From the database of publications to the WEB-system for accounting of scientists' intellectual activity results

Svetlana A. Vlasova, Nikolay E. Kalenov

Joint Supercomputer Center of RAS – branch of Federal State Institution «Scientific Research Institute for System Analysis of RAS, Leninskiy pr., 32a, Moscow, 119334, Russia

Abstract

The article describes a WEB-system developed by the authors that implements services related to the formation and provision of multifaceted information about the results of scientific activities (publications, copyright certificates and reports at scientific events) of employees of an organization or a group of organizations. The system is focused both on the end user interested in obtaining specific data, and on the administrative staff, who generates reporting materials for the parent organization. The information base of the system contains metadata on the following classes of objects: persons (authors), organizations and their subdivisions; publications at analytical, monographic and summary levels; copyright certificates; scientific events (conferences, symposia, seminars); reports. The system includes two modules – an administrative one (intended for entering and editing data) and a user one, which is a special search engine that searches for information, visualizes it, provides navigation among related resources and exports data. A distinctive feature of the system is the introduced concept of "equivalent" objects. Objects are considered equivalent if they are represented in the system by different metadata, but referring to the same physical entity. Such objects are "persons" corresponding to one author with different spellings of the surname in the bibliographic descriptions of publications; organizations with different variants of names; articles published unchanged in various languages. In accordance with modern requirements for reporting on publications, the system reflects the sources of research funding, as well as the affiliations indicated in the articles for each author.

Keywords

scientific works, scientific activity, automated system, database, management reports, network technologies

1. Introduction

The effectiveness of fundamental scientific research is based on assessments of the results of the intellectual activity of their employees, reflected, first of all, in scientific publications and reports at scientific conferences. Recent organizational decisions in this area are aimed not only at quantitative, but also at qualitative assessment of publications and reports. In this regard, for each organization, the issues of creating a toolkit are becoming more and more urgent, allowing in an automated mode to register the results of the intellectual activity of employees and to promptly generate the necessary reporting data.

Considerable attention was paid to the accounting of publications of scientific organizations employees, starting from the thirties of the last century – all libraries of academic institutes were obliged to keep card indexes of employees' works. But this was an activity primarily aimed at helping readers who periodically needed to provide lists of their works for certification, obtaining a new scientific degree etc. With the advent of computer technology in libraries, the filing cabinets of employees' works

SSI-2021: XXIII All-Russian Conference on Scientific Services & Internet, September 20-23, 2021, Moscow (on-line), Russia EMAIL: vlas.svetlana2013@yandex.ru (S.A. Vlasova); nkalenov@jscc.ru (N.E. Kalenov) ORCID: 0000-0003-1533-5850 (S.A. Vlasova), 0000-0001-5269-0988 (N.E. Kalenov)



began to be replaced by databases with which library employees worked. The development of the Internet made it possible to switch to network technologies for creating and maintaining databases of employees' publications and to significantly expand the area of their application [1–7].

However, most of the supported systems reflect publications in a "traditional" form — based on a standard bibliographic description. One of the problems arising in this case is the different spelling of the names of the authors of publications, primarily transliterated from Cyrillic to Latin. To obtain a complete set of publications of a particular author, it is required to formulate requests containing all possible spellings of surnames. This problem is typical not only for local publication accounting systems, but also for the largest world-class systems. So, the surname "Королёв" in the Russian version is sometimes written as "Королёв", and sometimes as "Королев", which is permissible by the Russian rules for writing such surnames (see Fig. 1).

In publications reflected in foreign databases, the surname "Королёв" is transliterated in most cases as "Korolev", sometimes as "Korolyov", sometimes as "Koroljov". For example, the WEB of Science Core Collection (WoS CC) database contains 1128 publications by "Korolev A.", 107 publications by "Korolyov A.", 5 publications by "Koroljov A." (Fig. 2).

It should be noted that the possibility of "integration" of different spellings of the surnames of one author is implemented in the ORCID system (Open Researcher and Contributor ID) [8]. However, a test search shows that, at least for a number of Russian surnames, the ORCID search engine does not perform satisfactorily. So, when searching by the last name of "Королев", the system returns 36 records (Fig. 3), and when searching by the last name of "Королев" – 9 records (Fig. 4).

Another problem associated with the analysis of the publication activity of employees of a particular organization is renaming the organization or changing its status, in particular, when merging with other organizations, which is typical for the current moment of reorganization of the Russian scientific infrastructure. Internationally, they are trying to solve this problem within the framework of the Research Organization Registry Community [9].

One of the functions of an automated system that registers the results of the intellectual activity of research workers across the organization should be the formation of reporting data that meet the requirements of the Russian Ministry of Science and Higher Education. According to the latest regulatory documents, the data on the publication activity of the organization should take into account the affiliation of each author indicated in the publication, and the source of research funding to which article is devoted. The report should also include speeches at scientific events, indicating the authors, status of the event and status of the speech. To assess the personal contribution of an employee to scientific activity in many organizations, when forming internal reports, it is required to indicate which of the coauthors presented the speech.

Analysis of modern publications and accounting systems for the work of scientific workers, presented on the Internet [10–13], showed that none of them solves the above problems.

Below is a description of the system for recording the results of intellectual activity, developed by the authors, in which the above functions are implemented. The presented system is the result of the development of works previously carried out by the authors in this direction [14–17].

This year, the system received its further development in connection with the new requirements for reporting data on the results of the intellectual activity of scientific workers.

The new version of the system provides registration of information about the publications of employees, received copyright certificates, reports made by them at scientific conferences, symposia, seminars. The system provides for the reflection of the data necessary for the formation of various internal and external reports of the organization, and also allows you to solve the above problems associated with the ambiguity of the presentation of the names of the authors of publications and the names of organizations.

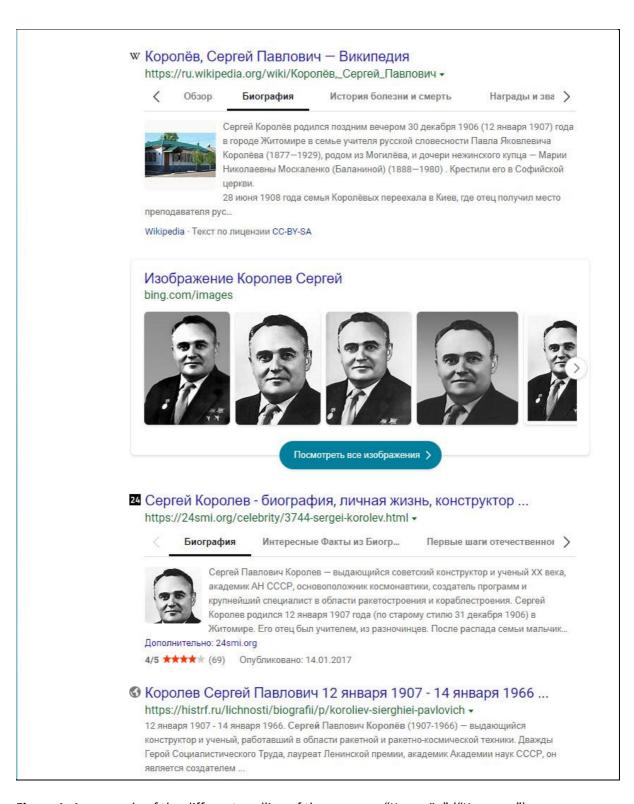


Figure 1: An example of the different spelling of the surname "Королёв" ("Королев")

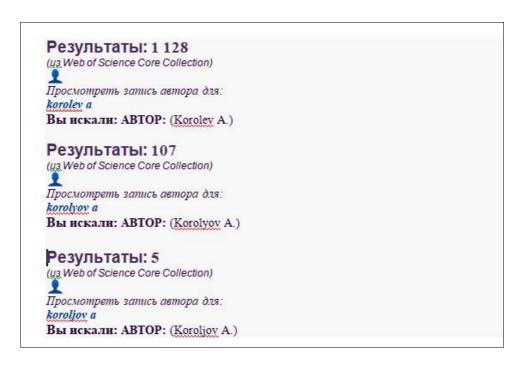


Figure 2: Various transliterations of the "Королев" surname in WoS CC

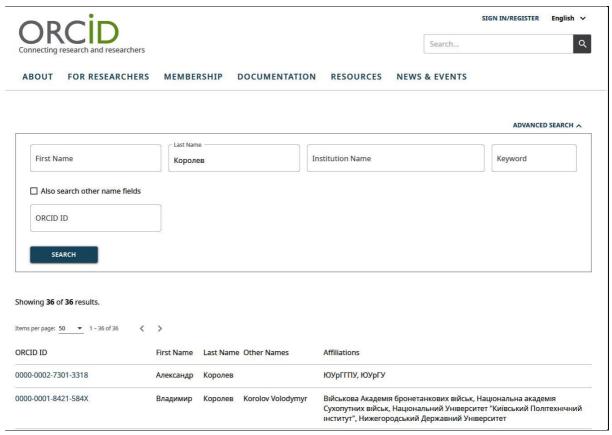


Figure 3: Search result by the name "Королев" in the ORCID system

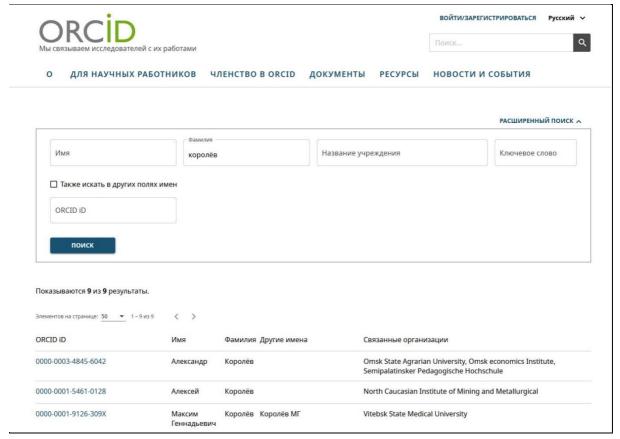


Figure 4: Search result by the name "Королёв" in the ORCID system

2. System structure

The system provides the creation and support of the following interconnected objects:

- ✓ publications at analytical and monographic levels;
- ✓ sources (journals and collections in which articles are published);
- ✓ conference reports;
- ✓ scientific events which reports were presented;
- ✓ persons (authors of publications and reports);
- ✓ organizations and their divisions.

In the new version of the system, in comparison with the previous one [16, 17] the metadata profiles of objects of the following classes are expanded.

Class "Person" - added person identifiers in ORCID, RSCI, Scopus, WoS systems.

Class "Source" – added ISSN, ISBN numbers.

Class "Publication" – added fields: number of the state assignment; information on grants that supported the research provided in the publication.

Class "Event" – added a field for the status of the event (Russian, international, regional, local).

Class "Report" – added fields: language (Russian, English); link to presentation of the report; link to the video of the speech.

In the new version of the system, instead of the connection between the objects "Person – organization", "Publication – person", "Report – person", the following links are implemented:

- ✓ Publication person organization (indication of the author's affiliation);
- ✓ Report person organizational unit;
- ✓ Report person speaker;
- ✓ Report person co-speaker.

3. Objects registrations in the system

Changes in the structure of metadata and object relationships have made it necessary to redesign the user interface of the registration process for publications and reports. Let's consider these processes in detail.

The first step of publication data input is entering its authors in the order presented in the publication. First, the required author is searched for in the system database by entering the initial fragment of the surname into the search line. The system will display a list of found persons (surnames are active links), as well as a link "New person". If the person you need is in this list, then you have to activate the corresponding link. The system will show the name of the person and related organizations that were registered earlier (Fig. 5). In the process of entering persons and organizations, you can point to equivalent records.



Figure 5: Entering the author and his affiliation in a new publication

After all the authors of a new publication have been entered, the system will provide a form for entering its metadata (Fig. 6). Depending on the type of publication (article, monograph, copyright certificate), the obligatory input of certain data fields is automatically checked. Mandatory to fill in when entering any type of publication are the title and year of publication. When entering an article, it is mandatory to indicate the source (journal, collection). The source input interface is the same as the author-related organization input interface described above. After all the necessary metadata has been entered, the publication will be registered in the system.

New report registration, as well as the introduction of a new publication, begins with input of its authors. The search for authors, the choice of organizations for them and their linking to the report are similar to entering the authors of the publication. For each author of the report, the status is indicated: speaker or co-speaker. Then the system provides a form for entering the metadata of the report: the title of the report; presentation type (plenary, sectional, poster, invited); language (Russian, English); address of the report presentation; address of the speech video recording. Figure 7 shows an example of entering the report metadata (the speaker's name is in bold).

After the report input information about linked with it conference (event) is entered (Fig. 8).

<u>Kalenov N.</u> Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)			
Sobolevskaya L. Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)			
Sotnikov A. Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)			
Add author			
Title of the article, bibliographic description Archival and Museum	Information as a Component of the Common Digital Space of Scientific Knowledge		
of the monograph, copyright certificate			
Publication type	article in conference materials		
Source	Proceedings of the 10th International Conference on Data Science, Technology and Applications Delete source		
Year of publication	2021		
Volume			
Nomer			
Pages	144-149		
Language of publication	English V		
Identifiers in external databases (DOI, WOS, Scopus)	DOI: 10.5220/0010512401440149 Scopus: 2-s2.0-85111780911		
State assignment number	0580-2021-0016 ×		
Grants (RFBR)	20-07-00773		
Full text address	publ/HH2238.pdf		

Figure 6: Registration of publication

N. Kalenov Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)				
G. Savin Научно-исследовательский институт системных исследований (НИИСИ РАН). Межеедомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН)				
I. Sobolevskaya Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)				
A. Sotnikov Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)				
Add author				
Title of the report	Some estimates of labor contribution for creating digital libraries			
Report type	sectional			
Language	English V			
State assignment number	0580-2021-0014			
Event	18th International Conference on e-Business	Delete event		
Presentation of the report	/bd/pr/2021/66.pptx			
Video recording of the speech				

Figure 7: Registration of report

Event title	18th International Conference on e-Business
Event type	conference V
Event status	international V
Venue (country)	
Venue (city)	Online
Dates of the event	year 2021 from 07 . 07 to 09 . 07 (day.month)
Link to the event website	http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=115968©ownerid=45217

Figure 8: Registration of event

4. User module of the system

The user block of the system (http://dirsmsc.ru/bd/) is a search engine that processes requests of varying complexity. Queries can include elements of all attributes of object metadata profiles, combined by logic operators "AND", "OR", "AND NOT". The logic of composing and executing queries by the system for various search fields is described in detail in a previously published work [16]. Here we will focus on the interface for providing the user with the information found and navigating through related resources.

The system provides the user with the opportunity to indicate the information about what classes of objects he wants to receive in response to his request directly in the resulting output. This can be data about publications, sources (magazines, collections), reports, events, persons or organizations.

The search interface allows you to process requests such as "find journals and collections in which in the period 2020–2021 articles by Joint Supercomputer Centre (JSCC)² staff were published, supported by Russian Foundation for Basic Research (RFBR)³ grants".

Scientific works of JSCC RAS employees			
Grant number ✓ РФФИ AND ✓			
AND V			
Organization V MCLI AND V			
AND V			
Report's title AND			
reports title AND V			
Year: from 2020 ✓ to 2021 ✓			
Show Journals / collections \checkmark by $20 \checkmark$			
Language All V			
Sorting Alphabet ✓ in direct order ✓			
Search Reset			
Records found: 12			
CEUR Workshop Proceedings (CEUR-WS.org) ISSN: 1613-0073			
Lobachevskii Journal of Mathematics ISSN: 1995-0802 (print version), 1818-9962 (electronic version)			
Proceedings of the 10th International Conference on Data Science, Technology and Applications ISSN: 2184-285X ISBN: 978-989-758-521-0			
Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering ISBN: 978-073814275-3			
Proceedings of the of the XXIII International Conference "Enterprise Engineering and Knowledge Management" (EEKM 2020), Moscow, Russia, December 8-9, 2020			

Figure 9: Request to search for sources

For this request (Fig. 9), 12 titles of journals and collections are issued, each of which is an active link, upon clicking on which a list of all publications from this journal (collection) available in the system is issued.

If in the request (Fig. 9) in the drop-down list "show" instead of "journals / collections" select "publications", the system will show a list of articles by the staff of the JSCC, in which links to RFBR grants are given. There are 26 such articles, they are presented in the form of a list (see the fragment shown in Fig. 10). Such list of publications contains their bibliographic descriptions and additional information (publication DOI, the state assignment ID and information about the grants allocated for research, reflected in the article).

² In Russian MCLI

³ In Russian РФФИ

Records found: 26		
1 <u>2</u>		
□ Igor Fiodorov, Alexander Sotnikov, Yury Telnov, Nixon Muganda Ochara Improving Business Processes Efficiency and Quality by Using BPMS // CEUR Workshop Proceedings (CEUR-WS.org), 2021 Vol. 2830 P. 327-336. DOI: 10.51218/1613-0073-2830-327-336. Scopus: 2-s2.0-85103226250 State assignment number: 065-2019-0014 PΦΦИ 19-07-01137 A, 19-57-60004/19		
□ Ivanov V.K., Palyukh B.V., Sotnikov A.N. Additive Criterion to Evaluate Object Innovation // Lobachevskii Journal of Mathematics, 2021 Vol. 42, - № 11 P. 2537 - 2544. DOI: 10.1134/S1995080221110111 Scopus: 2-s2.0-85118583562 State assignment number: 0580-2021-0016 PΦΦИ 20-07-00199		
□ Kalenov N., Sobolevskaya I., Sotnikov A. Archival and Museum Information as a Component of the Common Digital Space of Scientific Knowledge // Proceedings of the 10th International Conference on Data Science, Technology and Applications, 2021 P. 144-149. DOI: 10.5220/0010512401440149 Scopus: 2-s2.0-85111780911 PΦΦИ 20-07-00773		
□ N. Kalenov, I. Sobolevskaya, A. Sotnikov Mathematical modeling of the processes of interdisciplinary collections formation in the digital libraries environment // CEUR Workshop Proceedings (CEUR-WS.org), 2020 Vol. 2543 P. 391-398. Scopus 2-s2.0-85078449040 State assignment number: 0065-2019-0014 PΦΦИ 18-07-00893, 18-00-00372		
□ Nikolay Kalenov On the Formation of the Space of Scientific Knowledge Subject Ontology // CEUR Workshop Proceedings (CEUR-WS.org), 2020 Vol. 2784 P. 329-337. DOI: 10.51218/1613-0073-2784-329-337. State assignment number: 0580-2021-0016 PΦΦИ 20-07-00103		
□ Nikolay Kalenov, Gennadiy Savin, Alexander Sotnikov Fundamentals of Common Digital Space of Scientific Knowledge Building // CEUR Workshop Proceedings (CEUR-WS.org), 2021 Vol. 2990 P. 93-99. DOI: 10.51218/1613-0073-2990-93-99 Scopus: 2-s2.0-85119376588 State assignment number: 0580-2021-0016 PΦΦИ 20-07-00773		
□ Olga Ataeva, Nikolay Kalenov, Vladimir Serebryakov Ontological Approach to the Description of a Common Digital Space of Scientific Knowledge // CEUR Workshop Proceedings (CEUR-WS.org), 2020 Vol. 2784 P. 295-303. DOI: 10.51218/1613-0073-2784-295-303 PΦΦИ 20-07-00324, 18-00-00297, 18-00-00372		

Figure 10: Fragment of the found publications list

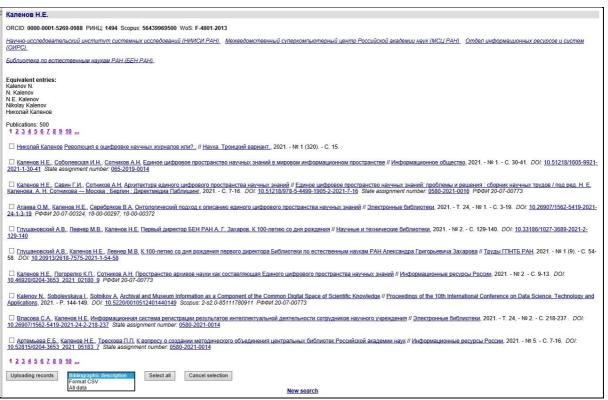


Figure 11: An example of the person information visualization

Authors, publication titles and names of the journal (collection) are active links leading, respectively, to information about these objects. Example of page with author information is shown in the Fig. 11. You can see there all the equivalent records for this person, as well as its ORCID and identifiers in the systems RSCI, Scopus, WoS, the names of all organizations associated with the person, its publications

and reports, are given out. Further, the system indicates the total number of person's publications registered in the system, and shows their descriptions. If the person is connected with several organizations, their names are active links. Clicking on the link of the selected organization will result in the display of publications of this person, in which this organization is indicated as the author's affiliation.

Having received a list of publications at his request, the user can mark the records of interest to him and unload them in one of three formats (bibliographic descriptions in text form, full information about the publication (in text form or in CSV format).

Descriptions of the conference reports found as a result of the search query contain: the authors of the report, the title of the report, a description of the corresponding event (Fig. 12).

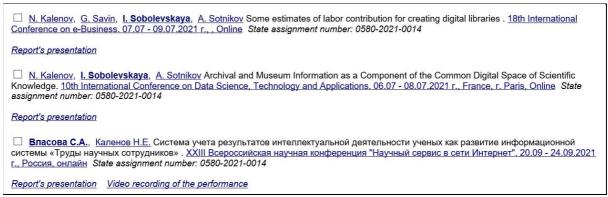


Figure 12: Example of conference report descriptions

The surname of the speaker is highlighted in bold in the description of the report. By clicking on the name of the event, you will be taken to its website. The names of the authors are active links, the transition to which will ensure the issuance of all of this author reports. If the addresses of the presentation and video recording of the speech are entered in the system, then the corresponding links will be located under the description of the report. As in the case of publications, the system allows you to unload the selected records of reports in various formats.

5. Conclusions

The presented automated system is currently operating in a technological mode at the JSCC of RAS. As of 01.12.2021 342 persons from 51 organizations, 818 articles published in 358 editions; 59 reports made at 35 events in 2018–2021 are registered in the system.

The system can be used in any scientific organization and provide invaluable assistance in the work of both researchers and the administration of the organization. A researcher in the search module of the system will be able to find his publications and conference reports for a given period of time and use this data for reports or compiling an article bibliography. Administration of the organization can receive comprehensive information about publications and speeches of employees for a given period of time, or data on articles carried out within the framework of a specific state assignment or grant.

6. Acknowledgements

The work was carried out at the JSCC RAS – Branch of Scientific Research Institute for System Analysis of RAS, within the framework of state assignment No. 0580-2021-0014.

7. References

[1] E. V. Beskaravajnaja, E. V. Dovbnja, S. S. Zaharova, Problemnoorientirovannye kollekcii. Formirovanie i analiz na primere bazy dannyh trudov sotrudnikov Instituta biofiziki kletki. Bibliografija 4 (2008) 30 –36 (in Russian).

- [2] S. S. Zaharova, Ju. A. Gureeva, Nauchnye publikacii: ot kartoteki trudov do bibliograicheskih profilej. Bibliosfera 2 (2017) 85–89. https://doi.org/10.20913/1815-3186-2017-2-85-89 (in Russian).
- [3] O. I. Levchenko, A. V. Solov'ev, Formirovanie bazy dannyh publikacij sotrudnikov Instituta fiziki tverdogo tela RAN, Informacionnoe obespechenie nauki: novye tehnologii: Sbornik nauchnyh trudov. M.: BEN RAN, 2015, pp. 215–221 (in Russian).
- [4] O. A. Rogoznikova, M. V. Danilin, Integracija bazy dannyh publikacij organizacii s indeksami nauchnogo citirovanija: realizacija sredstvami SAB IRBIS-64, Biblioteki i informacionnye resursy v sovremennom mire nauki, kul'tury, obrazovanija i biznesa: materialy Mezhdunarialy konferencii, (2015) (in Russian).
- [5] N. A. Mazov, V. N. Gureev, Bibliograficheskaja baza dannyh trudov sotrudnikov organizacii: celi, funkcii, sfera ispol'zovanija v naukometrii, Vestnik Dal'nevostochnoj gosudarstvennoj nauchnoj biblioteki 2 71 (2016) 84–87 (in Russian).
- [6] E. V. Kovjazina, BD trudov sotrudnikov kak sredstvo ucheta i prodvizhenija nauchnyh publikacij, Trudy GPNTB SO RAN 12 2 (2017) 336–343 (in Russian).
- [7] I. A. Pankratov, A. V. Ratushnyj, Proektirovanie informacionnoj sistemy dlja hranenija informacii o nauchnyh publikacijah, Vestnik molodezhnoj nauki Rossii 4 (2019) 31 (in Russian).
- [8] ORCID Open Researcher and Contributor ID. URL: https://orcid.org (accessed 1 December 2021).
- [9] Research Organization Registry Community. URL: https://ror.org (accessed 1 December 2021).
- [10] Baza dannyh publikacij IFTT. URL: http://www.issp.ac.ru/libcatm/publications_m.php (accessed 1 December 2021) (in Russian).
- [11] Publikacii sotrudnikov UlGU. URL: https://www.ulsu.ru/ru/page/page_1777/ (accessed 1 December 2021) (in Russian).
- [12] Publikacii sotrudnikov MIAN.
 URL: https://www.mi-ras.ru/index.php?c=mianpubs&l=0&jrnfilters[]=jhep (accessed 1 December 2021) (in Russian).
- [13] ISTINA (rukovodstvo pol'zovatelja). URL: https://docs.istina.msu.ru/getting_started/main.html (accessed 1 December 2021) (in Russian).
- [14] S. A. Vlasova, N. E. Kalenov, Novye podhody k formirovaniju baz dannyh publikacij sotrudnikov akademicheskih uchrezhdenij, Nauchnye trudy Instituta rukopisej Nacional'noj akademii nauk Azerbajdzhana 2 7 (2018) 85–94 (in Russian).
- [15] S. A. Vlasova, Avtomatizirovannaja sistema podderzhki korporativnoj bazy dannyh nauchnyh publikacij, Programmye produkty, sistemy i algoritmy 2 (2018) 42–46. https://doi.org/10.15827/2311-6749.27.311 (in Russian).
- [16] S. A. Vlasova, N. E. Kalenov, Informacionnaja sistema "Nauchnye trudy sotrudnikov akademicheskih uchrezhdenij", V sbornike: Nauchnyj servis v seti Internet trudy XXII Vserossijskoj nauchnoj konferencii. IPM im. M.V. Keldysha, 2020, pp. 152–165, https://doi.org/10.20948/abrau-2020-8 (in Russian).
- [17] S. Vlasova, N. Kalenov. Information System for Registering the Result of Scientific Institution Employees' Intellectual Activity, CEUR Workshop Proceedings 2784 (2020) 283–294, https://doi.org/10.51218/1613-0073-2784-283-294.