doi: 10.26114/pja.iung.454.2020.43.12

**Polish Journal of Agronomy**
**2020, 43, 125–131**

# The importance of „Open Science" for agriculture

*Monika Kowalik*

Department of Extension and Information, Institute of Soil Science and Plant Cultivation – State Research Institute
ul. Czartoryskich 8, 24-100 Puławy, POLAND

**Abstract.** Both agriculture as an research area and its closely related life sciences have developed rapidly over recent decades. The amount of knowledge that comes from day to day is vast and we are in the "Agriculture 4.0" era not by accident. Researchers representing agricultural science and research audiences often face choice – is open knowledge sources and data are appropriate source for their research or not. The discussion on the scope of open science resources is becoming intense, even though these initiatives are not new. The European Commission's position is clear – the results of EU-funded research should be as widely disseminated as possible. This will allow to have a broader discussion and develop new knowledge from existing data. It also applies to data relating to agricultural science. The study provides an overview of the current literature and relevant data sources available on the open science web for agricultural science. The article also describes harmful practices related to open access.

**Keywords:** open science for agriculture, open access, AKIS, open air, FAIR data

## INTRODUCTION

The purpose of the article is to organise the current knowledge regarding data sources and tools useful in agricultural sciences, which can be used in the open access formula. The beneficiaries of this solution, apart from scientists, are also experts, specialists, agricultural advisors providing services both in the state and private system, as well as farmers themselves, university or agricultural schools students. The paper presents not only domestic sources but also European ones and global solutions.

According to the research conducted by Harasim (2020), in the process of transferring knowledge, and especially the results of the latest studies, the system of state advisory services for agriculture still plays the dominant role, the advantages of which are: universality, non-profitability and socialisation. An innovative form of access to free data repositories and articles facilitates access to new knowledge and development of skills – both for scientists and advisers (O'Carroll et al., 2017). Such solutions are conducive to creating new knowledge and allowing expert discussions on the research already carried out and setting new directions; therefore, related to the demand for innovation (Lowndes et al., 2017).

The dominant direction in the development of agricultural advisory services was outlined in the study of the Standing Committee on Agricultural Research, emphasising the need for changes in the system of access to data for agriculture. Currently, a considerable portion of the information is unusable. Many public and restricted access data sets are difficult to decode due to the fact that their structure is unclear, and the metadata and contextual information are often missing. Parameter names are ambiguous or non-existent, units of measurement are omitted, and documentation is missing too. Information is often scattered across various data sources, using different formats. This situation is due to the fact that data is usually collected for a specific purpose, and no attention is paid to its preparation for use by other entities in the future. Depos-

List of abbreviations used in the text

| Acronym | Expanded form |
|---|---|
| AKIS | Agricultural Knowledge and Innovation Systems |
| FAIR | FAIR data are data which meet principles of findability, accessibility, interoperability, and reusability |
| OA | open access |
| OS | open science |
| RRI | Responsible Research and Innovation |

Corresponding author:
Monika Kowalik
e-mail: mkowalik@iung.pulawy.pl
phone: +48 81 4786 733

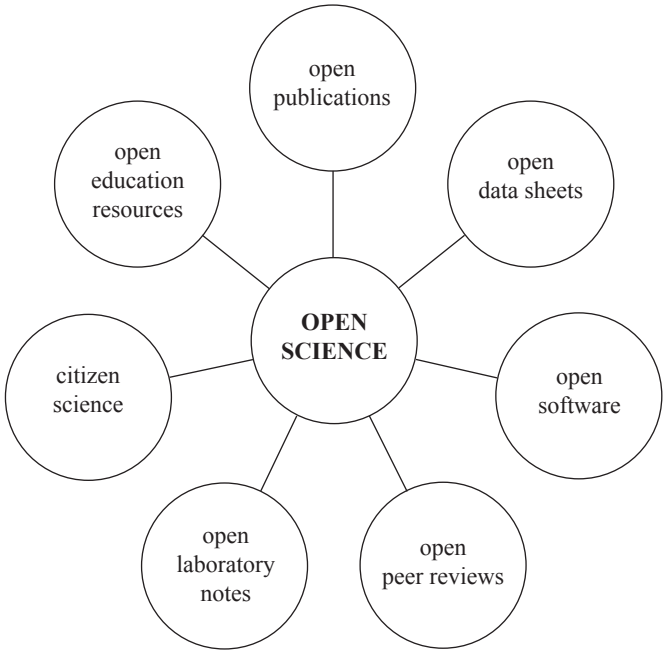


Figure 1. Scheme of activities within the Open Science initiative.
Source: Own elaboration.

iting information in dedicated databases (e.g., LUCAS for land use), where it is appropriately structured and described, promotes making greater use of the data obtained. Advanced tools available there allow for verifying, updating and grouping records and are even used by the original authors of the data (EU SCAR 2016).

## WHAT IS OPEN SCIENCE?

S. Hampton provided a clear definition of Open Science as a concept of transparency at all stages of the research process, combined with free and open access to data, software source code and documentation (Hampton et al., 2015). The need for scientists to share knowledge resources appeared as early as the 17th century, but the Open Science (OS) movement has seen dynamic development since the 1990s (Kuiper, Roeling, 1991). Its activity significantly contributes to the development of knowledge, exchange of scientific ideas, exchange of research results and experiences, and thus to scientific progress.

The Open Science movement includes several important activities, presented in Figure 1. Apart from open access to data (OA), its other elements are gaining importance, including citizen science, or sharing source data in scientific publications.

Successful promotion for the achievements of thousands of scientists worldwide in order to make them available to the widest possible audience has been observed for decades. This also applies to agricultural sciences and other natural sciences. Most of the data is collected at the request of the ministries responsible for scientific institutes, universities, relevant offices and institutions, and financed from public funds. Often, the data stored by various institutions is not widely available for reasons incomprehensible to the taxpayer. These are the fundamental problems of individuals or organisations trying to establish real and mutual cooperation, e.g., entities included in the national Agricultural Knowledge and Innovation Systems (AKIS). Much of this information must be made available with due regard to established principles, such as respecting the author's rights to conduct research. Open Science issues are closely related to the strategies and programs of the European Commission, including the principles of Responsible Research and Innovation (RRI). This is a relatively new approach in the context of European research and innovation policy, which aims to balance economic, socio-cultural and environmental aspects in innovation processes (Wilford et al., 2016). In the age of information overload and the growing number of scientific articles in various fields (Jinha, 2010), the quality of scientific research is of particular importance. The possibility of discussion in the framework of Open Science may contribute to increasing the value of this research.

The opening of research processes and making the research cycle transparent are fundamental requirements of Open Science. This involves not only collaborating in the use of research infrastructures which is concentrated around scientists, their work and their copyrights, but also with those institutions which have to review research results: libraries, research evaluation bodies and funding organisations. This leads to two principal challenges: possession of credible means allowing for linking researchers to their work (and with some other elements such as affiliation, academic activity, social activity) and how to present these linkages. In this situation, the provision of unique identifiers, such as ORCID, is a useful solution. ORCID is a register which, in its primary function, provides a unique identification number, the so-called ORCID ID. ORCID identification numbers are available to anyone who contributes to research, scientific projects and innovation. Natural persons perform the registration and create a personal ORCID record. The obligation to register an ORCID identifier in Poland has also applied to scientists in the field of agriculture for almost 2 years. Therefore, it is much easier to identify, e.g. an article and assign it to an appropriate author and unit (Dudek et al., 2019).

The first step towards defining the principles of open access should be the adoption of proprietary institutional policy of open access by each research unit and university, which should take into account the basic principles of data sharing, the so-called FAIR data. FAIR data, i.e., FAIR data, takes into

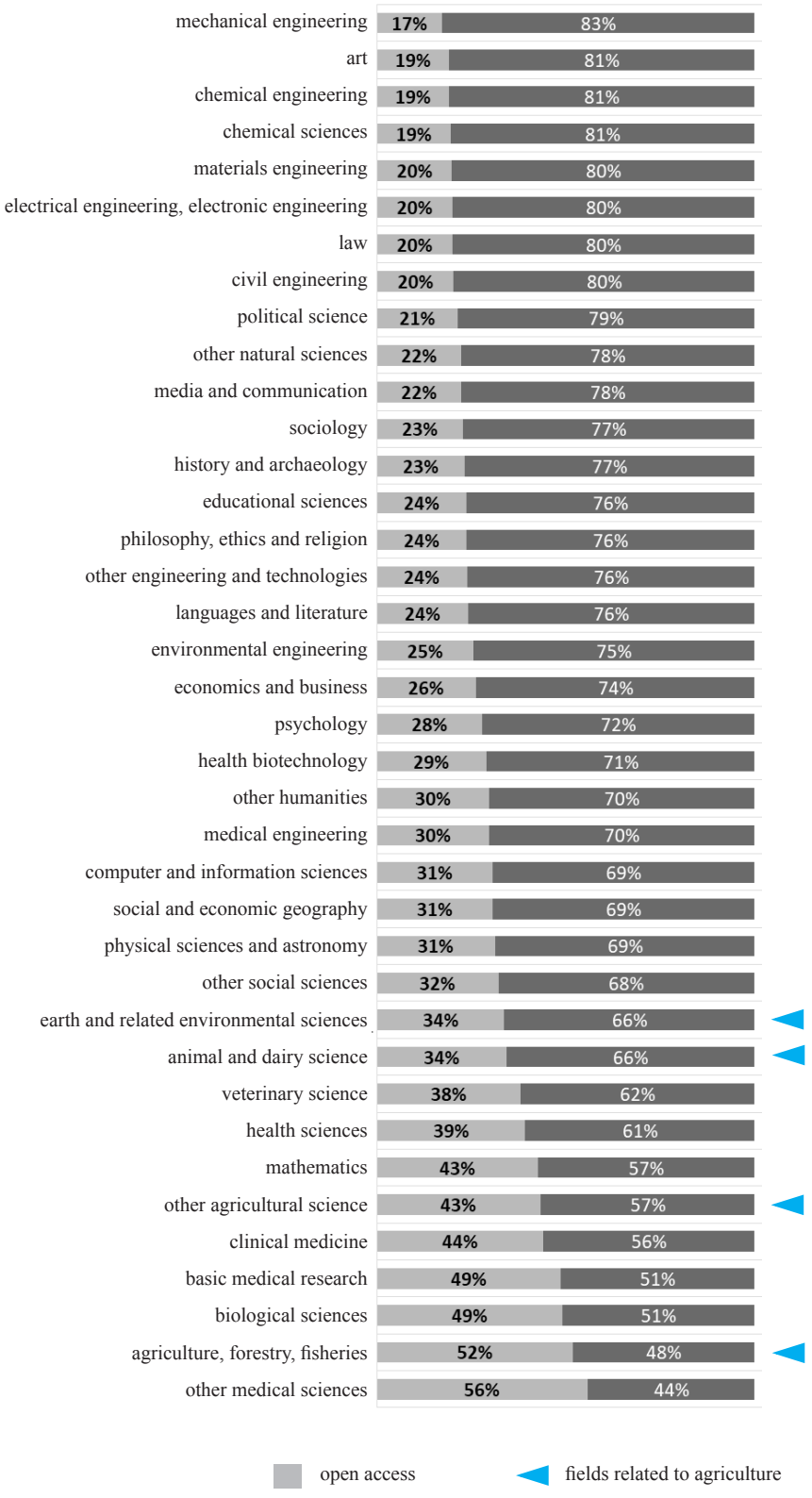


Figure 2. Percentage of Open Access publications by Fields of Science and Technology
Source: authors's compilation based on EC data (OA EC 2020)

account four basic principles, i.e., it must be: findable, accessible, interoperable and re-usable (EU SCAR 2019). One of the examples of the use of FAIR data is the European Commission Open Air website, offering a range of source data created during the implementation of projects financed by the European Commission (Open Air 2020).

One of the arguments for placing source data in open repositories which allow users to conduct verification is counteracting plagiarism. Such practices have been applied for many years by large publishing companies such as Elsevier or Scopus.

Open Access to publications refers to the possibility of having free access to scientific publications. They are divided into the following categories:

– Golden Open Access: the results of scientific research that are published in an open access journal,
– Green Open Access: research results that are published in a journal also available in an open access repository,
– Hybrid Open Access: research results that are published in a subscription-based open access journal with a transparent licence,
– Bronze Open Access: Research results that are published in a subscription journal and are available without a licence.

According to the data of the European Commission (OA EC, 2020), over the years 2009–2018 in the field of agriculture, forestry and fishery, 379,206 articles outside the OA and 409,475 articles in OA were published, where individual categories were respectively: Bronze OA – 8.59%, Hybrid OA – 3.84%, Green OA – 35.50%, Gold OA – 34.92%. In the Animal and Dairy Science category, outside OA 258420 (66%) and 134273 (34%) were published, where Bronze OA – 11.89%, Hybrid OA – 3.81%, Green OA – 16.02%, Gold OA – 12.45%. In the analyses of the European Commission in areas related to agriculture, one can enumerate the following categories: agriculture, forestry, fisheries, other agricultural sciences, animal and dairy science, Earth sciences and related environmental sciences. The share of each category is presented in Figure 2.

## Types of data and repositories, as well as OS examples for agriculture

A. Maps and cartographic data

One of the largest Polish portals offering free maps and databases is GEOPORTAL (Geoportal 2020). Among the data sets provided by the portal, the following layers should be listed:

– National Boundary Register
– Land and building registry
– Spatial development plans
– Digital terrain model.

B. Satellite hyperspectral imagery and derived data

The opening of repositories with already developed satellite data for users who want to use it in research, analysis or algorithms is a breakthrough in accessing data. Google Earth offers several layers of already developed satellite and aeronautical data sets which are available worldwide (Google Earth) in various resolutions. In particular, Google Earth Engine seems to be particularly useful for research purposes (Teluguntla et al., 2015). Among the numerous layers, the data on agricultural parcels and farmland for the whole world (GFSAD) is available. GFSAD is a NASA-funded project the aim of which is to provide high-resolution global data on farmlands and their water supply. GFSAD layers are obtained using multispectral remote sensing data (e.g., Landsat, MODIS, AVHRR), secondary data and data from farmlands and they are all intended to document the dynamics of the phenomena occurring in the growing season. At a nominal scale of 1 km, V0.1 provides a spatial layout of a disaggregated 5-class map of global farmland coverage. There is no information on the type of crop or the dominant crop. Farming intensity (annual, biennial, triennial and permanent crops) can be obtained for each pixel by using remote sensing data in a time series. The GFSAD1000 product from 2010 is based on the data from 2007-2012. High-resolution satellite data, including imagery from the following satellites: Sentinel-1, Sentinel-2, Sentinel-3 and Sentinel-5P, which are supervised by the European Space Agency (ESA) is also available in the OA formula on the platform of Copernicus Open Access Hub (Copernicus, 2020).

C. Computing clouds (e.g., Google, Amazon) and artificial intelligence (AI), including image recognition.

These technologies are developing highly dynamically. Powerful providers of cloud services often make them available for free to a limited extent. In this way, the user enjoys an opportunity to learn about the possibilities of modern tools, and the provider can verify the operation of their service and possibly collect data. Applications related to, for example, the recognition of plant pathogens are becoming more and more popular, although they are rather unavailable in the open access formula. The aforementioned services are mostly exploited in scientific branches related to agriculture and agricultural sciences in the scope of data analysis.

D. Digitised library resources and digital repositories

Works related to digitisation of existing digital resources have been carried out for many years. Among others, they are conducted by National Institute of Rural Culture and Heritage (in Polish: Narodowy Instytut Kultury i Dziedzictwa Wsi) – former Central Agricultural Library (CBR). According to the CBR data, in 2019 there were 125 electronic agricultural journals made available

by Polish publishers in the OA formula (CBR, 2019). The „Polish Journal of Agronomy" published by IUNG-PIB or the „Annals of Animal Science" of the National Research Institute of Animal Production in Krakow can be given as an example, but also a number of journals published in Polish, e.g., „Agricultural Advisory Issues" (in Polish: Zagadnienia Doradztwa Rolniczego) available on the website of the Agricultural Advisory Centre in Brwinów. A number of scientific institutions make not only their own journals available in a digital form, but also other publications.

E. Other national services

For many years, data and reports on agriculture have been available on the website of Statistics Poland (formerly Central Statistical Office; in Polish: Główny Urząd Statystyczny). It is possible to download data and reports from the Local Data Bank (BDL), which in the category of agriculture contains several sections, including data on utilised agricultural areas, archival data from general agricultural censuses, data on purchase prices, agricultural production, animal stock headcount (BDL 2020).

The website dane.gov.pl pursues the aim of the Central Repository of Public Information, indicated in the Act on Access to Public Information (Journal of Laws, 2020) as one of the modes for access to and re-use of public information. It contains 65 OA digital data sets for the agriculture category (as of 2020).

F. International websites operating in the open access formula

The re3data.org (re3data.org, 2020) website offers a range of open access raw data, including data in the agriculture category. The repositories are divided according to thematic categories and data types of the areas they come from.

Global Open Data for Agriculture and Nutrition (GODAN) also offers free and open access to data indirectly related to agriculture and to the repository of electronic publications – The Open Knowledge in Agricultural Development (OKAD) (goDAN, 2020).

## Negative effects of open data access

There occurs an increasing number of opinions on the deviated practices of OA application on the websites of widely recognised publishers and on the Internet. These include reports on the so-called predatory magazines. 1316 of them were selected and placed on the list published on the website predatoryjournals.com, which was initiated by the so-called Beall's List, created until 2017 by an employee of the Library of the University of California – Jeffrey Beall. Out of this group, 22 journals are closely related to agriculture (containing the prefixes agri-, agro-, farm- in the journal's title) (the authors' calculations based on LPJ, 2020). The criteria for considering a journal to be potentially predatory include:

- Excessive prices for publishing an article coupled with a lack of review or editorial supervision.
- Notifying authors about fees only after the approval of an article.
- Information sent to scientists by mass mailing spam in order to persuade them to publish or take positions on the editorial boards of journals.
- Swift approval of low-quality works, including fake ones.
- Placing scholars on the list of editorial board members without their consent or preventing their resignation.
- Putting names of false scholars as editorial members or authors on the list.
- Copying the graphic design and linguistic content of marketing materials as well as websites of legal, recognised magazines.
- Frauds or misuse of ISSN numbers.
- Providing false information on the location of the publishing operation.
- Falsified impact factors (IF).

Also, there appears and increasing number of reports and opinions from reputable scientists about reprehensible OA practices. As Poppema (2020) states, they lead to, among others, creating the so-called citation cartels where journals, authors and institutions conspire to over-cite scientific articles. In this situation, a huge responsibility rests on the editorial offices of journals and reviewers who should ensure that the quality of research is maintained. In each case of sharing data, not only for the purposes of agricultural sciences, it is indispensable to define:

- the manner in which the data will be collected,
- how the data will be disseminated,
- how the data will be archived for long-term protection of resources,
- how data can be appropriately reused, while respecting applicable regulations, including intellectual property rights (Abbà et al., 2015).

## Plan S

Plan S is an initiative launched in 2018 with the support of the European Commission and the European Research Council (ERC). Its main goal is to implement Open Access as soon as possible and introduce ten principles supplemented with detailed guidelines for scientific articles related to data openness (ESF, 2020; ON ,2020). The implementation consortium also includes the Polish National Science Centre among the other, 11 national agencies funding research that form the "Plan S coalition".

Since 2021, all scientific publications describing the results of research funded by public or non-public grants from national, regional or international funding institutions

and research councils must be published in open journals or on open platforms, or be made available immediately in open repositories, without temporary embargo.

## CONCLUSION

The dispersion of "suppliers" and national data sources still constitutes a problem for Polish agricultural advisory services, representatives of science and other AKIS members. Even if they are made available, integrating them is an enormous challenge, both in terms of time and money. Successive attempts to unify and create common platforms related to agriculture have not brought sufficient results, although new initiatives related to digitisation and attempts to design a common information platform in the framework units supervised by the Ministry of Agriculture and Rural Development are still being taken.

Open repositories have gained particular importance at the time of the threat and pandemic caused by the SARS-CoV-2 virus. Large groups of people working for agriculture have not always had the opportunity to work on company computers offering access to databases or repositories. Digitised open resources have greatly facilitated work under these specific conditions.

OS is one of the most important elements connected with the reduction of data access barriers for researchers, active citizens, pupils, students, teachers, experts and other people associated with agriculture. The next step in disseminating the already existing data in the OS system is to make the information about the existing resources as widely publicised as possible using media that provide broad access, such as websites or social media. Additionally, the attention of the entities responsible for data should be drawn to the widest possible presentation of summaries in the form of appropriately marked repositories (tagging). Owing to this solution, popular search engines, and thus final recipients, will have easier access to open data on Polish agriculture.

On the basis of the above-mentioned examples related to OS, it may be legitimately concluded that OS is essential for agriculture. The widest possible availability of data and tools may only translate into an increased public awareness in a beneficial way, and this may impact not only the economic efficiency of farms, but also the awareness of the great responsibility for the environment that rests on modern agriculture.

## REFERENCES

**Abbà S., Birello G., Vallino M., Perin A., Ghignone S., Caciagli P., 2015.** Shall we share? A repository for Open Research Data in agriculture and environmental sciences. EPPO Bulletin, 45(2): 311-316, doi: 10.1111/epp.12212.

BDL. Bank Danych Lokalnych, Główny Urząd Statystyczny, https://bdl.stat.gov.pl/BDL/dane/podgrup/temat (accessed: 29.12.2020).

CBR. Polskie e-czasopisma: Bezpłatne pełne teksty naukowe w internecie - open access, https://cbr.gov.pl/index.php/polskie-e-czasopisma, 2019 (accessed: 29.12.2020).

Copernicus. The Copernicus Open Access Hub Europejska Agencja Badań Kosmicznych, https://scihub.copernicus.eu/ (accessed: 29.12.2020).

**Dudek J., Bergmans J., Costas R., 2019.** ORCID: Open science monitor case study. Directorate-General for Research and Innovation, European Commission, Luxembourg ISBN 978-92-76-12547-1, doi: 10.2777/050794.

ESF, 2020. What is cOAlition S? European Science Foundation, https://www.coalition-s.org/ (accessed: 29.12.2020).

EU SCAR, 2016. Agricultural knowledge and innovation systems towards the future: a foresight paper. Brussels.

Geoportal. Geoportal Infrastruktury Informacji Przestrzennej Główny Urząd Geodezji i Kartografii, https://mapy.geoportal.gov.pl/ (accessed 29.12.2020).

goDAN - Global Open Data for Agriculture & Nutrition, https://www.godan.info/, 2020 (accessed: 08.12.2020).

**Hampton S., Anderson S.S., Bagby S.C., Gries C., Han X. et al., 2015.** The Tao of open science for ecology. Ecosphere 6, 7, art120, doi: 10.1890/ES14-00402.1.

**Harasim A., 2020.** Potrzeby z zakresu doradztwa w opinii rolników. pp. 129-135. In: Uwarunkowania i perspektywy rozwoju produkcji rolniczej w różnych regionach Polski: Program wieloletni 2016-2020. Wspieranie działań w zakresie ochrony i racjonalnego wykorzystania rolniczej przestrzeni produkcyjnej w Polsce oraz kształtowania jakości surowców roślinnych; eds: Kopiński J., Krasowicz S., Matyka M., Puławy, ISBN 978-83-7562-326-0.

**Jinha A., 2010.** Article 50 million: an estimate of the number of scholarly articles in existence. Learned Publishing, 23(3): 258-263, doi: 10.1087/20100308.

Journal of Laws of 2020, item 2176 [2020. Ustawa z dnia 6 września 2001 r. o dostępie do informacji publicznej, Dz. U. z 2020 r. poz 2176].

**Kuiper D., Roeling N., 1991.** The edited proceedings of the European seminar on knowledge management and information technology. 23rd and 24th November 1989, Wageningen, the Netherlands. Agricultural University, Wageningen, ISBN 90-72525-11-6.

**Lowndes J., Best B.D., Scarborough C., Afflerbach J.C., Frazier M.R. et al., 2017.** Our path to better science in less time using open data science tools. Nature Ecology & Evolution, 1, 0160, pp. 1-7, doi: 10.1038/s41559-017-0160.

LPJ. List of Predatory Journals, https://predatoryjournals.com/journals/, 2020 (accessed: 28.12.2020).

OA EC. Trends for open access to publications Komisja Europejska, https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science/open-science-monitor/trends-open-access-publications_en, 2020 (accessed: 29.12.2020).

**O'Carroll C., Kamerlin C.L., Brennan N., Hyllseth B., Kohl U. et al., 2017.** Providing researchers with the skills and competencies they need to practise Open Science. Open Science Skills Working Group Report. Directorate-General for Research and Innovation, European Commission, Publications Office of the European Union, Luxembourg, ISBN 978-92-79-69742-5, doi: 10.2777/121253.

OD. Otwarte Dane Centralne Repozytorium Informacji Publicznej, https://dane.gov.pl/, 2020 (accessed: 29.12.2020).

ON. Platforma Otwartej Nauki ICM UW, https://otwartanauka.pl/, 2020 (accessed: 29.12.2020).

Open Air, https://www.openaire.eu/, 2020 (accessed: 28.12.2020).

**Poppema S., 2020.** Citations Are the Currency of Science. https://www.the-scientist.com/critic-at-large/citations-are-the-currency-of-science-68204 (accessed: 28.12.2020).

re3data.org., 2020. re3data.org - Registry of Research Data Repositories, https://www.re3data.org/ (accessed: 29.12.2020).

**Teluguntla P.G., Thenkabail P.S., Xiong J.N., Gumma M.K., Giri Ch. et al., 2015.** Global Cropland Area Database (GCAD) derived from Remote Sensing in Support of Food Security in the Twenty-first Century: Current Achievements and Future Possibilities. pp. 131-159. In: Land resources monitoring, modeling, and mapping with remote sensing; ed. Thenkabail P., ISBN 9781482217957.

**Wilford S., Fisk M., Stahl B., 2016.** Guidelines for Responsible Research and Innovation. Centre for Computing and Social Responsibility, De Montfort University, Leicester.

Artykuł opracowano w ramach realizacji zadania 2.7 Programu Wieloletniego IUNG-PIB

| Author | ORCID |
|---|---|
| Monika Kowalik | 0000-0001-9272-3246 |

received – 16 December 2020
revised – 27 January 2021
accepted – 29 January 2021