ISSN 1063-7710, Acoustical Physics, 2019, Vol. 65, No. 6, pp. 778–786. © Pleiades Publishing, Ltd., 2019.
Russian Text © The Author(s), 2017, published in Akusticheskii Zhurnal, 2017, Vol. 63, No. 4, pp. 449–458.

INFORMATION

Acoustics Open Information System

V. G. Shamaev[a], * and A. B. Gorshkov[b]

[a] *Moscow State University, Moscow, 119992 Russia*
[b] *Moscow State University, Sternberg Astronomical Institute, Moscow, 119234 Russia*
**e-mail: shamaev08@gmail.com*

Received February 14, 2017; revised February 14, 2017; accepted March 15, 2017

Abstract—The problems associated with the absence in the Russian Federation of a single center for the collection and provision of information on Russian-language scientific publications are discussed. A technological scheme is proposed that is useful in creating such a center. The cluster of this center is described via example of acoustic topics, already implemented at the Department of Acoustics, Faculty of Physics, Moscow State University. The corresponding Acoustics portal (http://akdata.ru) is publicly available on the Internet. Entrance points of the portal are the following: the full-text archive of the Acoustics Journal, "Signal Information" on acoustics, and "Information Retrieval System 'Acoustics'. Russian-language sources". Work with the system is described. Scientometric data from analyzing the information content of the system are presented for journals, headings, and authors. Bradford's Law of Scattering of scientific publications is verified using acoustics topics. A list of authors demonstrating publication activity in the field of acoustics is given.

Keywords: Russian-language scientific literature, electronic documents, databases, information retrieval systems, Internet resources, full-text archives of scientific journals on the Internet, the Acoustics portal

DOI: 10.1134/S1063771019660047

The amount of information describing the current state of science is so large that, without modern information technologies, its effective use is not possible [1]. Therefore, along with traditional theoretical and experimental methods for researching the world around us, modern databases, which are an integral part of information retrieval systems, began to appear as a new tool of scientific knowledge. Any successful attempts to formulate and solve scientific problems using modern databases (DB) allow us to talk about the emergence of a new scientific direction, which consists in the development of methods for meaningful analysis of large amounts of information and, from them, obtaining new knowledge about the world around us and its laws.

Continuous replenishment of the database allows us to counteract the loss of information. If earlier, to replenish knowledge about the state of one's field of science, it was necessary to regularly read relevant specialized journals, as well as abstract journals, today it is enough to use a continuously updated information retrieval system (IRS). For chemists, the SciFinder IRS, which grew out of Chemical Abstracts, serves as such a system; for mathematicians, MathSciNet; for engineers, INSPEC; for astronomers, ADS (NASA); for physicists, Web of Science or Scopus (until 1991, RJ Physics of VINITI was successfully used [2, 3]), etc. However, in all of these western IRSs, there are practically no works by Russian authors, with the exception of those that are reflected in translated scientific journals or translated into English and sent by the authors to foreign publishing houses. The absence of Russian-language journals, and, consequently, Russian-language works in foreign databases, leads to their loss for a significant part of the scientific community. It is as though they do not exist.

This situation, to some extent, suited our scientific community in the 1990s. More correctly, some had more important things to deal with, and some were lulled by the euphoria of the opportunity to go abroad and the fact that "foreign countries will help us!" in the form of grants, etc. However, since the 2000s, this situation began to weigh on many of us and required change. Available sites of scientific journals appeared, among which we should note the portal of the Radiotekhnika Publishing House [4], the sites of journals of the Ioffe Physicotechnical Institute of the RAS [5], and the Siberian Branch of the RAS [6]. The Scientific Electronic Library project began implementation. Mathematicians, without waiting for an official decision, implemented an elegant initiative project called The All-Russian Mathematical Portal [7]. Therefore, it cannot be said that the Russian-speaking scientific community did nothing. In the same direction, our Acoustics portal was created, described below.

A significant number of scientific journals are devoted to one subject (for example, Acoustics Journal, Quantum Electronics, or Astronomical Journal). There are also polythematic journals publishing studies in several areas (Uspekhi Fizicheskikh Nauk, Jour-

nal of Experimental and Theoretical Physics). Because of this, only selected articles can be of interest to a particular reader. For many, it would be surprising that acoustics topics are scattered in more than 500 Russian-language magazines. Acoustics (understood in the broad sense of the word) is dealt with by no fewer scientific workers (including natural, humanitarian, and technical sciences) than the number of physicists. Therefore, the objective of creating the Acoustics portal is to provide open access and broad opportunities in the search for Russian-language information on acoustics to Russian and foreign researchers and engineers.

The Acoustics portal collects information from all sources available to us, including periodicals, books, proceedings of conferences and seminars, patents, regulatory documents, etc. All this information is placed in the form of a full-text archive of the Acoustics Journal, the only Russian academic journal entirely devoted to acoustics, and in the form of Signal Information. In the archive of Acoustics Journal, one can view the full text of the articles in the issue, as well as search by authors and headings. In Signal Information, every two months information is given (articles in journals, books, etc.), published since the previous issue was printed. Viewing is possible via journal issues, authors, and headings, as well as by the original layout of the issue published immediately on the site. Signal Information gives a look at the current state of acoustic research in the Russian-speaking world. But the main purpose of the portal is to support the Acoustics IRS. In the IRS, information is obtained in the mode of setting a search query, then viewing the issued information.

Such an information portal with a unique database, which contains information on acoustic publications of a current and retrospective nature, is one of a kind. It fully satisfies the purpose of our work: to provide on the Internet the information products on Russian-language works that fully reflect the subject area, to realize the search ability and, according to the opportunities provided, meet the requirements of specialists.

In each of the portal products connected by one design, we settled on an interface with a "minimalist style", which, in our opinion and based on experience with various information resources, is essential for users. All that is needed is in front of the user, not on different pages of the interface, and a minimum of "clicks" is required to start the search process in IRS or the issuance of specific information in the Acoustic Journal and Signal Information sections.

On the home page, with which the portal opens (http://akdata.ru), the user finds the Acoustics Information Retrieval System (Fig. 1). As one can see, eight fields are enough to search. It was possible to make seven of them, but, unfortunately, in some publications, for example, Reports of the Academy of Sciences, the volume numbers change throughout the year. In the specified edition, six times. The user can enter the name of the source, while, already with three characters entered, a suggestion appears, which is updated with each next character. If the user sees the desired edition in the list, they just click on it and the field will be filled immediately, if not, they should continue typing.

Then, one can enter the year (or range of years) by which the search will take place, and volume and issue number, if the information is searched in a periodical. The next field is "authors", then "keywords" and "rubric". One can fill in all fields of the IRS or any number of them. The system will indicate the total number of documents found and will provide the required information, but in an amount of not more than 100 documents. It is further proposed to refine the search parameters. The easiest way to do this is to define the years or add the names of the authors in the appropriate fields. One can also specify a heading, then, if being more precise is needed, a subheading.

By the way, if one specifies only the author, then, for the vast majority of authors, this is enough. Authors with more than 100 articles in the system make up 0.07%, or quantitatively 18 people at the beginning of 2017. Note that the author's last name is entered in the system's database as it appears in the article, therefore, for example, there are authors "Khmelev, V.N." and "Khmelyov, V.N." or "Semenov, A.G." and "Semyonov, A.G."—perhaps these are the same scientists, but this cannot be said with absolute certainty, and it is necessary to clarify each article with the authors. Nevertheless, when typing the names of the authors in the Acoustics IRS, the letters "e" ("e") and "ë" ("yo") are equivalent, and the system searches for both spellings.

Since all Russian-language sources are entered in the IRS, and not only those published in the Russian Federation, sometimes authors from the sovereign states of the former USSR like to flaunt their "independence" and try to transcribe their names in a national way. As a result, three of the five letters in the surname of Silin, P., are typed in Latin coding instead of typing all the letters in Cyrillic coding. Or Sorokotyaga, A.S., which suddenly began to be transcribed as Sorokotyaga, O.S., i.e., Aleksandr became Oleksandr. Of course, this makes it difficult to search, and the user can decide that the author is absent in the database. In this case, it is better to search by keywords or by the name of the source. Heading search can also help.

Search by Keywords

In our opinion, this search is useful if it goes along the entire text of the article or, at least, by title and resume, which is implemented in our IRS. Searching only by headings or by "keywords" that often accompany an article is unproductive. See our article in Uspekhi Fizicheskikh Nauk [8], as well as the article by specialists from VINITI RAS [9].

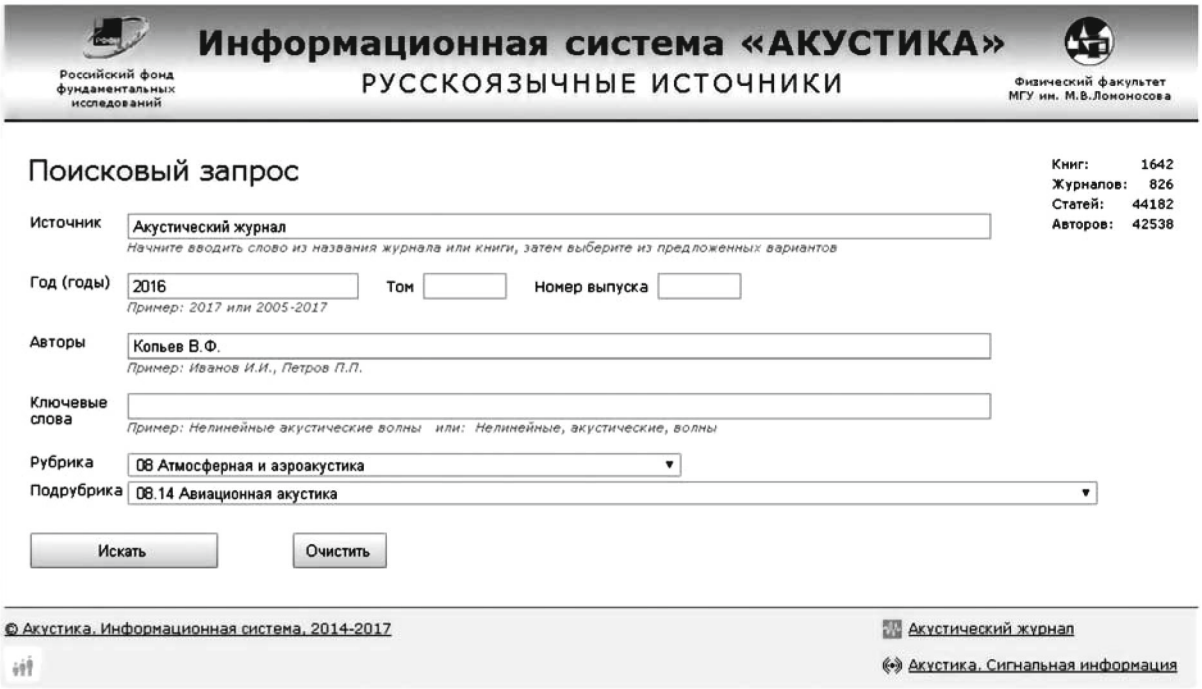

Fig. 1. Acoustics portal.

Search by Rubricator

Our information system also allows us to search by the rubricator used. The rubricator, like the "source", also appears as a suggestion when clicking on the input window. The rubricator has 16 rubrics, most of which are divided into subheadings. Search by rubricator, in our opinion, is extremely important, since the user can immediately take a look at most of the works on the topic in their historical development (from newest to earlier). If desired, the user also has the opportunity to follow the hyperlink to those rubrics that the editor considers necessary to mark as adjacent, or to receive all the articles of each of the authors. Note that the rubricator has three levels, if we take the first name of the Acoustics thematic area. Our extensive experience at VINITI suggests that three levels are probably enough, since the pursuit of their increase to 6–8, which was previously considered an achievement in the VINITI physics department, led to an inadequate picture when filling out the rubrics. As a rule, this depended on the professionalism, "taste", and responsibility of the editor, and did not correspond to the actual state of affairs in physics and made it difficult to find the needed works.

On the issue page, there is a complete bibliographic description of the document with a resume or abstract (Fig. 2).

In the framework of our project, for the bibliography we use the standard adopted in our previously developed technology for issues of the RJ Physics in the Department of Physics and Astronomy of VINITI.

The author's index, rubricator, article titles of the Acoustics Journal, and some other sources are provided with hyperlinks, which allows us to go to the page of the last articles of this author (up to 100), to the page of the last articles in this rubric (also up to 100), or the full text articles in pdf format if available in the system.

Lists of articles of authors or documents in the rubrics are generated automatically from the documents of the portal database. With filling the database, their filling also changes. To date, the database contains about 45 thousand documents over a 40-year period, and, according to Uspekhi Fizicheskikh Nauk, Acoustics Journal, and some others, the database contains all articles on acoustics topics for the entire period of their existence. In total, at the end of 2016, more than 40 000 authors were placed in the system database. Thus, the user receives all articles on acoustics from a given publication, all articles of a particular author, all articles of an author in a specific source, all articles for given years entered into the system, or all articles in a specific section, etc.

Let us briefly formulate the search directions in the Acoustics IRS.

Search for Publications

This is what the user usually needs. This search can be conducted in various ways. The easiest way is to enter the name of the publication in the Keywords field (a few consecutive words from the name are enough), and, almost certainly at the very beginning of the issue, one will receive this publication. It is not necessary to fill in other fields. Very rarely there are publications with the same name. Usually this hap-

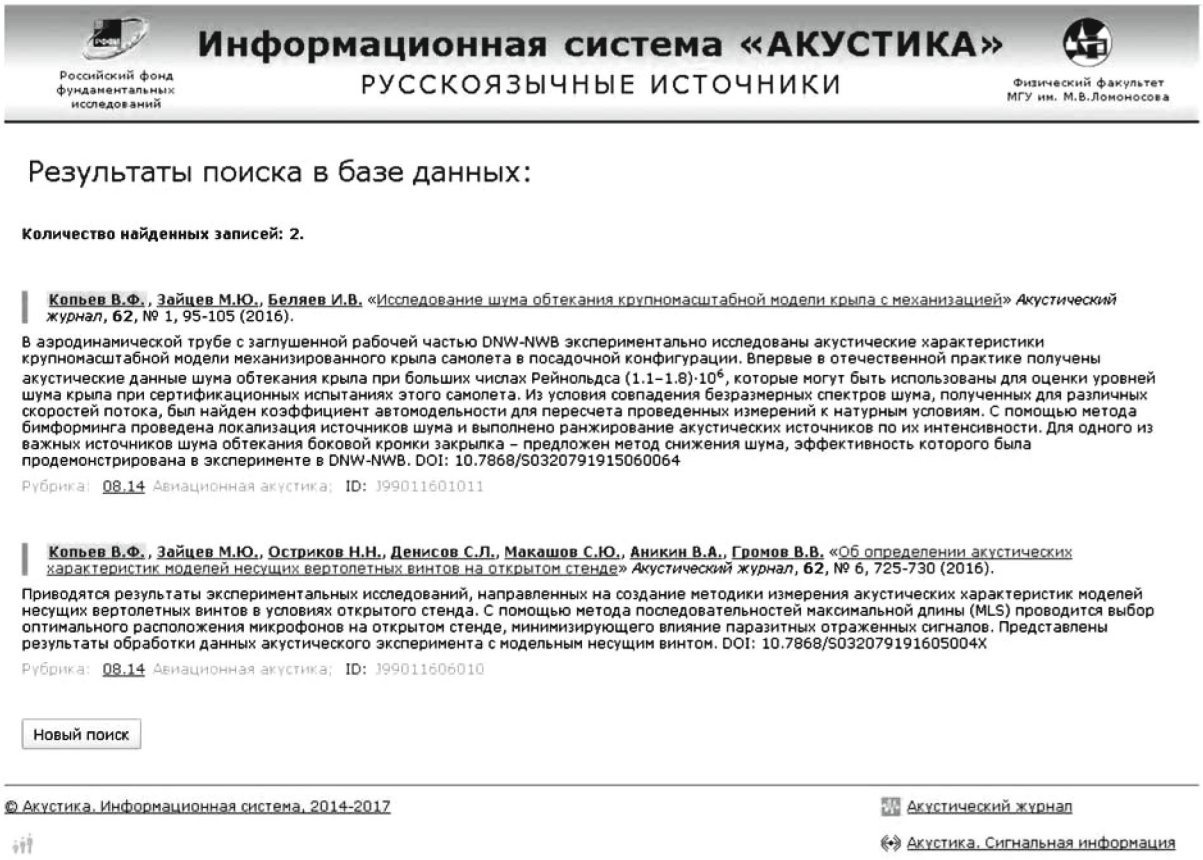

Fig. 2. Results of the output on a complex request: "Source", "Year", "Author", "Rubric", "Subrubric" (Fig. 1).

pens if the authors place versions of the same article in different publications, for example, in conference abstracts and in periodicals.

Search for Articles of Authors

If the user does not remember or does not know the exact name, then search by authors is possible. Upon receipt of more than 100 articles, it is necessary to clarify the year (or range of years) of publication or indicate the name of the journal.

Search for Acoustics Articles in Journals

It is necessary to enter the name of the journal in the Source field or select it from the drop-down list. Upon receipt of more than 100 articles, it is necessary to clarify the year (range of years) of publication.

Search by Rubricator

In the Rubricator field, it is necessary to specify the heading from the drop-down menu. If necessary, specify the Subheading from the adjacent drop-down menu. There are cases that, even with this refinement, the number of articles found exceeds 100. In this case, it is sufficient to indicate the authors or the interval of years of issue.

The Acoustics portal in the lower right corner has access points to the full-text version of the Acoustics Journal, the site structure of which we described in the Acoustics Journal [10], and to Signal Information (SI) [11]. The site of the Acoustics Journal archive is opened by all issues of the journal that have been published at the moment, from where it is possible to switch to the content of each issue with accompanying resumes. Each article is rubricated and, with the exception of the "content" section, is also placed in the corresponding section. There is a copyright index, which also provides the names of articles and their output. All this is connected by cross hyperlinks, which facilitates the search and transition to the full text of the desired article from the contents of the issue or any index. At the top right of the "content page", there are buttons to go to the previous or next issue of the journal. This option seemed to us very convenient.

Placing the journal's archive on the Internet and its presence there for the past five years, as we also expected from user reviews, has greatly simplified the search for journal articles. The online archive of the Acoustics Journal (http://www.akzh.ru) has a search system, not only in Russian, but also in English.

In the abovementioned article [10], we presented a 3D graph that gives the filling of rubrics by year of publication. This shows how the filling of rubrics

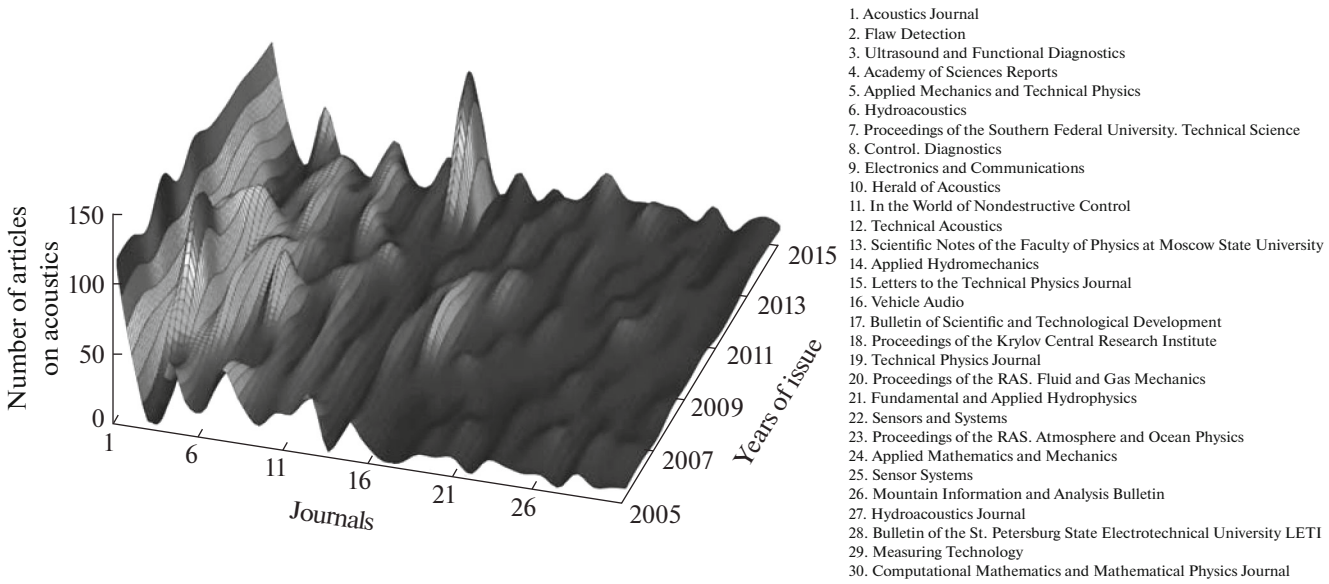

Fig. 3. Filling field of acoustics topics for the first 30 journals in 2005–2015.

(directions) changes over time. These data can be used for historical analysis of the development of various studies and for their forecasting. More details about the current state of the archive are described in the article [12].

We implemented the Signal Information website (http://akinfo.ru) as an information system for the current statistics of issued publications on acoustics with a frequency of two months. In addition, the SI preparation technology is used to edit the input retrospective information, because the Acoustics IRS filling goes both ways in time. The usefulness of SI is to provide an opportunity to glance at the current state of acoustics research in various directions. The site allows one to search by source, author, and rubricator, and to view the entire issue in pdf format. Almost simultaneously with the publishing of the next SI issue (and 25 issues have already been published since 2013), its information goes to the Acoustics IRS. Another argument in favor of creating this information Internet resource as Signal Information is related to the lack of any complete information on scientific research in the Russian-speaking world.

Created technologies for the preparation of these three information resources, in our opinion, are a complete system of information support for acoustics research in the Russian-speaking segment. They make it possible to ensure the full reflection of Russian-language scientific literature for monitoring and objective expert assessment of the scientific activities of Russian scientists, which can help strengthen the positive image of Russian science in the world.

Replenishment of the Acoustics IRS by current and archival documents today is carried out through the Signal Information also because the designed and used workstations for entering information into it and editing it are equipped with filters—formal-logical control of the input information, which facilitates flow control data entered into the database. In the future, the retrospective part in the database can also be supplemented by filling out existing, albeit scattered, databases, such as, the Science-MSU Information and Analytical System, the Math-Net.Ru portal, the Elibrary.Ru portal, and the conversion of the issues of the RJ Physics of the predigital era into digital form. The latter was demonstrated by us using issues of the RJ Mathematics of 1996–1997 [13]. In the future, archives of journals being created will also help.

Current information on fresh books, conferences, magazines, etc., collected in one place, allows, not only obtaining fresh statistic of scientific papers, as well as texts of articles from the Acoustics Journal, but also launching an information retrieval system based on a database that is formed and constantly being updated with current and retrospective documents.

We also note that, in addition to simply providing information, the Acoustics portal allows one to study the state of acoustics at the current time, to take data for subsequent analysis of trends in the development of a particular area of acoustics. In Figs. 3–9 there are examples of scientometric information that can be taken from this database. In [14, 15], we wrote that Bradford's Law of Scattering of scientific publications (the empirical regularity of the frequency distribution of articles in scientific journals), according to our data, should be written, not as $1 : n : n^2$ [16], but as $1 : n : n^3$. Having carried out the same research on the Acoustics database, which is the basis of the Acoustics IRS, we received confirmation of the earlier assumption (Fig. 7). From this we can conclude that, as when applying Zipf's Law (the empirical regularity of the distribution of words in the text) to different languages, Bradford's Law also has a dependence on the environment (the distribution of articles in scientific journals of a particular subject area).

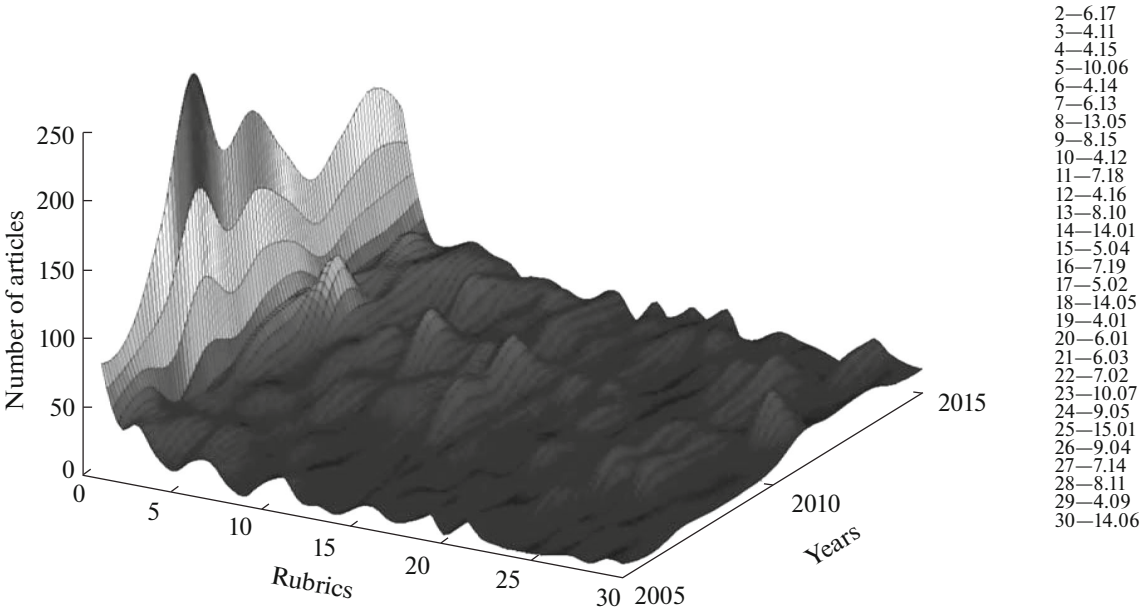

Fig. 4. Filling field of 30 rubrics of acoustics topics in descending order of the number of documents in them from 2005–2015.

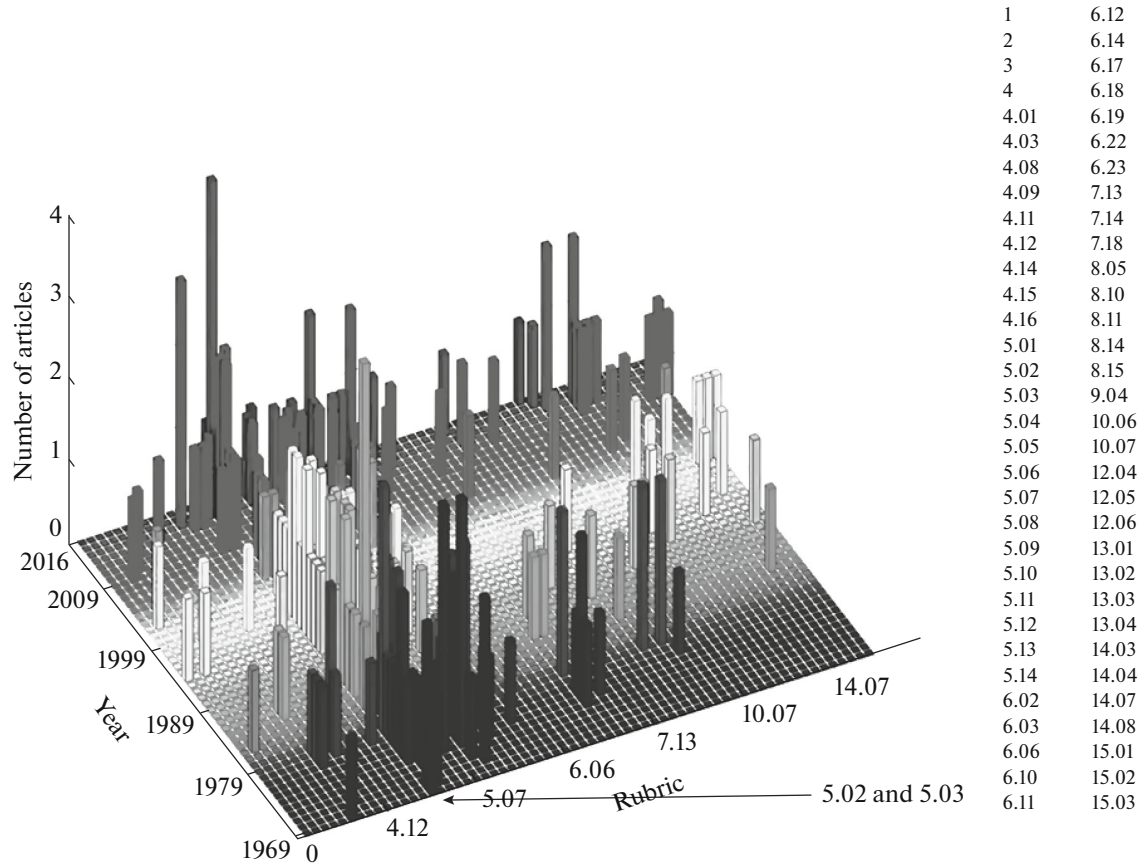

Fig. 5. Field of interest in the domain of acoustics for the entire period of public scientific activity of a leading author.

Let us pay attention to how the publication activity of the author (the number of articles N_P) affects the average number of coauthors N_{CA} in their works. Figure 8 shows data on the average number of coauthors for all 26475 Russian-speaking authors contained in the Acoustics database. It follows from this that this value is subject to great variation, although for the vast majority of publications the number of coauthors does

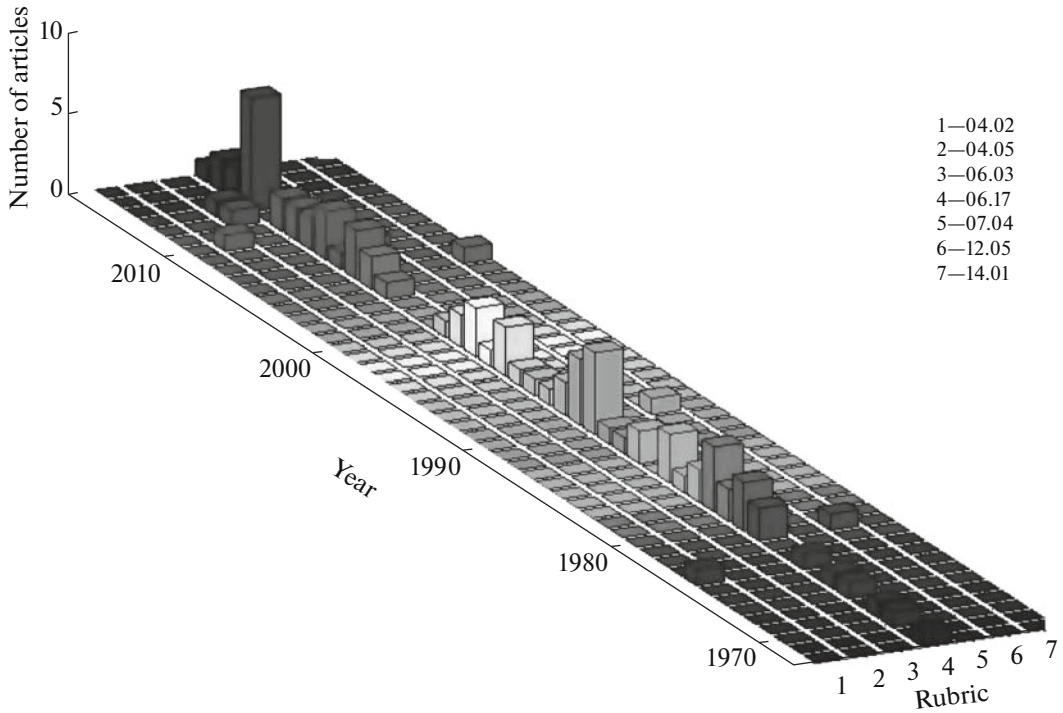

Fig. 6. Field of interest in the domain of acoustics for the entire period of public scientific activity of an author working in a narrow field of acoustics.

not exceed five. The average number of coauthors across all articles is 2.70. There are authors published in large groups of dozens of people. At the same time, there are single authors who have almost no coauthors throughout the entire scientific activity. For example, out of 135 works of Lapin, A.D., only five had one coauthor: Mironov, M.A. For Rudenko, O.V., who has the largest number of works, 257, the average number of coauthors is two, i.e., the average number of authors for articles is three. This also corresponds to his English-language works, where, according to information from Web of Science, the average number of coauthors is 1.58.

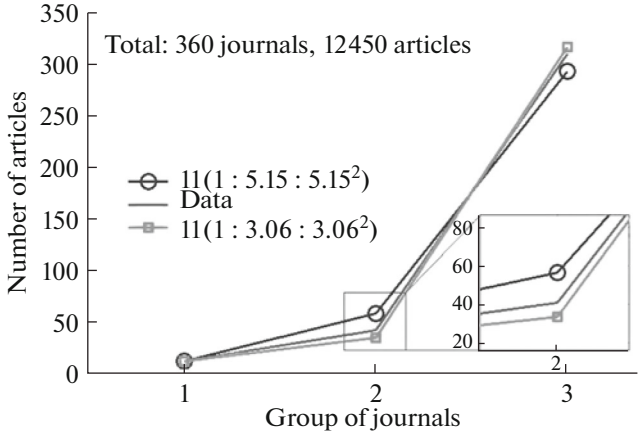

Fig. 7. Bradford's Law of Scattering of scientific publications is verified using acoustics topics.

We divided the authors by their publication activity into ranges of 10 articles. Author distribution over these ranges is shown in Fig. 9a. The vast majority of authors have 1 to 10 publications, and, with an increase in publication activity, the number of authors N_A decreases exponentially. Within each range, we determined the average number of coauthors $\langle N_{CA} \rangle$ for all authors. The result is shown in Fig. 9b. Vertical lines show the standard deviation. From the figure,

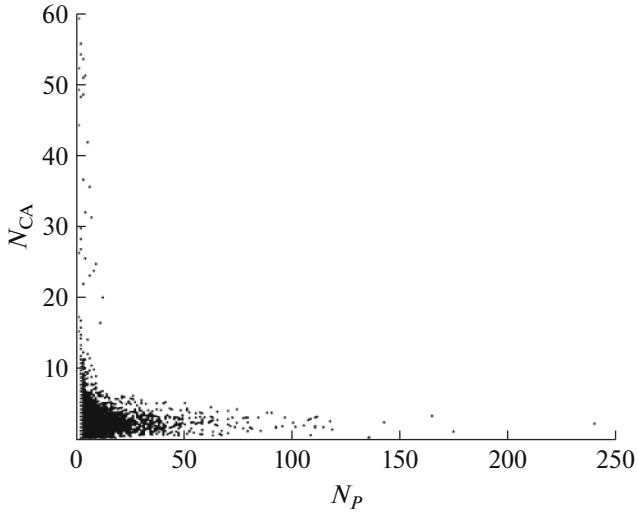

Fig. 8. Average number of coauthors N_{CA} in authors with number of articles N_P for all articles on acoustics topics located in the Acoustics database.

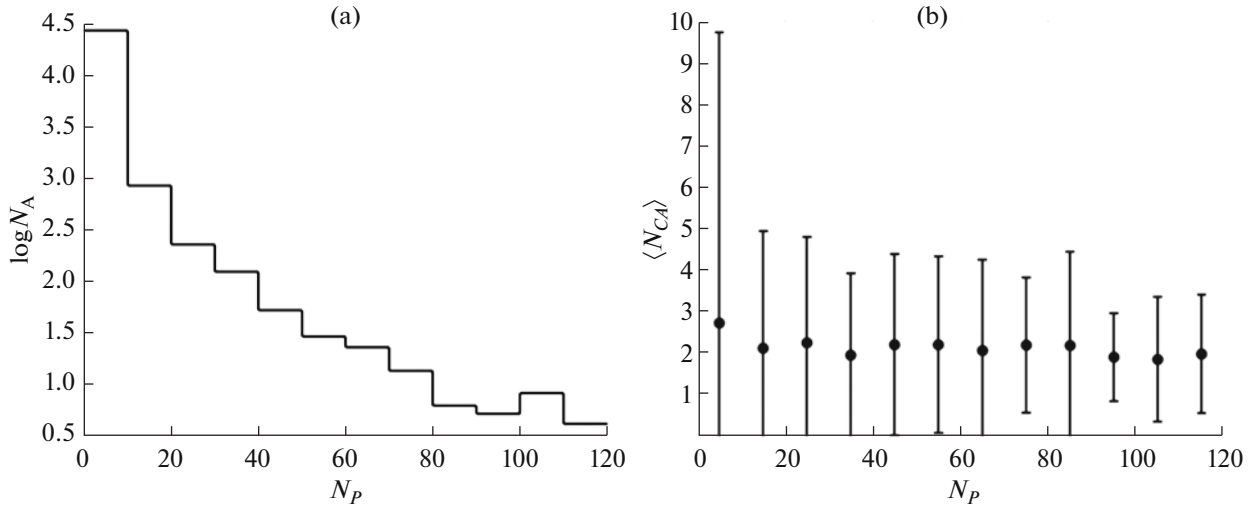

Fig. 9. (a) Number of authors N_A present in the Acoustics database, depending on their publication activity N_P; (b) average number of coauthors $\langle N_{CA} \rangle$ in the articles of authors present in the Acoustics database, depending on their publication activity N_P.

Table 1. An unclosed list of authors found in the IRS Acoustics, with a citation index of more than 1000 (data from ISI (Web of Sciences) with a retrospective depth up to 1986), with at least 10 publications in the IRS Acoustics

Surname, first name, patronymic	Citation index	Hirsch index	Primary place of employment
Balakshii, V.I.	1115	12	Faculty of Physics, MSU
Bunkin, F.V.	5449	33	Wave Research Center, GPI RAS
Vasil'ev, A.N.	4300	32	Faculty of Physics, MSU
Voloshinov, V.B.	1210	16	Faculty of Physics, MSU
Gavrilov, L.R.	1027	16	Andreev Acoustics Institute
Godin, O.A.	1566	18	University of Colorado, Boulder, Colorado, USA
Gulyaev, Yu.V.	5236	25	IRE RAS
Gurbatov, S.N.	1237	14	Faculty of Radiophysics, Lobachevsky State University of Nizhny Novgorod
Gusev, V.E.	4424	28	Université du Maine
Danilov, S.D.	1189	19	Alfred Wegener Institute for Polar and Marine Research in Bremerhaven
Zaitsev, V.Yu.	1112	18	IAP RAS
Karabutov, A.A.	3531	27	International Research Center of MSU
Kravtsov, Yu.A.	5899	33	GPI RAS
Kustov, L.M.	4542	31	Faculty of Chemistry, MSU
Landa, P.S.	2514	24	Faculty of Physics, MSU
Maev, R.G.	1163	15	Windsor U., Emanuel Institute of Biochemical Physics
Margulis, M.A.	2107	20	Andreev Acoustics Institute
Nazarov, V.E.	1266	18	IAP RAS
Ostrovskii, L.A.	4868	28	IAP RAS
Panchenko, V.Ya. (according to RSCI)	1237	11	Federal Research Center Crystallography and Photonics, RAS
Pelinovskii, E.N.	5202	30	IAP RAS
Preobrazhenskii, V.L.	1333	17	Wave Research Center GPI RAS
Pustovoit, V.I.	1262	14	Scientific and Technological Center for Unique Instrumentation, RAS
Romanov, V.P.	1491	17	Faculty of Physics, St. Petersburg State University
Rudenko, O.V.	5058	25	Faculty of Physics, MSU
Sazonov, S.V.	1151	15	National Research Center Kurchatov Institute
Sapozhnikov, O.A.	2071	23	Faculty of Physics, MSU
Sarvazyan, A.P.	3198		Artann Laboratories, Inc, Trenton, NJ, USA
Solodov, I.Yu.	1317	15	Faculty of Physics, MSU
Sutin, A.M.	2043	20	Artann Laboratories, Inc, West Trenton, USA
Talanov, V.I.	2977	24	IAP RAS
Khokhlova, V.A.	1471	19	Faculty of Physics, MSU
Charnaya, E.V.	1334	17	Faculty of Physics, St. Petersburg State University
Chirkin, A.S.	3066	26	Faculty of Physics, MSU
Shavrov, V.G.	2619	22	IRE RAS

despite the large scatter of the initial data, the average publication activity does not affect the number of coauthors, and this value stably lies in the range of 1.8–2.7 people, i.e., on average, each author has two to three coauthors for an article.

And last (but not least), we wanted to note in our work scientists working in acoustics with a citation index of more than 1000 as of January 30, 2017 [17] (Table 1).

The Acoustics IRS project gradually began to spread to other areas of physics and astronomy. As a first step, we uploaded to the IRS a full-text archive of the Results of Science and Technology VINITI (INT) in Physics and Astronomy. Earlier, we prepared a similar full-text archive for INT in mathematics [18], which found a place on the All-Russian Mathematical Portal (http://www.mathnet.ru) in the VINITI RAS section. To our regret, the check showed that, at the time of writing, the section of the VINITI site referenced by the portal was not working.

In conclusion, we note that the expansion of the Acoustics portal to all physics will provide access primarily to inaccessible scientific journals in Russian, which is a necessary part of the research work. This will allow us to find unresolved scientific problems, identify hot topics of science in its current state, generate new hypotheses, and establish mutually beneficial cooperation with foreign colleagues, whose search engines we are currently forced to buy.

We also inform you that the authors can replenish the IRS Acoustics by sending us lists of their published works (together with a resume), after making sure that they are not in the system. The lists will be checked by us, and, upon receipt of confirmation of the bibliography, will be entered in the IRS database. To do this, send an e-mail to akinfo2013@gmail.com with a bibliographic description of the work.

REFERENCES

1. A. I. Chernyi, *Russian Institute for Scientific and Technical Information: 50 Years of Serving Science* (Russian Inst. for Scientific and Technical Information Russ. Acad. Sci., Moscow, 2005) [in Russian].
2. V. G. Shamaev, Herald Russ. Acad. Sci. **81** (3), 321 (2011).
3. V. G. Shamaev, Nauchn.-Tech. Inform., Ser. 1. Organ. Metod. Inf. Rab., No. 7, 29 (2009).
4. http://www.radiotec.ru. Accessed February 16, 2017.
5. http://journals.ioffe.ru/. Accessed February 16, 2017.
6. http://sibran.ru/journals/. Accessed February 16, 2017.
7. A. B. Zhizhchenko and A. D. Izaak, Usp. Mat. Nauk **62** (5), 107 (2007).
8. V. G. Shamaev and A. B. Gorshkov, Usp. Fiz. Nauk **185** (11), 1235 (2015).
9. T. M. Leont'eva, I. P. Zhuravleva, and N. S. Pereverzeva, in *Proc. Int. Conference "Informational Society: Information Intelligence, Information Technologies", Moscow, October 24–26, 2007* (Russian Inst. for Scientific and Technical Information Russ. Acad. Sci., Moscow, 2007), p. 174.
10. V. G. Shamaev, A. B. Gorshkov, and A. V. Zharov, Akust. Zh. **59**, 283 (2013).
11. V. G. Shamaev, A. B. Gorshkov, and N. V. Shamaev, Akust. Zh. **60**, 109 (2014).
12. V. G. Shamaev, A. V. Zharov, and A. B. Gorshkov, Sci. Tech. Inf. Process. **34** (1), 10 (2007).
13. V. G. Shamaev, Sci. Tech. Inf. Process. **38** (2), 93 (2011).
14. V. G. Shamaev, Autom. Doc. Math. Linguist. **45** (1), 8 (2011).
15. S. C. Bradford, Eng.: Illustrated Wkly. J. (London) **137**, 85 (1934). http://journals.sagepub.com/doi/abs/10.1177/016555158501000407. Accessed February 8, 2017.
16. http://www.expertcorps.ru/science/whoiswho/ci86. Accessed February 12, 2017.
17. V. G. Shamaev and A. V. Zharov, Sci. Tech. Inf. Process. **34** (3), 88 (2007).

Translated by S. Avodkova