-19 cannot be ignored because the global health crisis caused by the COVID-19 pandemic has demonstrated to the whole world the urgent need to ensure equal access to scientific information, promote the exchange of scientific knowledge, data and facts, strengthening scientific collaboration and decision-making based on scientific evidence and knowledge (*Zastrow, 2020; Wellcome, 2020*).

Table 2: Thematic groups by the number of publications with datasets in Scopus (2011-2020)

Thematic groups	Number of publications
Protective effect; Magnetic properties; Coordination polymer	219
X-ray images; Clinical features; COVID-19	200
Nasopharyngeal smears; Serological tests; COVID-19	73
ARIMA; Mathematical modeling; COVID-19	60
Psychological support; Mindfulness; COVID-19	55
Pharmacotherapy; Prolongation; COVID-19	46
Hepatocellular carcinoma; Tumor microenvironment; Prognosis	43
Interleukin 6; Rheumatology; COVID-19	37
COVID-19 coagulopathy; Thrombocytopenia; Patients	29
Cell proliferation; Cerna; Invasion	25
Mother; Pregnancy; COVID-19	22
Neurological manifestations; Meningoencephalitis; COVID-19	18
Exchange rates; Prices; Bitcoin	16
Hematological malignancies; Cancer treatment; COVID-19	14
National minorities; Racism; COVID-19	13
Hip fractures; General surgery; COVID-19	13
Congenital heart defects; Children; COVID-19	12
Environmental pollution; Medical waste; COVID-19	12
Hypoxemia; Happy hypoxia; COVID-19	11
Children's dentistry; Dentists; COVID-19	11
Intravitreal injections; Teleophthalmology; COVID-19	10
Cats; Pangolin; SARS-CoV-2	10

Source: developed by the authors based on *SciVal* (n.d.) *data* (accessed on October 05, 2021)

Turning directly to the analysis of the current state of disseminating research data in Ukraine, it is fair to note the growth in the number of publications with datasets in Scopus over the past 10 years (2011-2020). Consequently, in 2020, the number of publications with datasets increased by 120% compared to 2011 (*Fig. 6*).

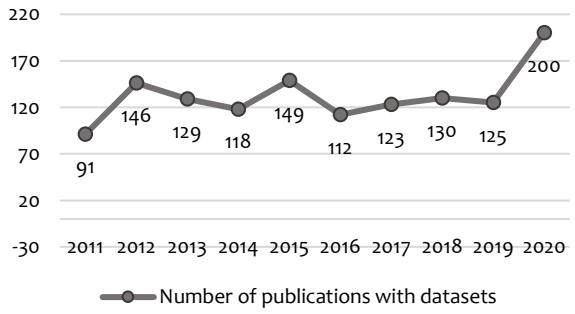

Figure 6: Number of Ukrainian publications with datasets in Scopus (2011-2020)

Source: developed by the authors based on *SciVal* (n.d.) *data* (accessed on October 05, 2021)

Continuing the study of publications with datasets in Scopus of Ukrainian scientists, it is advisable to analyze their affiliation with a research institution. Thus, data (*Tab. 3*) show that most of the publications with datasets in Scopus belong to scientists from the National Academy of Sciences of Ukraine (734 publications). The following positions are occupied by researchers from Taras Shevchenko National University of Kyiv (510 publications), the Institute of Single Crystals (268 publications), the Institute of Organic Chemistry of the National Academy of Sciences of Ukraine (210 publications), and V. N. Karazin Kharkiv National University (203 publications) (*Tab. 3*).

Table 3: Top 20 organizations in Ukraine by the number of publications with datasets in Scopus (2011-2020)

Institutions	Scholarly Output	Share of publications
National Academy of Sciences of Ukraine	734	21.69%
Kyiv National Taras Shevchenko University	510	15.07%
Institute for Single Crystals	268	7.92%
NASU - Institute of Organic Chemistry	210	6.21%
V. N. Karazin Kharkiv National University	203	6.00%
Enamine Ltd	189	5.59%
National Science Center Kharkov Institute of Physics and Technology	133	3.93%
Ivan Franko National University of L'viv	118	3.49%
NASU - Bogolyubov Institute for Theoretical Physics	81	2.39%
NASU - Institute for Scintillation Materials	65	1.92%
NASU - Institute of Bioorganic Chemistry and Petrochemistry	63	1.86%
Vernadsky Institute General and Inorganic Chemistry	58	1.71%
NASU - Bogatsky Physico-Chemical Institute	50	1.48%
NASU - Main Astronomical Observatory	43	1.27%
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"	43	1.27%
NASU - Pisarzhevskii Institute of Physical Chemistry	40	1.18%
National University of Pharmacy	33	0.98%
Odessa National University	32	0.95%
Danylo Halytsky Lviv National Medical University	30	0.89%
NASU - Institute of Molecular Biology and Genetics	25	0.74%

Source: developed by the authors based on *SciVal* (n.d.) *data* (accessed on October 05, 2021)

According to the sample (*Tab. 3*), it should be noted that 95% of publications with datasets were carried out by scientists from academic institutions of Ukraine.

It should be noted that over the past 10 years, the research activities of both world scientists and researchers in Ukraine have been aimed, among other things, at ensuring the sustainable development goals SDG 1 – SDG 17. In general, during 2011-2020, scientists worldwide conducted 116,264 studies in this area, which is 28% of the total number of publications with datasets in Scopus (n.d.). The share of research by Ukrainian scientists on ensuring sustainable development goals of the total number of publications with datasets in Ukraine is 7%.

Our research has shown that the thematic research structures of scientists from the world and Ukraine have the same priorities. Thus, more than 50% of the studies of both world scientists (70,302 works) and Ukraine (58 works) were carried out in the medical industry and are aimed at ensuring the goal of SDG 3 "Good Health". The share of publications on ensuring the

development of clean energy SDG 7 "Affordable and clean energy" is 17% (17 works) of the number of studies of Ukrainian scientists and 5.4% (7,324 works) of studies with datasets of scientists from the world. 9% (9 papers) of research by Ukrainian scientists and 7.1% (9,650 papers) of scientists from all over the world are aimed at ensuring the SDG 15 "Life on land" (*Fig.* 7).

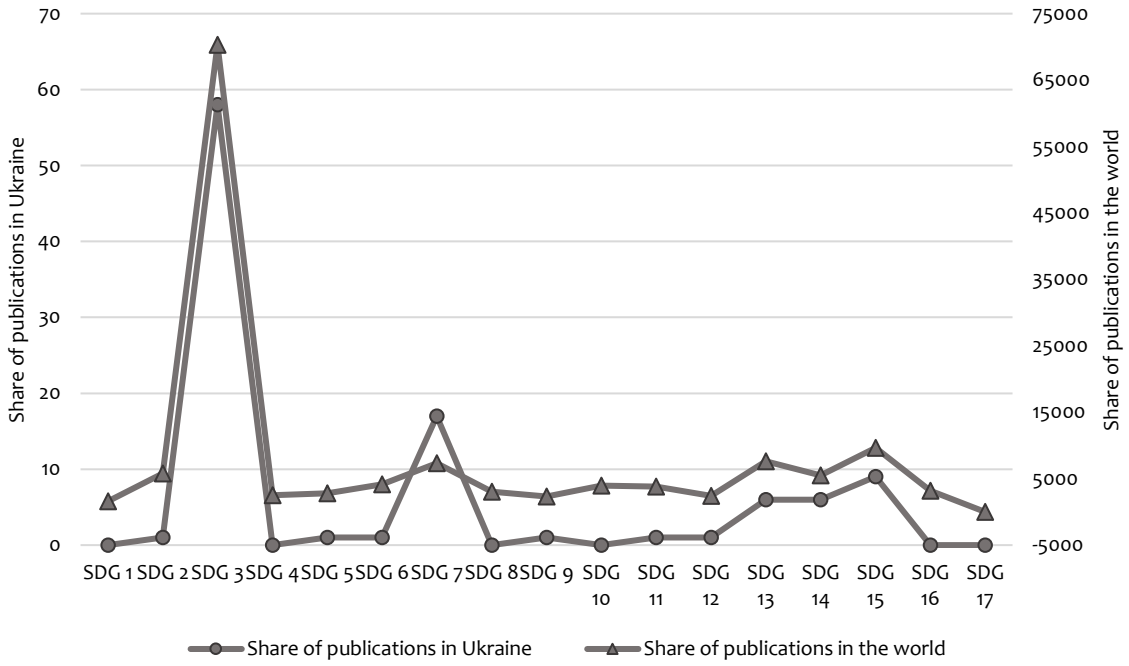

Figure 7: Number of publications with datasets by scientists from Ukraine and the world in the area of SDG 1–SDG 17 in Scopus (2011-2020)

Source: developed by the authors based on *SciVal* (n.d.) *data* (accessed on October 05, 2021)

Thus, this trend in disseminating research to ensure sustainable development goals (Fig. 7) once again confirms the relevance of disseminating primary data in the medical field of knowledge.

7. Conclusion

Thus, our study demonstrates the development of open science around the world: the leading countries are introducing their National Initiatives, the fundamentals for which are the settlement of issues of dissemination and long-term storage of research data; the share of data repositories is growing every year; bibliometric analysis has shown an annual increase in the publication of research datasets around the world.

A problem in the development of open research data in Ukraine is the lack of a single conceptual document regulating the issues of their publication and reuse; lack of national and institutional FAIR data repositories; poor technical support for the organization of such storage facilities.

For the further integration of Ukraine into the European Research Area, it becomes necessary to create specific plans to implement open science and research data, programs to control their quality and availability.

8. Funding

This study received no specific financial support.

9. Competing interests

The authors declare that they have no competing interests.

References

Amsterdam Call for Action on Open Science. (2016). *Information from the Government of The Netherlands*. Retrieved from https://www.government.nl/documents/reports/2016/04/04/amsterdam-call-for-action-on-open-science.

Bloom, T., Ganley, E., & Winker, M. (2014). Data access for the open access literature: PLOS's data policy. *PLoS Medicine, 11*(2), e1001607. doi:10.1371/journal.pmed.1001607.

Chan, L., Cuplinskas, D., Eisen, M., Friend, F., Genova, Y., Guédon, J.-C., & Hagemann, M. (2002). *Budapest Open Access Initiative*. Retrieved from http://www.opensocietyfoundations.org/openaccess/read.

Dorch, B. (2012). *On the citation advantage of linking to data: Astrophysics*. Retrieved 11 September 2021 from https://hal-hprints.archives-ouvertes.fr/hprints-00714715v2.

Dryad. (n.d.) *Official site.* Available at: https://datadryad.org/stash/.

Durman, N. A., & Tochtarova, I. M. (2017). Open data as an instrument for informational transparency of public authorities. *Teoriia ta praktyka derzhavnogo upravlinnia i mistsevogo samovriaduvannia*, (1). Retrieved from http://nbuv.gov.ua/j-pdf/Ttpdu_2017_1_15.pdf. [in Ukrainian]

Enago. (2018). China Open Science and Open Data Mandate Released. *Enago Academy*. Retrieved from https://www.enago.com/academy/china-open-science-open-data-manadate-released/.

Hajduk, G. K., Jamieson, N. E., Baker, B. L., Olesen, O. F., & Lang, T. (2019). It is not enough that we require data to be shared; we have to make sharing easy, feasible and accessible too! *BMJ Global Health, 4*(4), e001550. doi:10.1136/bmjgh-2019-001550.

Hauschke, C., Nazarovets, S., Altemeier, F., & Kaliuzhna, N. (2021). Roadmap to FAIR Research Information in Open Infrastructures. *Journal of Library Metadata*, 1–17. doi:10.1080/19386389.2021.1999156.

Henneken, E. A., & Accomazzi, A. (2011). *Linking to data-effect on citation rates in astronomy*. arXiv.org. Retrieved from https://arxiv.org/abs/1111.3618

Holdren, J. P. (2013). *Memorandum for the heads of executive departments and agencies: Increasing access to the results of federally funded scientific research*. Retrieved from http://web.archive.org/web/20160115125401/https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.

Hrynaszkiewicz, I. (2019). Publishers' Responsibilities in Promoting Data Quality and Reproducibility. *Handbook of Experimental Pharmacology*, 319–348. doi:10.1007/164_2019_290.

Karpenko, O. V., & Ryzhenko, O. V. (2018). Digital transformation and good data governance in Ukraine. *Bulletin of the NAPA, Series «Public Administration»*. 4(91). Retrieved from http://visnyk.academy.gov.ua/pages/dop/82/files/7c9c905c-cfa1-42de-ab8d-1395fb2ec8ba.pdf.

Kovtun, M. S., & Khriakova, N. O. (2018). Vidkryti dani v Ukrayini: sutnist' ta stan rozvytku [Open data in Ukraine: the essence and state of development]. *Forum Prava*, (2). 63-73. doi:10.5281/zenodo.1286078. [in Ukrainian]

McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., ... & Yarkoni, T. (2016). Point of view: How open science helps researchers succeed. *elife*, 5, e16800. doi:10.7554/eLife.16800.001.

MONU. (2021). Pro zatverdzhennia dorozhnoi karty z intehratsii naukovo-innovatsiinoi systemy Ukrainy do yevropeiskoho doslidnytskoho prostoru. *Ministerstvo osvity i nauky Ukrainy*. Retrieved from https://mon.gov.ua/ua/npa/pro-zatverdzhennya-dorozhnoyi-karti-z-integraciyi-naukovo-innovacijnoyi-sistemi-ukrayini-do-yevropejskogo-doslidnickogo-prostoru. [in Ukrainian]

Nature Editorial. (2018). Everyone Needs a Data-management Plan. *Nature*, 555(7696), 286–286. doi:10.1038/d41586-018-03065-z.

NIH. (2002). NIH Announces Draft Statement on Sharing Research Data. *National Institutes of Health*. Retrieved from https://grants.nih.gov/grants/guide/notice-files/NOT-OD-02-035.html.

Normile, D. (2018). China asserts firm grip on research data. *Science*. doi:10.1126/science.aat8311.

Pienta, A. M., Alter, G. C., & Lyle, J. A. (2010). The enduring value of social science research: The use and reuse of primary research data. *MLibrary*. Retrieved from https://hdl.handle.net/2027.42/78307.

Piwowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *PeerJ*, *1*, e175. doi:10.7717/peerj.175.

Plan S. (2018). *Part III: Technical Guidance and Requirements*. Retrieved from https://www.coalition-s.org/technical-guidance_and_requirements/.

re3data. (n.d.). *Official site*. Retrieved from https://www.re3data.org/.

Redalyc L, Clase R, In-Com Uab S. (2003). *Berlin declaration on open access to knowledge in the sciences and humanities. 2003*. Retrieved from https://openaccess.mpg.de/Berlin-Declaration.

SciVal. (n.d.). *Official site*. Retrieved from https://www.scival.com/home.

Scopus. (n.d.). *Official site*. Available at: https://www.scopus.com.

Second French Plan. (2021). Second French Plan for Open Science. *Ouvrir la Science*. Retrieved from https://www.ouvrirlascience.fr/second-national-plan-for-open-science/.

Tarasiuk, A. (2017). Vidkryti dani ta inshi dani u publichnomu dostupi: pravovi aspekty [Open data and other data in open access: law aspects]. *Informatzia i pravo*, 2(21), 59-65. Retrieved from http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Infpr_2017_2_9.pdf. [in Ukrainian]

UNESDOC Digital Library. (2019). *Preliminary study of the technical, financial and legal aspects of the desirability of a UNESCO recommendation on Open Science*. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000370291.

van Wezenbeek, W. J. S. M., Touwen, H. J. J., Versteeg, A. M. C., & van Wesenbeeck, A. (2017). *Nationaal plan open science*. doi:10.4233/uuid:9e9fa82e-06c1-4d0d-9e20-56202259a6c65

Vasylenko, A. Yu. (2019a). The development and implementation of the open science policy in the EU countries: example of France. Public Administration: Theory and Practice, 0(1), 71–77. doi:10.36030/2311-6722-2019-1-71-77.

Vasylenko, A. Yu. (2019b). The establishment and implementation of the state open science policy in the EU countries (example of Netherlands). *Bulletin of the National Academy of Public Administration under the President of Ukraine*, 0(4(95)), 56–62. doi:10.36030/2310-2837-4(95)-2019-56-62.

Vasylenko, A. Yu. (2020). Public open science policymaking in the baltic countries (example of Latvia and Lithuania). *Bulletin of the National Academy of Public Administration under the President of Ukraine*, 0(2(97)), 76–82. doi:10.36030/2310-2837-2(97)-2020-76-82.

Wellcome. (2020). *Sharing Research Data and Findings Relevant to the Novel Coronavirus (COVID-19) Outbreak*. Retrieved from https://wellcome.ac.uk/coronavirus-covid-19/open-data.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., … Bourne, P. E. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1). doi:10.1038/sdata.2016.18.

YODA. (2011). *The YODA project*. Retrieved from https://yoda.yale.edu/.

Zastrow, M. (2020). Open science takes on the coronavirus pandemic. *Nature*, 581(7806), 109–110. doi:10.1038/d41586-020-01246-3.

Zhang, L., & Ma, L. (2021). Does open data boost journal impact: evidence from Chinese economics. *Scientometrics*, 126(4), 3393–3419. doi:10.1007/s11192-021-03897-z.