Point of View

The Uniform Electronic Knowledge Space Revisited

A. B. Antopol'skii* and D. V. Efremenko**,#

Institute of Scientific Information on Social Sciences, Russian Academy of Sciences, Moscow, Russia
*e-mail: ale5695@yandex.ru
**e-mail: efdv2015@mail.ru
Received August 27, 2017

Abstract—The set of issues associated with the formation of a single Russian electronic knowledge space is considered. This task, first formulated in the Russian Presidential Decree On Approval of the Foundations of a State Cultural Policy, has become particularly urgent in the context of the so-called Big Challenges, preconditioned by the development of scientific knowledge and technologies. To solve this task, it is necessary to develop a tool kit based on the Semantic Web technology, using scientific electronic libraries and other data arrays, as well as classifications, thesauruses, ontologies, and systems of metadata and other tools for subject area representation. The creation of a single Russian electronic knowledge space may be referred to nationwide projects, and the RAS scientific organizations and the country's leading universities should play the leading role in it.

Keywords: single Russian electronic knowledge space, scientific information, infosphere, open access, state scientific–technical policy, traditional and digital communications.

DOI: 10.1134/S1019331618010070

The development of science and technology has brought modern civilization close to a technological shift that entails complex and irreversible changes in the interrelations of society, nature, the world of technical artifacts, and virtual reality. The problem is not only in the potential revolutionizing (or destabilizing) influence on humans, society, and the environment of technologies such as the Internet of Things, Big Data processing tools, 3-D printing, nanobiosensorics, brain implants, artificial photosynthesis, smart cities and electricity grids, geoengineering, etc. We are primarily speaking about the necessity to understand adequately systemic interactions that undergo qualitative changes because of stormy technological progress. As the Long-Term Strategy for the Scientific and Technological Development of the Russian Federation, approved by Russian President V.V. Putin on December 1, 2016, envisages, "It is necessary to make the country ready for big challenges, which have not yet been manifested and have not yet gained broad public recognition, and to anticipate opportune assessment of risks predetermined by scientific and technological development" [1].

Science-generated big challenges precondition the vector of development of not only society in general but also science itself as a social subsystem and institu-

tional mechanism of expert support for political decision making. To this end, the efficacy and adequacy for such challenges of the research information support infrastructure that exists in Russia are questioned again. Respectively, the necessity arises to assess the condition of the *scientific infosphere*, by which we understand the totality of information resources, services, and institutions that ensure scientific communication.

Today, as well as over the past quarter of a century, there has been no reason to speak about a single vector of development of information processes in Russian science. We may speak about structural and technological metamorphoses or significant trends [2], but a clearly articulated strategy is absent. At the same time, the changes themselves are truly fundamental.

Science as a social subsystem has fully experienced diversified consequences as humanity has broken away to the "Internet Galaxy." The most visible indicators of this are mass digitalization of scientific communication, the avalanchelike growth of communication channels, the development of social networks and collaborations, and the appearance of resources for collective creation. The scale of these changes indicates the advent of a digital revolution, accompanied by revolutionary technical changes during information production and transfer, as well as by fundamental socioeconomic consequences, which transform the modus operandi of contemporary society [3].

Aleksandr Borisovich Antopol'skii, Dr. Sci (Eng.), is Chief Researcher of the RAS Institute of Scientific Information on Social Sciences (RAS INION). Dmitry Valer'evich Efremenko, Dr. Sci. (Polit.), is a Deputy Director of RAS INION.

However, Russian science and its infosphere have also experienced the effect of political and socioeconomic transformations, through which our society and state passed on the cusp of the 1980s and 1990s. First, political and ideological censorship became history, which, no doubt, favored the intensive development of social and humanistic knowledge in our country, although, at the same time, it seriously weakened the positions of centers of scientific information, which provided researchers with exclusive access to "ideologically dubious" information flows. Second, the state has lost monopolistic positions in the production and distribution of scientific knowledge and information. Many players not associated directly with the state and with opposite interests have appeared in this field. In particular, commercial entities—integrators; commercial electronic libraries; encyclopedic, vocabulary, and inquiry services; etc., for whom the provision of users with access to scientific information is inseparably associated with profit making—are beginning to play an increasingly large role.

Simultaneously, public structures, more precisely, representatives of the network civil society, many of whom advocate open access to and the lifting of most barriers in the way of scientific information flows, enter this process [4, 5]. Thus, there are grounds to characterize the information space of Russian science as a complex dynamic system, distinguished by nonlinear interactions and self-organization processes.

The latter does not at all mean that it is impossible to control the information space of Russian science. The self-organization processes can and should be given a certain vector, concentrating resources in attractors that ensure further structuring of the scientific infosphere. Without downplaying the importance of the public and commercial sectors of the infosphere of Russian science, we believe that only the state can solve successfully such tasks in the present-day conditions.

The general problem here is associated not so much with the scientific infosphere as such as with the state scientific and technological policy. Obviously, the motive to reform the scientific sphere after the disintegration of the Soviet Union was the desire of government structures and specific individuals responsible for the scientific and technological policy to withdraw from the Soviet model of science organization, reducing costs but attaining a general increase in the efficacy of scientific activity. Apparently, the reform of the Russian Academy of Sciences was also aimed at this. We say "apparently" because in 2013 and later, there was no clear explication of its goals that could have won the confidence of the scientific community.

Four years later, we can state the success of the RAS reform from the fiscal point of view, as scientific institutions were "built into" the hierarchical vertical of scientific institutions, subordinating academic institutes directly to the Federal Agency for Scientific Organizations (FASO). However, the structural and organizational model of scientific activity remains Soviet. Neither the redistribution of resources in favor of leading universities (which, overall, surely encouraged the growth of their publishing activity) nor the formally declared transfer of state support to the level of individual laboratories ensured a qualitative transformation of science organization in Russia. A fundamentally new Russian model does not exist; and there have been no explicit and consistent attempts to copy the American or German institutional prototype of scientific activity. Even the very successful Chinese experience of science governance remains practically unclaimed: science governance in China was largely based on the Soviet pattern, and during the period of reforms by Deng Xiaoping and his successors, it was adapted to new realities, ensuring technological modernization and innovative activity in the Chinese economy [6−9].

A document very important for Russian science, the Long-Term Strategy for the Scientific and Technological Development of the Russian Federation, was adopted in late 2016 [1]. It fixes the principle of the "integrity and unity of Russia's scientific and technological development," considering such development as a common task of scientific and educational organizations, industrial enterprises, and other institutions that conduct scientific research and scientific−technical and innovative activities. This should quite logically entail the principle of a single information space for Russian science. Strangely enough, however, the strategy does not focus on this, even when enumerating problems that hinder Russian science from reaching the necessary level of competitiveness. At the same time, the strategy stresses that one of significant factors of development of science in the Russian Federation is "a sharp increase in the amount of scientific and technological information, the appearance of principally new methods of work with it, and the change in organizational forms and hardware and software tools for R&D." Consequently, one task of Russia's scientific and technological development is to form "an effective system of communication in science, technology, and innovation," to ensure increased openness of the economy and society to innovation, and to promote conditions for the development of science-intensive businesses. In addition, it is true that nothing has been said about the role of scientific and technical information, and amid information components of the infrastructure of Russian science, it names only "access of research groups to national and international information resources" [1].

Unfortunately, the strategy does not bring any clarity to the issue of the future institutional and organizational model of Russian science and, specifically, the role of the RAS. Most probably, this means that the near future will see no principal changes in science at higher education institutions and in the system of academic institutes. Probably, the initiative to consolidate

institutes will be carried out further; they will become fewer in totality, but the remaining scientific organizations will grow numerically without radical changes in the general principles of functioning. At the same time, we may expect steps aimed at greater coordination and formation of collaborations of science at higher education institutions and academic institutes.

The framework of this scenario allows for large players of scientific-information activity, preserving the opportunity for their optimal labor division and work coordination. It is very important to orient it at forming a single integral system of information support for scientific labor, at least, within a network of academic scientific institutions and ideally on a national scale. Labor division between participants in scientific-information activity should fully reflect the specifics of various types of scientific knowledge, but the priority is to reach systemic integrity.

Naturally, one cannot enter the same river twice. Obviously, it is pointless to revive the former State System of Scientific and Technical Information (SSSTI) due to the profound change in the socioeconomic and technological conditions, as well as in the external environmental factors. The new system should, first, ensure the preservation of knowledge accumulated at institutions of social memory, higher educational establishments, and scientific organizations; second, provide researchers, students, and postgraduate students with access to scientific knowledge and information; third, use state-of-the-art network technologies to ensure multifunctional information support for research; fourth, guarantee the on-the-fly reflection of new knowledge; fifth, develop tools for the analytical study of the production processes of scientific knowledge, scientific information, and communications of the scientific community.

The regulatory and conceptual basis for the development of the scientific infosphere could, probably, be the concept of a *single Russian electronic knowledge space* (SREKS), which first appeared in Presidential Decree no. 808 of December 24, 2014, On Approval of the Fundamentals of the State Cultural Policy in the following formulation: "The formation of a single Russian electronic knowledge space based on digitized book, archival, and museum funds, collected into a National Electronic Library and national electronic archives on various subject areas and spheres of creative endeavor" [10]. Later, the SREKS concept resurfaced in the Federal Law On Librarianship, adopted in 2016, which states [11],

> The purpose of creating a library is to preserve historical, scientific, and cultural treasures of the peoples of the Russian Federation; to ensure conditions for improving the intellectual potential of the Russian Federation; to popularize Russian science and culture; and to form the basis for creating a single Russian electronic knowledge space.

In May 2017, a draft of the updated Provisions on the National Electronic Library (NEL) [12] attempted to give an extended interpretation of the SREKS concept. This document considers NEL as a basic structure that forms the totality of mutually integrated information systems and other resources that function on the basis of single technologies and principles that ensure semantic cohesion of the context, as well as tools for user-requested information search and retrieval.

The draft NEL Provisions also state the goal of creating the knowledge space: "the formation of a single, integral, and authoritative totality of accumulated human knowledge, the improvement of the intellectual potential of the Russian Federation, the popularization of Russian culture and science at home and abroad." It is assumed that the priority in forming the knowledge space will be given to documents in Russian and other languages of Russia's ethnic groups. In addition, the main principles of knowledge space formation are enumerated:

• unlimited user access to information contained in the knowledge space;

• gratuitousness of user access to information;

• authenticity and authoritativeness of information;

• semantic coherence of information.

The main components of the electronic knowledge space, except for the NEL itself, include the following:

• a nationwide scientific and educational interactive encyclopedic portal;

• information resources containing electronic copies of the documents of the Archival Fund of the Russian Federation, access to which is unlimited in line with Russian legislation;

• the federal information system the State Catalog of the Museum Fund of the Russian Federation;

• information resources containing electronic copies of objects and collections of Russian museums;

• information resources containing electronic copies of audiovisual documents owned by organizations that are authorized to keep them permanently in line with Russian legislation;

• a technology for an information classification, search, and retrieval system.

As for the sequence of steps to form an electronic knowledge space, after the NEL, it is assumed that a system of information classification, search, and retrieval will be created, the NEL-based information systems and resources, which function by the same principles and rules, will be mutually integrated, and the electronic copies of documents contained in the information classification, search, and retrieval system, as well as in information and telecommunication networks, will be connected semantically. Finally, it is assumed that an Internet Knowledge portal will be created to ensure information retrieval on requests

from the users of information systems and resources included in the knowledge space.

In this article, we do not set the task of analyzing the SREKS concept critically, although many of the above formulations raise questions and require additional specifications. Let us opt for legal and administrative logic: being mentioned in the federal law, the concept of a single Russian electronic knowledge space thus acquires a legitimacy; therefore, we will focus on practically and technologically acceptable interpretations.

The most crucial SREKS aspect, in our opinion, is functional. The very composition of functions to be implemented within this space should determine the mechanism of its creation, as well as the structural and sustaining components.

We can propose the following list of tasks aimed at the development of Russia's modern scientific infosphere and the creation of SREKS:

• the development of tools to form the electronic scientific knowledge space using the semantic network methodology, the existing scientific electronic libraries, and other electronic scientific information arrays, as well as classifications, thesauruses, ontologies, metadata systems, and other means of subject area presentation;

• the formation of a state-of-the-art system of operational reflection in SREKS of new knowledge obtained by Russia's scientific organizations and higher educational institutions and represented electronically, including the automatic formation of electronic editions, repositories, encyclopedias, directories, and dictionaries;

• the development and implementation of technologies to evaluate the efficacy of research and scientific organizations using infometric methods and semantic technologies;

• the development and implementation of technologies of automatic identification of promising points of growth and new trends in Russian science;

• the development of principles and technologies of open science, open access, and open data, as well as the formation of the corresponding infrastructure;

• the coordination and optimization of access to foreign information resources (this is the only one of the above tasks that is being implemented at present [13]);

• ensuring the preservation of the scientific knowledge stored in electronic resources of scientific information;

• the coordination of activity in forming documental, factographic, encyclopedic, and vocabulary scientific electronic resources, electronic scientific publications, and their support;

• the formation of the SREKS control system based on public—private partnership and the development of organizational, economic, and legal support for the SREKS creation;

• the organization of multiaspect and multifunctional information support for scientific research based on state-of-the-art network technologies;

• the improvement of the efficiency of social costs of the formation, support, and use of electronic scientific resources;

• the creation of a system to analyze and foresee the scientific infosphere.

Amid the SREKS components mentioned in the draft NEL provisions quoted above, the *information classification, search, and retrieval system* is of special interest. Considering the intensive development of cognitive information technologies in recent years, it is safe to assume that NEL should be supported by a general scientific Russian-language ontology, which is formed and supported using a semantic network and associated open data. At the same time, the concentration of large scientific-information arrays in NEL, other national and industrial electronic libraries, archives, and databases is an objective condition for providing the above ontology with an interactive function. It should ensure interaction with information-search languages that are used to form the existing electronic information resources. In addition, as it is created, it is necessary to consider different classifications for heterogeneous resources: for library resources, it is primarily the Universal Decimal Classification (UDC) and Library-Bibliographical Classification (LBC); for patent resources, the International Patent Classification (IPC); for electronic libraries of dissertations and extended abstracts, the Higher Attestation Commission (HAC); for many resources of scientific and technical information, the State Rubricator of Scientific and Technical Information (SRSTI); and so on.

The logical and linguistic basis for SREKS should be a single Russian-language ontology of scientific knowledge, which includes the lexicon and paradigmatics of classifications, thesauruses, metadata systems, and other semantic means, practically used to form national electronic resources: librarian, archival, museum, and other arrays of scientific and educational information. In addition, the source material can, in our opinion, be a system of conformity of scientific classifications, created as a result of previous work [14—16] and furnished with glossaries of definitions taken from various encyclopedias. It is important to maintain continuity with many practically used classifications.

As SREKS is created, it is necessary to conduct an integrated analysis of scientific resources and to study trends in their development, processing technologies in Russia and abroad, knowledge extracts, formalizations, and presentations primarily based on a semantic network. As a pilot project, it is possible to consider the monitoring of academic information resources in

the humanities and social sciences, organized at RAS INION, which shows that over 16000 information arrays of various types have been created and supported in the institutions of the FASO/RAS system [17]. The results of this monitoring will help determine the specific parameters of general systemic processes within SREKS: digitization programs; digital information protection; and navigation, aggregation, and integration of information resources.

At the first stage of project development, which should consider the experience of the existing knowledge formalization systems, e.g., Socionet [18], it is possible to propose and implement experimentally a SREKS model that

• optimizes the SREKS structure, turning on the distribution of centralized, industrial, regional, thematic, and local resources;

• implements mechanisms of navigation and semantic search in the distributed information space;

• automatically integrates into the SREKS model electronic catalogs of institutions of social memory (scientific libraries, archives, and museums), science management databases, citation indices, and other secondary information resources;

• coordinates the activity of forming and developing scientific electronic libraries, including the development of Russia's Scientific Heritage electronic library, and digitizes the documentary collections of academic scientific institutions in cooperation with NEL;

• forms a single Russian ontology of scientific knowledge as a system of associated open data based on the existing thesauruses, classification schemes, dictionaries, encyclopedias, and metadata systems;

• multifunctionally uses SREKS for bibliometric analysis of scientific activity, including the identification of promising trends and points of growth in Russian science and its participation in the international division of labor;

• reveals duplication, plagiarism, low-quality materials, "predatory" journals, and other negative phenomena arising in scientific communications;

• examines investment pitches for research by topicality and novelty parameters;

• selects and stores indefinitely valuable electronic scientific resources not included in the funds of contemporary memory institutions (primarily scientific resources of digital origin).

One of the important systemic tasks of development of the infosphere is rational distribution of scientific information resources between centralized and decentralized services. Now they are not just decentralized but, in fact, scattered to full dissipation. Developers lack due coordination; they often do not even know that colleagues have duplicate resources. In addition, not just mutual information sharing is principal but the motivation of developers for collaboration and collective creation of information resources and services. Corresponding organizational—economic mechanisms are necessary to this end, e.g., endowments.

An important feature of the state-of-the-art infosphere is a transitional state between the traditional and digital forms of communication, when many resources and services are duplicated in both forms. In these conditions, it is necessary to think over the composition of centralized services that could ensure coordination of information activity of academic institutions and optimize its costs. The centralized services may include the following:

• an automated system of editorial and publishing preparation of FASO/RAS scientific journals;

• a centralized repository of electronic versions of scientific journals;

• a system to coordinate the digitization of document collections of scientific institutions;

• a cumulative catalog of scientific documents, including the service of corporate cataloguing to aggregate the existing electronic catalogs of scientific institutions, libraries, and archives;

• a navigator of electronic scientific resources, including analysis of their use based on infometric and expert methods;

• a databank of linguistic tools used to form SREKS, including ontologies, classifications, dictionaries, authority files, metadata systems, etc.

The organizational issue concerning SREKS creation requires special discussion: a nationwide interdepartmental project. Since we are speaking specifically about the scientific knowledge space, the leaders of this process, in our opinion, should be FASO/RAS institutions and universities. The driving forces could be collaborations of leading universities and academic research institutions. Such an example is Moscow State University's National Center of Competences in the Digital Economy, the RAS Federal Research Center of Informatics and Management having contributed greatly to its creation in 2017.

The RAS institutes of scientific information, the All-Union Institute of Scientific and Technical Information (VINITI) and the Institute of Scientific Information for Social Sciences (INION), should play a significant role in implementing the project. We believe that, in the medium term, academic scientific information services should transfer the center of gravity of their work from the primary level of analytical and synthetic processing of scientific information (its compression to an abstract format) to the level of multifunctional analysis of scientific-information flows, scientific knowledge production, and scientific communication processes. Of special importance in the

context of great challenges will be deep integration of scientific-information activity and research in science studies.

Academic scientific-information services have obviously irreparably lost the role of the main link between scientists and scientific information flows, but they can be mediators that link institutional science with the network environment. The principally important point is that integration of various forms of scientific and educational communications, including access to publications, reviewing and discussing results, editorial-and-publishing processes, archival and librarian activities, scientometric services, etc., happens today in the Web 2.0 format. To this end, it is important for academic scientific-information centers to participate actively in network communications and to take on the functions of monitoring the scientific blogosphere, as well as domestic and foreign network services. The previous achievements and vast experience of the institutes of scientific information will guarantee solving this large-scale task at a high professional level.

The SREKS creation program should encompass libraries, archives, and museums under the Russian Ministry of Culture and the Russian Ministry of Education and Science, as well as large asset holders of other departments and nongovernment institutions. For example, participation is obligatory for the Yeltsin Presidential Library, the Federal Institute of Industrial Property, the Scientific Electronic Library, the All-Russia Research Institute of Hydrometeorological Information–World Data Center, and other major participants in the scientific infosphere.

Finally, in discussing the problems of the single Russian electronic knowledge space, we cannot but touch on the issue of *open data* and *open access*. We believe that the broader public and scientific community should be informed of research achievements and results in an open data format (an exception here is, no doubt, classified materials). The principle of open access, in turn, is quite crucial for further development of the infosphere of social sciences. Under the Russian conditions, it is necessary to make an unambiguous choice in its favor. This is our stand. It is predetermined not by ideological but primarily pragmatic considerations associated with the interests of society, Russian science, and social memory institutions.

The implementation of the open access principle in our country is hardly a simple task. The general approach is that new knowledge, obtained during the activity of scientific institutions and higher education institutions financed from the budget, should reach the public domain (considering the above restrictions). It is necessary to initiate an extensive dialogue between academic institutes, universities, libraries, museums, archives, scientific publishing houses, and representatives of civil society and commercial entities

to promote in a coordinated way the principles of open access and open science. It will stimulate the elaboration of a clear-cut position of the public authorities regarding information support for scientific and technical activity. Ideally, the functioning of the scientific infosphere and the work of creating a single Russian electronic knowledge space should be based on a specific social contract that would develop the state-owned, public, and commercial sectors, each of which is necessary for successful information support of scientific labor.

REFERENCES

1. The Long-Term Strategy for the Scientific and Technological Development of the Russian Federation. http://sntr-rf.ru/materials/strategiya-nauchno-tekhnologicheskogo-razvitiya-rossiyskoy-federatsii-na-dolgosrochnyy-period/. Cited August 25, 2017.

2. A. B. Antopol'skii and D. V. Efremenko, *The Infosphere of Russia's Social Sciences* (Direkt-Media, Moscow, 2017) [in Russian].

3. C. Freeman and F. Louçã, *As Time Goes By: From the Industrial Revolutions to the Information Revolution* (Oxford Univ. Press, Oxford, 2001).

4. "The Hague Declaration on Knowledge Discovery in the Digital Age," Bibliotekovedenie, No. 4, 83–87 (2015).

5. P. N. Materu, Open Source Courseware: A Baseline Study (2004). http://siteresources.worldbank.org/INTAFRREGTOPTEIA/Resources/open_source_courseware.pdf. Cited August 25, 2017.

6. K. Jonkers, Mobility, *Migration and the Chinese Scientific Research System* (Routledge, London, 2010).

7. W. A. Wells, "The returning tide: How China, the world's most populous country, is building a competitive research base," J. Cell Biol. **176** (4), 376–401 (2007).

8. Y. Xie, C. Zhang, and Q. Lai, "China's rise as a major contributor to science and technology," Proc. Nat. Acad. Sci. U. S. A. **111** (26), 9437–9442 (2014).

9. A. V. Vinogradov, E. A. Salitskaya, and A. I. Salitskii, "Science and technology in China: Modernization accomplished," Herald Russ. Acad. Sci. **86** (1), 48–56 (2016).

10. http://kremlin.ru/acts/bank/39208. Cited August 25, 2017.

11. http://fzrf.su/zakon/o-bibliotechnom-dele-78-fz/st-18.1.php. Cited August 25, 2017.

12. http://www.unkniga.ru/images/docs/2017/polozhene-neb-2-variant.pdf. Cited August 25, 2017.

13. Organization of a centralized (national) subscription for scientific and educational entities to information resources of global informational and analytical databases, electronic books, and scientific journals (2017). http:минобрнауки.рф/проекты/базы-научного-цитирования. Cited August 25, 2017.

14. N. E. Kalenov and V. N. Beloozerov, "Formation of terminological dictionaries for the lexicon of classifica-

tion systems," Nauchn.-Tekhn. Inform., Ser. 1. Organ. Metod. Inf. Rab. **42** (3), 60−70 (2015).

15. Yu. M. Arskii, I. Yu. Nikol'skaya, and S. M. Gonnova, "Formation of a system of thematic classification to develop information exchange in the scientific and technological sphere," Nauchn.-Tekhn. Inform., Ser. 1. Organ. Metod. Inf. Rab. **42** (9), 23−26 (2015).

16. A. B. Antopol'skii, V. N. Beloozerov, T. S. Markarova, and E. Yu. Dmitrieva, "Establishing correlations between the rubrics of the State Rubricator of Scientific and Technical Information (GRNTI) and rubrics of other systems of classification of scientific and technical information," Nauchn.-Tekhn. Inform., Ser. 1. Organ. Metod. Inf. Rab. **42** (3), 3−19 (2015).

17. A. B. Antopol'skii, "On the organization of monitoring of information resources of academic entities associated with the humanities and social sciences," in *Digital Humanitaristics: Resources, Methods, Investigations. Proc. of the International Scientific Conference in Two Parts* (Perm. Gos. Nats. Issled. Univ., Perm, 2017), Part 1 [in Russian].

18. On the Socionet draft scientific information system (2016). https://socionet.ru/idea.htm. Cited August 25, 2017.

Translated by B. Alekseev