ABSTRACT

**Scientific objective:** Journals as one of the basic channels of scientific communication should support researchers in the process of open sharing of research data. Their publicity positively affects the quality of scientific research, reduces the costs of conducting them, and promotes the establishment of scientific cooperation. The significance of this issue prompts to carry out research on the strategy of managing research data undertaken by Polish and foreign journals. The paper examines this issue on the example of a group of 198 Polish and 95 foreign journals in the field of historical science. **Research methods:** The strategy for managing research data was examined by analyzing instructions for authors published on the websites of journals representing historical science, which were on the list of journals scored by the Ministry of Science and Higher Education and foreign journals with the Impact Factor index. In the instructions, the author looked for principles for managing research data. **Results and conclusions:** The analysis shows that journals representing historical science are reluctant to adopt the principles of policy for managing research data. This is especially visible in the case of Polish journals, although also among foreign ones with an established Impact Factor indicator, the implementation of appropriate practices is not a common phenomenon. **Cognitive value:** The study shows one of the overlooked aspects of the functioning of scientific journals, especially in the context of discussions on providing open access to scientific publications and research data. It also justifies the need to implement best practices in scientific journals related to sharing research data.

KEYWORDS

journal of historical science, open access, research data
Open access to research data has been an issue widely discussed in the scientific community for several years. Despite some objection from researchers who treat the data they collect almost as private property (Boulton, Rawlins, Vallance, & Walport, 2011), the discussion seems to be moving not towards whether to share them with other scientists, but how to share them using available digital technologies that currently create great opportunities in this area. The argument is the large number and popularity of repositories gathering research data (e.g. Figshare, Dryad). This trend will also be strengthened by scientific policy making research funding conditional on the obligation to provide access to research data generated during the scientific project (“Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020,” 2016).

Research data is a very diverse group of materials that were created as a result of research activities, or were generated not for scientific purposes, but are used by researchers. The Horizon 2020 research funding program, implemented by the European Commission, states that research data "Refers to information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion, or calculation.” For example, they included statistics, results of experiments, measurements, field observations, survey results, interview recordings, and images (Jones, n.d.; “H2020 Program. Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020,” 2017).

The benefits of sharing data are undoubted. Access to them makes it easier to control the research results achieved, minimizing attempts to falsify them (Boulton et al., 2011). In addition, it is associated with giving other researchers the possibility of different interpretations, discovering new research directions, testing hypotheses or developing datasets based on existing ones. The possibility of establishing cooperation between researchers from various institutions and the issue of reducing the costs of scientific work is also important, as the availability of data eliminates the need to collect them again (Uhlir & Schröder, 2007; Tenopir et al., 2011).

However, certain issues that block wider access to data cannot be overlooked. Collecting and developing a resource in the right format takes time, and the publication of such a prepared dataset is often not treated as part of the scientific output, which discourages some researchers from sharing data. In addition, some scientists are afraid of more efficient use of data by other researchers or are concerned about the legal protection of the shared resource (Borgman, 2010; Reichman, Jones, & Schildhauer, 2011).

The development of information technology already allows wider access to research data. Generally speaking, there are at least three methods for publishing data: making them available on a dedicated repository; publication of data in a special type of scientific journal—Data Papers; publication in the form of supplemental material that accompanies the main text (Pampel & Dallmeier-Tiessen, 2014).

Data Papers are scientific journals that publish only descriptions of research data, omitting the interpretation of the collected material. The dataset may be on the publisher’s journal server or may be available in another repository, but the reference in the text should clearly link to it. The publications characterizing the datasets usually undergo a normal review procedure, in which the reviewer assesses the correctness of the methods and techniques used to gather the research resource. Each publication describing a dataset should contain information about how the resource was collected, by whom, when, and where, and who is the owner of the data (Chavan & Penev, 2011). Researchers (Candela, Castelli, Manghi, & Tani, 2015) have identified 116 publications of this type published by 15 different publishers.

In turn, Enhanced Publications (Kędzierska, 2015) are most often scientific journals, in which the published papers are accompanied by additional materials: data, animations,
graphics, videos, etc. This enhancement, however, is implemented to varying degrees in scientific journals. Often, these additional materials are not sufficiently integrated with the main text or are published in a format that hinders their reuse (Pop & Salzberg, 2015). Increasingly, however, publishers encapsulate a paper in a digital version with solutions that make it easier for the user to receive it, e.g. interconnection of various batches of text, placing references to external electronic resources, interactive content, e.g. data tables, software for their processing (Bardi & Manghi, 2014), readers’ comments or connection with social media (Degkwitz, 2014).

In the humanities, the issue of sharing research data is not yet as developed as in the natural or even social sciences (Buddenbohm et al., 2016). Re3data.org has only 242 repositories in the humanities category, which is 10% of all registered resources (“re3data.org—Registry of Research Data Repositories,” n.d.). There is no doubt, however, that the situation will change gradually in this respect, because representatives of the humanities: historians, linguists, literary scholars are increasingly working with digital data, e.g. digitized versions of source texts, digital maps or operate on corpora.

Scientific Objective
Scientific journals are one of the basic channels of scientific communication. Therefore, the growing pressure to provide access to research data raises questions about the strategy for managing research data in this group of publications. In the paper, this issue was examined on the example of a group of Polish and foreign journals representing historical science.

Periodicals were selected for the analysis of Polish journals, the descriptions of which are recorded by the Arianta database (“Arianta—Scientific and Professional Polish Electronic Journals,” n.d.) developed by Aneta Drabek and Arkadiusz Pulikowski. The database contains a total of 202 journals that have been assigned to the “Historical and Archival Sciences” category. The analysis only contained those that were included in the ministerial assessment of scientific journals covering the years 2013-2016, a total of 198 journals (“List of Scientific Journals Containing the History of the Journal from Published Lists for the Years 2013-2016,” n.d.).

Strategies for managing research data in Polish journals representing historical science have been confronted with the principles that apply in foreign journals with the established Impact Factor (IF). Based on the Journal Citation Reports (“Journal Citation Report 2018,” 2019) a group of 95 journals from the “History” category was selected.

The study aimed to determine whether journals representing the area of historical science have a developed strategy for dealing with research data. To this end, instructions for authors published on the websites of the analyzed journals were examined. References to the term “Research Data” and related rules for managing this type of material were sought. The aim was to answer the following questions:

1. Do publishers have a policy for managing research data, and is it clearly formulated on the journal’s website?
2. Does the journal allow one to send supplemental materials to the main text? These are data that for technical reasons could not be included in the paper, but are necessary to understand the text or verify the results of the study (Landesman, 2013). This is also a slightly different form of managing research data, but not always, especially in the case of paid journals, this type of material is generally available.
3. Do publishers require authors to publish research data?
4. Will publishers suggest publishing research data in specific repositories?
An analysis of Polish journals was conducted in September 2017. Incorrect descriptions of resources were corrected in March 2019. The study of journals with an established Impact Factor was carried out between June 25 and July 30, 2019.

**Research Review**

The importance of open access means that scientists often take up in their research the issue of how journal publishers manage research material. A bibliographic query showed that researchers often analyze this issue in relation to journals representing social sciences, while they relatively rarely consider it in the context of the humanities.

Extensive research was conducted by, among others British scientists. As part of the Journal Research Data (JoRD) project, carried out in 2012-2014, researchers analyzed the state of play of research data in 371 scientific journals from various fields of knowledge which were listed on the Journal Citation Report. Their findings show that 50% of journals did not apply any policy, only 24% of identified policies can be classified as so-called strong, i.e. containing a detailed description of how to manage data (Sturges et al., 2015).

In another 2015 project (“Journal Research Data Policy Registry”), which analyzed publishing policy in 250 journals representing various areas of science, researchers noted a slight increase in journals (by 7%), compared to previous analysis, which implemented a policy for managing research data. In the case of social sciences, 40% of journals had this policy. About 30% of publishers imposed on authors the obligation to deposit data in public repositories, but in this group of journals representing social sciences constituted only less than 11%. (Naughton & Kernohan, 2016). The cited analyzes, although covering a wide group of journals, representing various fields of knowledge, do not show what the situation looks like in the case of the humanities.

Herndon and O’Reilly compared the presence of policies of managing research data in 2003 and 2015 in a group of social science journals listed on the Journal Citation Report list (20 journals with the highest Impact Factor index selected from the following categories: finance, economic sciences, international relations, political science, sociology). The authors found that in 2015, in the group they analyzed, 39% of publishing houses have implemented policies for managing research data. Although this percentage is not large, it represents a significant increase compared to 2003, when only 10% had such rules. They also noted that the largest increase occurred in the group of economic journals, among which 65% had a policy for managing research data (Herndon & O’Reilly, 2016).

Analysis of management of research data among the highest-rated journals in the field of social sciences (50 journals with the highest Impact Factor index selected from the following categories: political sciences, international relations, economic sciences, sociology, history, psychology, anthropology) listed on the Journal Citation Report was also carried out in 2017 by a group of researchers from Harvard University. In the group of 291 journals, just over 50% had an implemented policy for managing research data. The largest percentage concerned economic journals (74%), while in the political and psychological sciences group it did not exceed 60%. Journals representing historical science were characterized by the lowest number of data management policies implemented (only 18%). A relatively small group of journals required disclosure of data. In the case of economic sciences, the percentage of such journals was 36%, and in the case of sociology, it did not exceed 10%. Among historical journals, no one required authors to publish research data (Crosas et al., 2018).

A different methodology makes it difficult to compare the results of the discussed texts. Researchers achieved relatively similar results in the case of economic journals. The biggest
discrepancy concerned journals representing political science and international relations. Crosas identified 60% of journals from this group having a policy for managing research data. Herndon and O’Reilly found such a policy only in less than in 33% of journals (Crosas et al., 2018; Herndon & O’Reilly, 2016).

Vlaeminck and Herrmann analyzed a group of 346 economic journals listed on the JOURQUAL 2.1 compiled by the German Academic Association for Business Research and on the list of economic journals created by Bräuninger, Haucap, and Muck. Vlaeminck and Herrmann identified about 35% of journals that have implemented a policy for managing research data. In this group, approx. 21% were journals whose policy was classified by the authors in the “data policy” category imposing an obligation on authors to provide research data for the purpose of replicating the study. The second category were the journals (about 14 percent) using the “data availability policy.” It also requires the authors to provide data, but they accompany the paper in the journal as so-called supplemental material (Vlaeminck & Herrmann, 2015; Bräuninger, Haucap, & Muck, 2011).

Researchers from Korea noted a low percentage of Korean scientific journals that implemented the principles for managing research data. In this respect, only 13% of the editorial staff of the surveyed journals gave a positive answer. But in almost 77% of journal cases only recommended sharing research data without obliging authors to disclose it (Kim, Yi, & Huh, 2019).

Researchers also analyzed how to deal with research data in terms of repeatability. Vlaeminck as part of the EDaWaX project examined 141 economic journals from Austria, Germany, and Switzerland. Only less than 21% of journals have implemented a policy of sharing research data, and even fewer, less than 8% set the rules for sharing data in such a way that it is possible to repeat the study. The author also drew attention to the criteria that should be considered by replication policy, including the obligation to provide research data or a description of how to obtain them, provide information about the formula and software used for data analysis, or provide data in an open format to facilitate long-term storage (Vlaeminck, 2013). In turn, Eubank, researching 24 papers submitted to the Quarterly Journal of Political Science, found that it is impossible to repeat the results of the research in almost 60% of cases (Eubank, 2014). Key also conducted research in this direction. She analyzed nearly 590 paper in the field of political science. Almost 500 contained research data, but the full data, i.e. the material used in the study and the formula used to perform the calculations, were present only in less than 60%. The author’s findings show that full data is more common in journals that have a mandatory policy for sharing research data (Key, 2016).

The researchers also pointed out in their analyzes the relationship between the presence of policy of conduct with research data and the Impact Factor indicator. For example, British researchers have noted that journals with a high Impact Factor rate were characterized by a more specific data management policy (Sturges et al., 2015). Vlaeminck and Herrmann pointed out that journals with implemented data sharing policy in almost 96% of cases had an IF rate set, when among journals without such a policy the percentage of publishers with IF did not exceed 60%. (Vlaeminck & Herrmann, 2015). The relationship between the publication of additional materials accompanying the paper, including research data and the higher IF index, was also noted by researchers analyzing journals in the field of library and information science (Aleixandre-Benavent, Moreno-Solano, Ferrer Sapena, & Pérez, 2016).

The analyzes also concerned types of policies for managing research data. Blahous and others in a group of 534 journals (selected from the SCImago Journal & Country Rank list, research data repositories and websites of scientific societies) representing 17 scientific disciplines identified 65% of those that defined a policy for managing research data. The authors classified
each of the analyzed policies into one of three categories: strong (imposing an obligation on authors to publish research material); poor (publisher postulates sharing data, but this is not mandatory). The third category concerns managing supplemental materials that may contain research data, but the publisher does not impose an obligation to add them to the paper. In the analyzed group, the largest part were journals (over 60%), which contained guidelines regarding supplemental materials. Almost 20% were included in the group of strong policies. In the group of journals representing social sciences, humanities, and art sciences only just over 3% had a strong policy of managing research data (Blahous, Gorraiz, Gumpenberger, Lehner, & Ulrych, 2016). In turn, Lepthien and Zenk-Möltgen, analyzing sociological journals, noticed that only 5% of them had rules for handling research data individually specified for each periodical. For 67% of publications, reference was made to the general principles of data management, which the publisher has defined for all journals it publishes. Almost 28% of sociological journals did not contain any guidelines regarding the management of research data. Researchers also pointed out that in most cases, data management instructions are very general and they do not provide authors with information on how to share them (Lepthien & Zenk-Möltgen, 2014).

The researchers’ attention is also drawn to the issue of supplemental materials. Borrego and Garcia, analyzing the guidelines for authors in journals representing library and information science, determined that 36% of them contained instructions for publishing additional materials (Borrego & Garcia, 2013). Some publishers, due to the large number of such materials reported, forgo their publication or publish only selected ones (Borrego & Garcia, 2013; Landesman, 2013). They are not always reviewed or reviews are very cursory (Pop & Salzberg, 2015).

The researchers also point out (Candela et al., 2015) that publishers are getting away from the practice of adding materials to papers in favor of publishing them in external repositories. The repository must meet certain specific conditions, e.g. enable long-term storage or be generally available.

The results of the cited analyzes, although based on different methodologies, indicate that an increasing number of publishers are trying to regulate access to research data to varying degrees. Although publishers recommend rather than order authors to share the collected research material.

**Analysis Results**

**Polish Journals Representing Historical Science**

In total, the websites of 198 Polish journals were reviewed, which were classified in the Arianta database as “History and Archival Sciences.” Some publications had an interdisciplinary character. Less than 13% of them were included in the category of “Cultural Studies.” The group consisted of journals in the field of philological sciences, political sciences, religious studies, and art did not exceed 6%.

The publishers of the analyzed journals were mostly universities (less than 39%) and scientific societies (about 28%). Institutes of the Polish Academy of Sciences published less than 14% of journals. The publishers also included museums, archives, Polish Academy of Arts and Sciences, and research institutes.

More or less extensive instructions intended for authors appeared in 175 journals (88.38%). In the case of journals registered on the ERIH list, only 4 out of 28 (14.28%) did not contain any guidelines for the authors.

None of the instructions analyzed included any reference to the issue of research data, even by recalling this term. Therefore, publishers did not necessarily raise the issue of depositing data in external repositories, nor did they condition the acceptance of the paper from the publication of research materials.
This does not mean, however, that the authors of the texts published in Polish journals representing historical science do not publish in repositories. In Figshare one can identify, among others materials from the *Kultura–Historia–Globalizacja* (Eng. *Culture–History–Globalization Journal*) and *Kultura i Historia* (Eng. *Culture and History Journal*). However, these are full texts, not research material used in the study (Hylewski & Burdzik, 2014).

The guidelines for authors concerned mainly the formatting of the text to be submitted to the journal, the method of preparing footnotes and bibliography. In addition to the remarks mentioned above, the publishers repeatedly raised the issue of preparing additional materials (illustrations, tables, charts), which were to supplement the main text. In the surveyed group, 80 publishers (40.40%) recommended sending separate materials supplementing the main text.

It should be added that almost 35% of publishers recommend sending high-resolution graphic materials, and less than 20% ask for files with statistical data, which makes these materials a potentially valuable source of data, which after processing can be made available on the website of the journal or in an external repository. In this case, it is probably more about issues related to preparing the text for printing, and not about making the data available for reuse, as publishers do not mention this in the instructions for the authors.

**Foreign Journals Representing Historical Science**

In the group of 32 publishers, most journals representing historical science (19 each) are published by Taylor and Francis Group (T&FG) and Cambridge University Press (CUP). Publishers also include Oxford University Press, Sage Publications, John Wiley & Sons, Brill, Edinburgh University Press. All of the analyzed publications published instructions for authors on their websites.

In total, 13 of 95 journals (less than 14%) defined the rules for managing research data, including in the sections for authors called Data Sharing Policy (T&FG), Data Sharing and Accessibility (John Wiley & Sons), Data Availability Policy (CUP). They differ in the detail of the description. The most extensive instructions for authors in the section devoted to research data can be found on the websites of journals published by T&FG.

But even if the publishing house has comprehensively solved the policy of dealing with research data, not all journals apply it. Only 11 out of 19 T&FG journals analyzed have implemented specific policies for handling research data. Springer Nature has developed five types of data handling policies, but none of them has been implemented in the *Cliometrca* journal (“Research Data Policy Types | Authors | Springer Nature,” analyzed in the text). Therefore, publishers leave the editors of historical journals some freedom in shaping the practices of handling research data.

In none of the 13 journals did the editors make publication of the text conditional on the obligation to publish research data. They only encouraged sharing data, enumerating the benefits of making it public.

T&FG introduced a new policy on how to handle research data in 2018 (“Understanding Our Data Sharing Policies,” 2019). The publisher has implemented 5 types of policies for managing research data. The least restrictive is the basic one used by all 11 journals representing historical science. The authors are encouraged to share data, there is no obligation to add a statement on the availability of data or assign a DOI number to the collection. The author’s decision also depends on which license the dataset will be made available for. For other policies, e.g. adding a data availability statement is mandatory. The publisher has prepared 10 types of statements that specify where and under what conditions the data will be available to the user, e.g. after the embargo period expires, at the user’s request or that the data comes from the public domain (“Data Availability Statement—Statement Layout Template,” 2017).
The most open policy assumes full access to research data under Creative Commons license (CC0 or CC BY) and compliance with Fair Data Principles (“Data Sharing Policies,” n.d.). FAIR principles have been formulated by a group of archivists, librarians, scientists, and publishers gathered in the FORCE11 organization working for changes in the creation and sharing of knowledge. They define the principles of such data handling so that they are easy to find, available, interoperable and re-usable (FAIR—to be Findable, Accessible, Interoperable, Reusable) (Wilkinson et al., 2016).

In the group of 19 journals published by CUP, only the “International Review of Social History” defined the rules for managing research data, but did not do it in such an extensive way as T&FG. The Data Availability Policy indicates that data should be archived in repositories that have a Data Seal of Approval certificate (replaced by CoreTrustSeal from 2017) (Leeuw, 2019) certifying that the repository guarantees appropriate data storage and the possibility of their re-storage use. The CUP also notes the appropriate documentation that accompanies the data that will allow the study to be repeated. By default, the data is made available under the Open Access license (“Instructions for Contributors,” n.d.).

The policy of sharing data was briefly defined by the Australian Journal of Politics & History by John Wiley and Sons, encouraging only authors to publish research data in relevant repositories and to add a statement on data availability (Australian Journal of Politics & History, n.d.).

A slightly larger group of journals are those that in the instructions for authors try to regulate the issue of supplemental materials to the main text—33 out of 95 (almost 35%) journals provide the ability to send and publish materials supplemental the main text. Also 11 journals published by T&FG, which have implemented a policy of sharing research data, allow sending additional materials to the text.

Sometimes instructions for sending this type of data are accompanied by information about the benefits of sharing research data, such as increasing the chance of other researchers finding the paper and more citations (“Enhancing Your Article with Supplemental Material,” 2015). Some publishers (Journal of British Studies, Journal of Global History) emphasize that sending data serves the possibility of repetition of the study (“Information for Contributors,” n.d.) (Journal of British Studies, n.d.).

For publishing additional data on Figshare, or on SP (“Supplemental Material—Guidelines for Authors,” 2015), 5 out of 7 journals from Sage Publications encourage, but in the latter case data may not be available to the general public if the journal requires purchase a subscription.

None of the publishers requires authors to publish research data in a specific repository. However, some (Brill, T&FG and SP) have created their own subdomains on Figshare to publish such data (“SAGE Journals Data Repository,” n.d., “Taylor & Francis Group Data Repository,” n.d., “Brill Online Data Repository,” n.d.).

Summary
The analysis shows that journals in the field of historical science are reluctant to adopt the principles of policy for managing research data. This is especially visible in the case of Polish periodicals, although also among foreign journals with an established Impact Factor indicator, the implementation of appropriate practices is not a common phenomenon. The analysis carried out and the results of other studies discussed have shown, for example, that journals representing historical science differ significantly in sharing research data from publications representing social sciences.

The aim of the study was not to identify the causes of the described issue, but it seems that the reason may be some ignorance of this issue among publishers, as well as technological and
financial barriers that are difficult to overcome for the scientific community, especially in the area of the humanities. Already in 2012, the European Commission in its analysis devoted to access to scientific information in digital form, pointed out that for publishers the most important barrier in providing access to research data is the lack of money to finance the appropriate infrastructure necessary for their storage. Another significant obstacle is the lack of appropriate national or regional strategies regulating the management of research data (European Commission & Directorate-General for Research and Innovation, 2012). If we add to this certain concerns of the scientific community related to the publication of data, we will get a picture of numerous issues that hinder the process of opening research data. However, there is no in-depth study that would analyze the problem of sharing data from the perspective of various humanities disciplines.

Polish journals have a long way to go before joining the increasingly widespread current of open access. For now, this issue is overlooked, even by periodicals with a relatively high position in the scientific community.¹ Therefore, it seems necessary to start work in this area, because the amount of digital data, also in the humanities, will constantly increase and the pressure of the scientific community on ensuring access to this data will grow. For publishers, T&FG can be a role model, which has prepared several types of policies for managing research data and clearly explained to the authors the rules for sending and sharing materials collected by them.

It will take some time to prepare and implement appropriate policies for managing research data, but publishers can now encourage authors to share collected materials with other users and identify the benefits of doing so. This should be followed by the promotion of repositories, along with tips on how to share data in them. In addition to the Figshare website referred to in the paper, Polish researchers can, for example, use RepOD—Open Data Repository (“About the Repository—RepOD,” n.d.), and historians dealing with the history of old Poland have at their disposal the Atlas Fontium portal (AtlasFontium, n.d.).

Bibliography


¹ Polish publishers are not even fully using the opportunities offered by the popular Open Journal System (OJS), which is used by less than 10% of the editorial staff. The system allows convenient management of additional materials supplementing the main text. Little more than half of the publishers using OJS (less than 56%) recommend authors to send this type of material separately (tables, illustrations, charts).


Leeuw, L.D. (2019). Data Seal of Approval (DSA) [Dataset]. https://doi.org/10.17026/dans-28z-njx


