The 15th National Information Day: Open Science, Open Data, Open Access, Bulgarian Open Science Cloud

Peter Stanchev[1, 2][0000-0002-0063-0024], Aneta Karaivanova[3][0000-0002-6493-7981],
Yanita Zherkova[4][0000-0003-2681-2034], Mikaela Stancheva[4][0000-0003-2389-2702],
Jordan Iliev[5][0000-0002-0218-8540], Desislava Paneva-Marinova[1][0000-0001-5998-687X],
Georgi Simeonov[1]

[1] Institute of Mathematics and Informatics - Bulgarian Academy of Sciences, Sofia, Bulgaria
[2] Kettering University, Flint, USA
[3] Institute of Information and Communication Technologies –
Bulgarian Academy of Sciences, Sofia, Bulgaria
[4] Ministry of Education and Science, Sofia, Bulgaria
[5] National Centre for Information and Documentation, Sofia, Bulgaria
stanchev@math.bas.bg, anet@parallel.bas.bg, y.zherkova@mon.bg,
m.stancheva@mon.bg, j.iliev@bpos.bg, dessi@cc.bas.bg,
gsimeonov@math.bas.bg

Abstract. The paper gives an overview on the current landscape and the activities on both national and institutional level regarding Open Science, Open Access to scientific information, Open Data, and the Bulgarian Open Science Cloud.

Keywords: Open Access, Access to Knowledge, Open Science.

1 Introduction

The 15th National Information Day is dedicated to demonstrating successful open access (OA), open data sharing and open science projects and initiatives; strategies for the promotion of OA; advancements in the Bulgarian Portal for Open Science and the national repository; copyright and licenses under OA; best practices and examples of policies for the development and promotion of open access at national and institutional levels, etc. The event will support academic staff, research project managers, policy-makers and staff working in research funding bodies. It will demonstrate the main concepts for the implementation of OA principles and propose a national OA roadmap, key terms, strategies, business models, work plans, copyright and licenses, best practices, examples and policies for OA, open data and open science.

The European Open Science Cloud (EOSC), the Common European Data Space for R&I, is the EU's flagship initiative for the digital transformation of research. EOSC is a key element of the EU's Open Science Policy. It is recognised as one of the Common European Data Spaces and is therefore a main component of the European Strategy for Data, enhancing the EU's leadership in the global data economy. Open Science is at the

centre of European research policy. Open Science has a wide spectrum of axes, from open access to publications, FAIR and open data, open-source software, sharing services and e-infrastructures outside the domain they have been created for, and by engaging the wider public through citizen science projects.

OpenAIRE AMKE is a non-profit organization with a mission to promote open scholarship and improve discoverability, accessibility, shareability, reusability, reproducibility, and monitoring of data-driven research results, globally. The organization operates a European e-infrastructure offering a diverse set of public services to accelerate the adoption of Open Science and is supported by a network of experts placed in key national organizations across European countries, the National Open Access Desks. The users of OpenAIRE services include researchers, research communities, policy makers, research-intensive organizations, SMEs, universities, libraries, and citizen scientists. OpenAIRE is also a key implementer of EOSC policies.

2   Bulgaria OA Journals, Repositories and Portals

Data for the Open Access Bulgarian repositories in 2024 are shown in Figure 1.

Open Access Bulgarian Repositories 2024	Software	Link
New Bulgarian University Repository	EPrints 3.4.3	http://eprints.nbu.bg
BulDML at IMI-BAS	DSpace 7.3	http://sci-gems.math.bas.bg
Digital repository of Sofia University	DSpace 1.8.2	https://research.uni-sofia.bg
D.A.Tsenov Academy of Economics - Digital Library	DSpace 6.3	http://dlib.uni-svishtov.bg
Academic Research Repository at the Burgas Free University	Dspace 1.8.1	http://research.bfu.bg:8080/jspui/
Central Medical Library - MU, Sofia	DSpace 7.4	http://cml4.mu-sofia.bg:4000/home

Medical Academic Repository - Medical University - Varna Prof. Dr. Paraskev Stoyanov	Dspace 6.2	https://repository.mu-varna.bg/
BGOpenAIRE at IMI-BAS	DSpace 7.3	http://www.bg-openaire.eu
Bulgarian Portal for open science (BPOS)	Elasticsearch	https://bpos.bg/
Technical University of Sofia (E-Publications)	Technical University	https://e-university.tu-sofia.bg/e-publ/search.php
Digital Library SU 'St.Kliment Ohridski' *https://sudigital.org/en Manuscripts, Archives, Books, Periodicals, Maps, Pictures*	CDS Invenio v0.99.1	https://sudigital.org/lib/
NALIS Repository - The founders of the National Academic Library and Information System Foundation (NALIS Foundation) are the Central Library of BAS (CL-BAS), the Sofia University "St. Kliment Ohridski" (SU) and the American University in Bulgaria (AUBG).	DSpace 6.2	http://digilib.nalis.bg/xmlui/
Repository of the University of Chemical Technology and Metallurgy	DSpace 7.6	https://rlib.uctm.edu/home
CL-BAS Repository	DSpace 6.2	http://dspace.cl.bas.bg/xmlui/

Fig. 1. Open Access Bulgarian repositories in 2024.

The total number of OA journals in Bulgaria (DOAJ: Directory of Open Access Journals) is 96 – Figure 2.
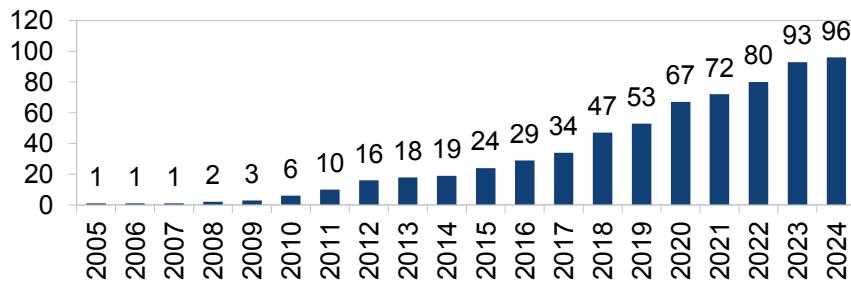
Fig. 2. The total number of OA journals in Bulgaria.

The existing Bulgarian National (open datasets) Portals are:

- **Bulgarian portal for open science** (https://bpos.bg/). The National repository is multidisciplinary and welcomes doctoral dissertations and peer reviewed manuscripts or published versions of scientific publications, created by researchers.
- **Open Data portal** (*https://data.egov.bg/*). Access to public open data sets in machine-readable format. It contains data sets from many organizations.
- **National Spatial Data Portal** – Infrastructure for Spatial Information, INSPIRE (https://inspire.egov.bg/) includes data sets, services and maps.
- **National Statistical Institute** (https://www.nsi.bg) like the Open portal provides free and unrestricted access, but for selected data sets in machine-readable data formats in CSV, JSON.

3. The Main Open Science Activity in Bulgaria – 2023-2024

National Priorities within the OS Initiative

- Provide the necessary trusted infrastructure for long-term storage, sharing and access to scientific publications and other research results from publicly funded research.
- Foster new skills and competences.
- Raise awareness on Open Science and ensure the active involvement of researchers, universities and research organizations.
- Create a National Open Science program to consolidate objectives and solutions.

National Policy and Infrastructure

Bulgaria has adopted a national Open Science Plan in January 2021. The general aim of the plan is to provide the research community and general public with access to scientific publications and research data in an open and non-discriminatory way at the earliest stage possible and ensure the possibility for their use and re-use. The plan provides for the adaptation of national policies, initiatives and research infrastructures to creation of conditions for inclusion and interoperability with the European Open Science Cloud.

Actions are being taken to create and operate a public platform, as well as institutional repositories directed at facilitating open access of scientific publications and data with a view to moving from paid subscriptions for scientific publications to open science. The established National program "Stimulation of publication activity in authoritative international scientific journals and open access to scientific information" aims to stimulate publication activity with a focus on publications in authoritative scientific journals, referenced or indexed in Web of Science or Scopus, as well as to promote the culture of sharing scientific information with open access and its adoption as a standard practice by the research community. Universities and scientific organizations with PhD programs accredited by the National Agency for Assessment and Accreditation and a profile in the Bulgarian Open Science Portal can apply for funding under the national program. A necessary condition for participation is that the applying organization must have and implement internal mechanisms to stimulate the open access publication activity of their research groups. Researchers are then provided with additional remuneration, rewards or other financial incentives for the authorship of such scientific publications.

In January 2023, Bulgaria joined the EOSC SB statement recognizing EOSC as the overarching transverse European Data Space for research. It thus shall be implemented as orthogonal and supplementary to the thematic European Data Spaces, which in turn should capitalize on expertise and solutions developed in the context of EOSC; recognize EOSC to be an accelerator of the digital transition, comprising FAIR digital objects as well as people with data science and management competencies; reaffirms that the EOSC governance underlines the importance of the sovereignty of the research sector as FAIR data provider in EOSC; invite to opening EOSC to the public sector at large and to commercial sector partners dedicated to open science, considering them as providers to and users of the EOSC ecosystem, which is developed to be a public good; call all research and innovation stakeholders to commit to making open science the new norm, as well as to develop and deploy the necessary science policies and supplementary financial investments to achieve this goal; commit to support progression towards open science and the advancement of EOSC implementation by facilitating monitoring, exchange of best practices, and promotion of mutual learning across countries and communities.

The Ministry of Education and Science (MES), together with the European Commission, the European Open Science Cloud Association and the EOSC Steering Board organized a hybrid thematic tripartite conference on March 28, 2023, in Sofia, Bulgaria. The aim of this event was to promote strategies and activities on the topic of open access

to scientific information and to give insight into the progress of implementing plans to achieve coherent national and European policies and practices. Open access to scientific information is essential for the dissemination of scientific knowledge in Europe. It has an impact on science, industry and society and is essential for a more effective use of pre-existing knowledge of scientific value. The panelists gave and international dimension to the event with views from the perspective of EOSC and the European Commission on the subject together with experience and quality insight on best practices from Sweden, the Czech Republic, Austria, Spain, Poland, Italy and France.

The Ministry of Education and Science has established a taskforce for the development of a national open science program. The National Science Centre of Poland (NCN) and MES have recently engaged in a form of bilateral cooperation in the field of open science, particularly concerning the European Open Science Cloud. The meeting served as an excellent opportunity for sharing EOSC-related practices, to discuss challenges, and opportunities for future collaboration between Poland and Bulgaria in advancing the open science agenda. This initiative serves to indicate a key role of bilateral/multilateral cooperation in driving forward agenda by sharing experiences in a similar geopolitical context, as both countries aim to contribute to the development and implementation of EOSC, fostering innovation and research excellence in Central Europe.

Open Data and Cloud Computing National Initiative in Bulgaria

The National Initiative Open Data and Cloud Computing (NI ODCC) is established by the members of National Centre for High-performance and Distributed Computing (NCHDC), object of the National Roadmap of RIs. The coordinating organization is the Institute of Information and Communication Technologies, Bulgarian Academy of Sciences (www.iict.bas.bg) (IICT - BAS) and the partners are Sofia University "St. Kliment Ohridski" (www.uni-sofia.bg), Technical University – Sofia (www.tu-sofia.bg) , Medical University – Sofia (www.mu-sofia.bg), University of Plovdiv "Paisii Hilendarski" (www.uni-plovdiv.bg), University of Library Studies and Information Technologies – Sofia (www.unibit.bg), Institute of Mathematics and Informatics - BAS (www.math.bas.bg), National Institute of Geophysics, Geodesy and Geography - BAS (www.niggg.bas.bg), Institute of Mechanics - BAS (www.imbm.bas.bg).

Recently, IICT conducted a survey to categorize the needs of Bulgarian researchers when it comes to open science and use of such resources. Additional questions were added to the survey to get a picture of the background, experience and high-performance computing and DATA usage of the aforementioned. Although the study was open to anyone interested in the field, the target group was HPC users of the NCHDC's resources. The survey was completed by 85 participants. Results show that: A significant part of users is interested in long-term data storage (5 years – 42.2 % of respondents, 10 years - 9.6 %) This can be tackled by dedicating separate resources and maintenance protocols, as well as by including training materials accordingly.

A significant interest in using HPC in the cloud (64.4 % of respondents). There is a need to respond by deploying relevant services and providing appropriate training. 43.7 % of the users from academia are familiar with FAIR principles for data storage, but

318

56.3 % seems to be unfamiliar. There is a definite need for dedicated training. The survey helped us gather more useful information regarding software used for data processing, type of systems used for data storage, access of metadata, etc. 25 % of participants have started using these technologies fairly recently or are yet to be trained. The majority of users have stated significant experience, so the emphasis of training can be put on developing more advanced users. Still, there is a tangible need for entry-level training as well.

The good news for researchers who need access to data storage systems is that in the last year the members of the initiative, through participation in projects funded by the Operational Program "Science and Education for Smart Growth", acquired a significant amount of new equipment with substantial data capacity. The largest new facility is the infrastructure complex in IICT which includes the supercomputer HEMUS with peak performance of more than 3 petaflops (ranked 360th in the November 2023 - top 500 list of most powerful supercomputers in the world), a data storage and processing system with a capacity of more than 6.72 petabytes, high performance server systems for application development and digitization facilities. For an efficient and convenient user experience, all layers of the computing ecosystem are well integrated, including network, computing resources, data, software, services and tools. This complex environment requires diverse expertise to achieve its potential, but also opens distinct possibilities, especially in high performance data analytics and artificial intelligence, as well as traditional HPC applications.

The Bulgarian Portal for Open Science

The Bulgarian Portal for Open Science (BPOS), managed by the National Centre for Information and Documentation (NACID) has continuously been developed in the past year. As of May 30, 2024, the number of documents available on the portal has increased to over 63 000. By access rights they are dispersed as follows: open access – 56 282, metadata only access – 6 432, restricted access – 1 025, embargoed access – 33. These data contributions are being provided by 113 content providers: research organizations (59), higher education institutions (43), publishers of scientific literature (7), the Bulgarian OpenAIRE repository, as well as various independent researchers. Even several funding agency reports, provided by the National Research Fund, can be found., Access is provided even to dissertations defended since 2018 via an automated interface, established between BPOS and the Register of Academic Staff in the Republic of Bulgaria, also managed by NACID with their number reaching over 6500. Among the various resources available, the number of datasets and data papers is still very limited.
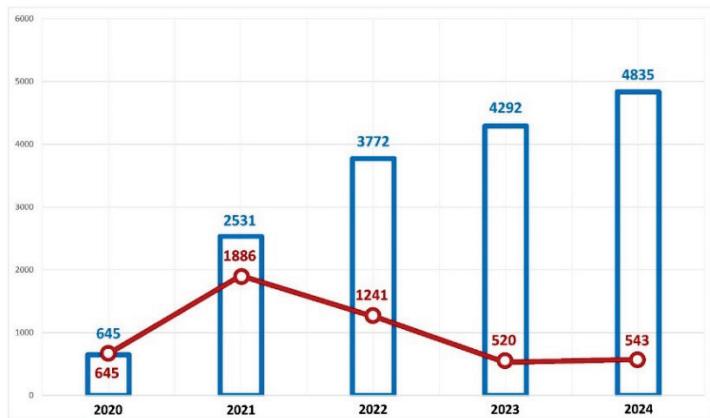
Fig. 3. Registered individual users by year (2020 – 2024). The number of new users for the year (the line) and total number of users for the year (bar chart).

The total number of registered individual users has reached 4 835 (as of May 30, 2024) (Figure 3). By category, they are distributed as follows: 4 347 scientists, 291 moderators and 197 organization administrators. This demonstrates sustained interest of the scientific community in BPOS. It is important to mention that BPOS is, still, the single trustworthy repository in Bulgaria, certified by CoreTrustSeal. A representative of BPOS is a member of the Assembly of Reviewers of CoreTrustSeal.

The brand-new Promotion of Research and Innovation Act (published in the Official Gazette, Special Issue 39 of 1 May 2024, pp. 10-30) introduces some legal definitions, related to open science. Open science is defined as transfer of knowledge by providing access to research data and wide dissemination of research results that are financed via public funds. It is established that publicly funded research is based on the principle of open science, without contradicting the principles of intellectual property protection. It is important to note that there is a separate chapter devoted to open science promotion and the BPOS. It is stated that the Ministry of Education and Science, through NACID, establishes and maintains the Bulgarian Portal for Open Science. The portal is considered a national repository of scientific information, resulting from research funded in whole or in part by public funds; moreover, it is also the single point of access to such information. The procedure for publishing and storing data and publications in BPOS shall be determined by regulation of the Minister of Education and Science.

Under the "Research, Innovation and Digitalization for Smart Transformation" Program, an upgrade of several features of the portal in 2024 is planned. This aims to build a platform of cooperation between research institutions and businesses, by improving some systems maintained by NACID, i.e.: the Register of Academic Staff, the Register of Scientific Activity, and the Bulgarian Portal for Open Science.

Training and Support

The Bulgarian Portal for Open Science team regularly participates in trainings, seminars, scientific conferences and other appropriate forums. Among the ones held in the past year, we should pay special attention to the "Open Access Publishing of Scientific Information" online training (June 7, 2023). This event was jointly organized by NACID and the Science Directorate of the Ministry of Education and Science and involved representatives of universities and research organizations in Bulgaria.

The main functionalities of the Bulgarian Portal for Open Science and the European research infrastructures for open science were presented, focusing on metadata, persistent identifiers and licenses, highlighting the benefits of open science. During the discussion that followed, the participants had the opportunity to discuss ideas for the improvement and further development of BPOS.

More than 200 open science publications were approved and published in the hybrid Elsevier journals under the bilateral transformative agreement between Bulgaria and Elsevier.

Outreach and Dissemination

- In the frame of OpenAIRE, we work in support of researchers and research-supporting organizations to facilitate a smooth shift into open science research practices, easing the transition with services, training, and appropriate consultations:
- Support via creation, adoption, and implementation of open science policies, aligned with European research policies and global standards.
- Facilitate interoperability between services, research outputs and EOSC.
- Offer consultation and guidance to publish in open publications, data, and software.
- Monitor outcomes, policy uptakes and the scientific, economic, and societal impact OpenAIRE activities have on the research community in Bulgaria.
- Thanks to OpenAIRE Services, the research community has access to:
- ARGOS a tool for creating data management plans.
- ZENODO a catch-all repository able to archive not only publications, but also datasets, codes, and software.
- AMNESIA facilitates the anonymization of data, helping users share datasets protecting personal data.
- MONITOR a service showing indicators for open science and publication costs.

Task Forces

- Participation in the European Open Science Cloud – Steering Board;
- Participation in the EOSC Association - https://eosc.eu/members/sofia-university-st-kliment-ohridski.
- Participation in the Research Data Alliance.
- Participation in the OpenAIRE AMKE (IMI-BAS).

- Participation in Skills4EOSC (IICT-BAS).
- Participation in the EOSC TF on Data Interoperability (IICT-BAS).
- Participation in the Open Science Fair, Madrid (IMI-BAS).
- Participation in RICH Europe - 1st Symposium of Research Infrastructures, Madrid (IMI-BAS).

4. The 15th National Information Day: Open Science, Open Data, Open Access, Bulgarian Open Science Cloud. Topics for discussion

- Bulgarian Roadmap for Open Science.
- Stimulation of publication activity in respectable international scientific journals and open access to scientific information.
- BPOS and the national repository.
- Institutional repositories and other e-infrastructures.
- FAIR data requirements and tools.
- Transformative agreements with publishers.
- Open data.
- Open sources.
- Open access indicators.
- Best practices.
- Contribution to solving issues with harmonization of national legislation and practices.

Acknowledgments.

References

O'Neill, G., Lebhardt, F., & Grapengiesser, I. M. (2024). *Brochure of Use Cases on EOSC and Open Science (Version V1)*. Zenodo. https://doi.org/10.5281/zenodo.10822698

OpenAIRE. (2022). *OpenAIRE Science. Set Free. 3-year Strategy 2023-2025*. https://www.openaire.eu/openaire-strategy-2023-25