Assessment of the Relevance of Journals in Research Libraries Using Bibliometrics (a Review)

V. N. Gureev and N. A. Mazov

Information and Library Center, Institute of Petroleum Geology and Geophysics, Siberian Branch,
Russia Academy of Sciences, Novosibirsk, Russia
e-mail: GureyevVN@ipgg.sbras.ru, MazovNA@ipgg.sbras.ru
Received October 30, 2014

Abstract—Originally, most bibliometric research is aimed at the improvement of library collection-management methods and at the development of new methods for the selection of documents for a collection. This is also true of the impact factor, which was designed to evaluate journals before their inclusion in the collections of research libraries. Thus, the core of bibliometrics is librarianship. However, in recent years bibliometric studies are mainly conducted for the evaluation of scientific developments and individual scientists. In this case, a significant part of such research is carried out by workers at research libraries. Paradoxically, these approaches are not used in librarianship itself. With the expansion of the range of available tools, the development and use of bibliometric methods for the analysis of scientific information in libraries, especially in the acquisition process, once again has become relevant, which is consistent with the changes in science and publishing themselves. This has been stated by both foreign and domestic experts.

INTRODUCTION

In a study that was conducted in 2012 by EBSCO it was stated that with the use of modern analytical systems based on bibliometrics librarians can obtain comprehensive data on the information needs of researchers, which are necessary for the subsequent selection of journals for a library; recently, studies in this direction have become more active [1]. A. Meadows, an employee of the Wiley publishing house, has stated the high value of data about information needs [2]. According to Meadows, the publishers themselves have begun to provide such information on the basis of webometrics, usage of journals, and traditional bibliometric indices, which makes it possible to identify the needs of modern scientists to a certain extent. Employees of the Central Library of the Pushchino Science Center of the Russian Academy of Sciences assign a significant role to the study of the information needs of scientists using bibliometric methods. They point out that the study of information needs is one of the most important steps prior to all subsequent actions of libraries and information agencies on the organization of information and library support of users [3].

In addition to the use of information systems in their work, librarians have begun to provide new information services that are designed to help with the flow of information based on these systems [4]. An interesting project should be noted that proposes to use alternative indicators (metrics) [5]. Scientists themselves search for possible solutions on information retrieval [6], which indicates the lack of currently available services.

Nevertheless, currently all the modern services are associated with bibliometric analysis, which is probably one of the most advanced methods that makes it possible to identify the level of satisfaction of scientists with the current subscriptions more objectively and accurately in comparison with formal indicators and peer review. This can greatly optimize the acquisition of library collections.

A Russian researcher in the field of scientometrics, I.V. Marshakova, has identified three main advantages of bibliometric analysis in comparison with other methods [7].

1. Bibliometric analysis makes it possible to cover science in general.

2. It is possible to use different methods of analysis based on the material of international bibliographic databases (DBs).

3. In contrast to questionnaires, interviews, etc., in bibliometric analysis the "reified phenomena" are the source material, when papers have already been published and cited. Thus, the method becomes more objective.

The American researcher W. Moore reached a similar conclusion by comparing expert evaluation of

journals with an evaluation based on the citation indices [8]. Moore found a general correlation between the results of the two evaluation methods. However, this researcher stated that the indices that are based on citation are preferable for the assessment of the quality of journals and have a wider range of coverage of journals with different qualities.

This list can be supplemented with one more item that is associated with the elimination of the subjective opinions of acquisition librarians, which was pointed out by V.V. Shilov and G.M. Vikhreva [9−11]. According to Vikhreva, the elimination of the subjectivity of the acquisitions librarian in the selection of documents has become particularly relevant because of the need to constantly study new document flow phenomena and to promptly develop criteria for their evaluation [9].

Bibliometric analysis is carried out by special software according to specified algorithms (earlier it was conducted manually). It uses the mathematical formulas and methods of statistical analysis and is based on various input data. Primarily, scientific publications and citations can serve as such data. Consequently, the main bibliometric indicators can be divided into three classes; these are indicators based on:

(1) the number of publications (author, team, and organization);

(2) the number of citations obtained by these publications;

(3) the ratio of the numbers of publications and citations.

A detailed analysis of indices that are currently used can be found in the paper by A.V. Tsyganov [12].

To date, many varieties of this type of analysis has been developed. The most popular indices are based on the total number of citations ("simple bibliometrics" according I.V. Marshakova [7]), are identified by co-citation analysis ("structural bibliometrics" according to Marshakova), and are based on the impact factors of journals.

For collection acquisitions, citation analysis is the most commonly used bibliometric method, while the analysis of an author's publication rate is used less frequently. However, this method can be used only at large organizations with a large number of publications that can be used to make a representative sample. One of these studies was conducted by P. Davis at Cornell University. The author analyzed the journals in which the papers of the employees of the university had been published in the last 5 years. On this basis, Davis built a list of journals and proved the validity of Bradford's law [13]. In Russian library practice this method, in conjunction with a number of others, is used at the Central Library of the Pushchino Science Center, which is a fairly large entity comprising ten research and educational institutions [3].

Content analysis that can be applied to acquisitions and is not based on citation analysis can be another bibliometric method. Content analysis is the thematic analysis of papers that come from an organization with the goal of the identification of the information needs of researchers. Its results can then be used to collect and analyze thematically identical arrays of international publications and the selection of the journals with the largest number of relevant papers.

THE USE OF IMPACT FACTORS IN THE SELECTION OF JOURNALS

The impact factor, which in the Russian literature is also referred to as the factor or influence index, is a journal's rank, which was originally considered as a criterion for the selection of scientific periodicals for library collections [14]. According to Fedorets, the impact factor occupies an intermediate position between the assessment of the use (demand) and peer-reviewed scientific evaluation of a journal [15].

In the first years after the publication of the impact factors of journals research libraries showed a significant interest in this indicator. It should be noted that librarians conducted various experiments using the impact factor [16, 17]. For example, in Russian library practice the impact factor is actively used, along with expert assessment and demand for periodicals, at the Russian National Library [18, 19], at the Higher School of Economics for analyzing the quality of the contents of journal databases [20], and at VINITI RAS as one of the most important indicators of the multicriteria evaluation of journals [15, 21].

As the popularity of impact factors grew, the critical feedback also increased. Accumulated data on the disadvantages of the use of impact factors were described by M. Amin and M.A. Mabe [22].

The main problems that limit the use of impact factors include the following:

(1) It is not always possible to reproduce an impact factor, since the exact algorithm for its computation is not disclosed and it is unclear what types of publications Thomson Reuters considers in its formula [23, 24]. It is noted that rarely cited publications, for example, letters, are not counted in the denominator of an impact factor, but they are counted in the numerator if they are cited, which does make it possible to consider an impact factor as the average number of references per paper [25, 26].

(2) In calculating an impact factor, papers that are not cited are not taken into account, although they may represent a significant part of all publications. Thus, the editor of *Nature*, P. Campbell, when discussing the impact factor, gave the following figures: in 2004, 89% of the impact factor of the journal was formed by only 25% of the most cited publications [27].

(3) The impact factor lends itself to manipulation by the editorial boards of journals, which makes its

value doubtful. The most commonly used manipulations include self-citation, publication of the abstracts of all the papers of the journal for the previous year, an increase in the proportion of reviews, selection of manuscripts on popular topics, and quoting of papers that were published in the journal but that are not included in the denominator of the formula [26, 28].

(4) The 2-year period for the calculation of an impact factor that is arbitrarily used by Thomson Reuters is an issue [25].

(5) The inapplicability of impact factors to journals in national languages is a problem, since they are most compatible with English journals [29]. Therefore, it is recommended to create regional indices for the comparison of journals.

(6) The very foundation of the calculation of impact factors attracts specific criticism, i.e., papers are quoted only by scientists and for research purposes only. Therefore, the opinion of the journal of the rest of its audience is not taken into account. This, in particular, was mentioned by the editors of the top-rated journal *PLoS Medicine*. They stated that the audience of their journal also includes patients, health officials, school teachers, and nongovernmental organizations [23].

It is noteworthy that the function of the evaluation of scientific results, which was added to the impact factor later, has also been often criticized, especially by the scientific community [5, 30, 31].

The impact factor (like many other "global" parameters) is more applicable to acquisitions in universal large libraries, university libraries, and to a much lesser extent for acquisitions in the small libraries of scientific organizations. Impact factors have been successfully used as one of the criteria for the selection of journals, for example, in VINITI [15]. In smaller organizations it is also necessary to consider a number of factors that are relevant only to a particular organization [13, 32, 33]. In particular, the discrepancy of the impact factor for the use of journals in small libraries has been pointed out [34], because it is calculated based on average values around the world and the specific features of organizations are not taken into account [13].

Because of these drawbacks, attempts have been made to find a replacement for the impact factor. Thus, M. Thelwall urged that the nonscientific audiences of journals should be taken into account during the evaluation of their impacts. Thelwall is one of the initiators of the use of alternative indices that are based on the information collected on the Internet [35]. Attempts to find an alternative to the impact factor using bibliometric methods have also been made, for example, on the basis of the publications that are included in 10 or 25% of the most cited papers in the world [36].

A team of European researchers developed the SNIP and SJR journal indicators [37], which are used in Scopus and SciVal databases of Elsevier. The SNIP (Source Normalized Impact per Paper) is the impact per paper normalized by the source. It takes the differences in disciplinary characteristic features that can be used when comparing journals in various fields into account [38]. The SJR is the SciMago journal rank. It is an indicator of the reputation of a journal. Its calculation method is similar to the method of Google PageRank [39]. This indicator weighs the importance of a citation depending on the discipline, quality, and reputation of the journal from which a citation is taken. Thus, citations are not equivalent to each other. The SJR also takes the differences in the work styles of researchers in different disciplines into account and can be used to compare journals in different fields of knowledge. In contrast to the impact factor of a journal, the SNIP and SJR use a 3-year period of citations.

A recent study compared three journal indicators, the impact factor, SNIP, and SJR, with the expert analysis of journals [40]. In order to carry out a large-scale study, data that were collected by the Australian Research Council were used under the program aimed at the improvement of science in Australia. From 2007 to 2010, 700 researchers ranked 20 712 journals [41]. Another group of researchers compared these data [40] with three bibliometric journal indicators. The highest correlation of the expert review was established for the SNIP, despite the fact that compliance was observed in all three cases. This points to the prospects of using cheaper and faster bibliometric methods of information processing, as well as the need to improve traditional bibliometric indicators, because the SNIP gives better results in comparison with the impact factor.

A similar study was carried out for journals in librarianship, when nine bibliometric journal indicators, including the impact factor, were compared with the subjective expert assessment [42]. The study also pointed to the significant compliance of ranked lists based on expert analysis and bibliometric indicators.

It should be noted that with the advance of the electronic document flow the connection between citation and the impact factor started to weaken, although this seems paradoxical. As noted by a group of Canadian researchers led by G.A. Lozano, the reason is that publications in the electronic form became accessible to readers who associate them with a certain journal much less [43]. As a result, this leads to the fact that journals with high ratings lose a share of the publications with the highest citation rates; this share is growing in journals with lower impact factors. The declining share of high rating publications on economics in high-rank journals was noticed by the American expert G. Ellison [44]. However, the researcher goes further and relates this phenomenon as a whole with the outflow of publications from scientific journals and transition of researchers in economics to other forms of information dissemination.

Meanwhile, at present, information workers and librarians, while recognizing some drawbacks of the

impact factor, believe that it is too early to reject it and that it is much more accurate measure in comparison with the indicators that are based on a simple count of the total number of citations [45]. Researchers are encouraged to continue to use it in the acquisitions of research libraries as well. However, they should use common sense and not overlook the problematic aspects of this indicator that have been identified [45, 46].

THE USE OF REFERENCE BIBLIOGRAPHY (CITATION ANALYSIS)

The method of citation analysis in library acquisition has long been known. It has undergone a number of different modifications during its existence. It makes it possible to clearly outline the information needs of users, to identify tendencies of their development, to specify the core of journals that it is necessary to subscribe to, and to determine the sufficient depth of the subscription archive. It should be noted that before the Web of Science database was first released on CD-ROM, all studies were carried out manually, which was very time consuming. In addition to the tasks that concern the selection of new journals to collections of research libraries, the important task of writing off journals was also solved. It freed space on shelves, as well as placing the most used journals next to their subscription. With the transition of journals into the electronic form, these problems were solved (although new problems that are primarily related to archiving appeared). At the same time with respect to acquisitions, bibliometric studies gained new impetus and became much more accessible to the employees of libraries.

Differences in the application of the citation analysis method primarily consist in the choice of the source from which references are taken. These sources can be divided into global ones, i.e., those based on citations from journals and journal groups, and local sources, i.e., based on the citations of the publications of scientists who work in an organization served by a library [32].

Citation Analysis Based on the References from Primary Sources of Scientific Information

The citation analysis method was first applied to the acquisition of collections in 1927 by the chemical college librarians P.L.K. Gross and E.M. Gross [47]. Given the need for the rigorous selection of scientific periodicals for the collection of a small research library and the attempt to move away from expert assessment, which according to these authors is liable to the subjectivity of an expert, they proposed the following method of selection of journals. Based on the *Journal of the American Chemical Society* librarians compiled a list of all the journals that were cited in it in a given year, distributed them by year, and deleted references to the *Journal of the American Chemical Society* itself.

This journal was selected as the most significant in the chemical field; according to these researchers references from it could point out the information needs of the scientific community in the field of chemistry for the next 10 years.

After analyzing the list, the authors concluded that it is wrong to rely solely on citation data and that it is necessary to take other factors into account, such as the age of the journal, since younger journals that have fewer references will be still more necessary for the library because of their prospects and lower prices. The authors then drew attention to the possibility of evaluating the archival journal volumes based on the concentration of references to preceding years.

The study that was conducted by Gross and Gross was repeated 2 years later by librarians of another organization based on the example of journals in mathematics [48]. In their study they used not one but several journals for citation analysis, in order to cover a number of countries and languages and to have a more objective picture. Just 1 year later, in the same college a study was conducted in the field of electrical engineering [49], which was also based on several journals. This was partly due to the smaller number of references in individual journals than it was in *The Journal of the American Chemical Society*. In later research on geological journals using this method, P.L.K. Gross also broadened the base to six journals [50]. In the following decade, the Gross method in one form or another was repeated in a number of areas: radio physics [51], different areas of medicine [52—54], and other disciplines.

This method is still used. For example, in a paper that was published in 2007 American librarians described the preparation of a ranked list of 116 international journals in librarianship based on citations from 11 leading journals in this field [55].

Citation Analysis Based on References from Review Journals

In the 1970s, the Gross and Gross method was adjusted by the Indian researcher I.N. Sengupta [56], who proposed *Annual Reviews* as the basis for references. This choice was based on the fact that these journals are international, cover several areas of a certain discipline, and publish papers from all over the world on the most relevant areas of different disciplines. In addition, the reviewing nature of the published material involves a large number of references. The author also noted that in *Annual Reviews* new papers obtain citations much faster than in other journals. As a result, according to Sengupta, the cumulative list of references from papers of this journal reflects the diversity of countries and languages and makes the overall picture objective and representative. The author strengthened his argument with an analysis of biomedical journals; in the 1980s he conducted

another study in the field of biophysics using the proposed method [57].

Methods of reference analysis based on one or more journals were further developed in studies in which the authors rejected using groups of journals in favor of the entire set of journals from the database of Journal Citation Reports (JCR). The method of co-citation analysis was used in these studies [58, 59].

Co-citation Analysis Based on Data from Journal Citation Reports

With the advent of the electronic version of the database of the American Institute for Scientific Information and, in particular, of the Science Citation Index and the JCR database, it became possible to carry out more sophisticated studies based on co-citation when quantitative bibliometric studies developed into qualitative studies [7]. One of the first studies on the acquisition of library collections was the work of the American researcher K.W. McCain [59]. In this paper, the author created a core ranked list of journals in the field of genetics based on the analysis of data on the mutual citation of journals from the JCR database and without a starting point in the form of a set of papers or any journal. The author suggested that one use the subject categories of the JCR database as research journals in a particular field.

Citation Analysis Based on Secondary Sources of Scientific Information

In 1950, C.H. Brown from the American Library Association proposed an algorithm for ranking journals that was similar to the Gross and Gross method; it differed only in the fact that it was based not on a particular journal but a secondary bibliography [60]. In Brown's opinion journals whose abstracts frequently occur in abstracting journals (e.g., *Chemical Abstracts*) and thus pass some kind of expert evaluation should be of the most interest for acquisitions librarians. Brown proposed the method as a response to the shortcomings of the Gross and Gross method, in particular, the optional correspondence of the demand for the journal's citations, inadequate coverage of journals in other languages, and higher citation of old journals in comparison with new ones. However, a drawback was also found in the method that was proposed by C.H. Brown; in refereed sources journal titles with a greater publication frequency or larger volume would occur more often [56]. Therefore, the use of abstract papers as a basis for bibliometric analysis in the acquisition of collections does not provide an objective picture.

Citation Analysis Based on References from the Scientific Papers of Research Groups

This field began to develop in the 1970s. Thus, the American specialist E.F. Hockings used references

from technical reports of employees of multidisciplinary laboratories in the field of electronics for the ranking of journals. He documented the effectiveness of the Bradford law in a multidisciplinary profile library [61], whose cost effectiveness was proven previously by B.C. Brookes [62].

J. Ash, an employee of the Yale University Library, considered citations from master's and doctoral dissertations in order to optimize the acquisition of collections. Ash also demonstrated the effectiveness of Bradford law in the distribution of journals [63]. At another American university a study was conducted to consider references from dissertations for research library acquisitions [64]. In general, dissertations as the only source of references for their subsequent analysis were considered to be unsuitable for the optimization of acquisitions. They showed significant changes in the selection of cited journals in short periods of time, which, in turn, strongly depended on the dissertation topic, which did not always coincide with the profile of the organization and subject-typological plan of the library acquisition [63, 65, 66].

It is noteworthy that Bradford's law was effective for all scientific fields, including the cluster of the social sciences. It was resistant to variations in data-collection methods and to various data sources that are used by different researchers, and it did not depend on the breadth or narrowness of the scientific field and the age of the discipline, as was demonstrated in the work of the American experts C. Drott and B.C. Griffith [67].

A series of studies in this area was conducted by K. McCain (partially in collaboration with J.E. Bobbick) [59, 65]. Thus, in a paper that was published in 1981, McCain described a citation-analysis method that is based on references from the works of university employees, which were divided into three types: papers, dissertations, and coursework of postgraduates. Sets of cited journals were defined based on each of the three groups and then lists were compiled. It should be noted that for all three groups the core of the journals appeared to be common, whereas the peripheral journals were different. The authors associated this fact with the participation of different researchers in research programs. As the authors pointed out, a high citation index of a journal is direct proof of its use and the need for the library to subscribe to it, while journals with low citation indices can be considered as candidates for exclusion from the collection. An analysis of the distribution of references over the years made it possible for the authors to determine the year for each of the core journals after which they can be written off.

In Russian practice the citation analysis method was used at the RAS LNS [68], at the library of the Institute of Catalysis of the RAS [69], and at the Institute of Petroleum Geology and Geophysics of the SB RAS [70]. The work of N.A. Slascheva and Yu.V. Mokhnacheva from

the Pushchino Research Center of the Russian Academy of Sciences is of particular interest [3, 71]. In particular, the dissertation of Mokhnacheva describes a method for identifying the information needs of scientists based on two approaches [72]:

(1) First, citations for a certain period are analyzed based on papers of researchers that are served by the library.

(2) The citations of all the papers in a specific citation database are then analyzed by the same subject and for the same period.

A set of worldwide papers is revealed by author's keywords from papers of researchers of the organization that is served by the library. The method of content analysis was described by Mokhnacheva in earlier works [71]: based on the two obtained lists a single core is found, which is required to meet the information needs of scientists.

A citation analysis based on references from papers of the organization that is serviced by the library was used by the authors of this paper [73, 74]. We selected a local source of references, based on staff papers rather than an international database (e.g., one based on a journal). This choice is justified by the fact that every scientific organization is unique, existing and newly formed research groups and schools both work, special studies are carried out, and their own development strategies are developed. Under these conditions, drawing a repertoire of periodicals based on references from a particular journal in a given scientific field would mean that it is applicable to all of the dozens of organizations around the world that work in the same subject area. However, this egalitarian approach ignores the significant features of different organizations, although they are conducting research on a single discipline.

These features include national ones. Thus, in a list of journals that is based on references from the papers of employees, national journals will occur in the core, even with low impact factors, since they play an important role in the acquisition of scientific information in a particular country. In a list that is based on references from a specific journal these papers are likely to be in lower positions. Other features that are taken into account by a list that is based generally on references from the literature rather than by a list based on references from journals themselves include the uniqueness of scientific developments themselves and their corresponding information needs. In order to conduct various experiments researchers can, for example, require information from adjacent fields on an ongoing basis, which will be reflected in their citations and consequently in the resulting list of journals.

The inapplicability of bibliometric indicators of the global level to the specific needs of small libraries have already been pointed out by librarians. For example, the American researcher F. Davis discussed this in relation to the faculties of universities [13]. The need to consider the local characteristics of an organization, which should be reflected in the profile of acquisition of library collections, was also pointed out by the American expert T.E. Chrzastowski [33]. In particular, for these reasons some librarians refuse to use impact factors when evaluating journals for small collections.

It is important to note that scientists cite not all literature that is read for scientific purposes [75], but they select the most valuable sources for themselves, that can be considered as a kind of assessment and modeling of the group of sources that they need.

As noted by N.I. Podkorytova, the determination of journals based on references from papers of scientific organization employees is carried out in line with the involvement of the users themselves in the process of acquisition of collections, which is becoming a popular tendency [76]. By selecting certain references scientists are already indirectly involved in the selection of the best journals.

The Disadvantages of Citation Analysis

Despite the many advantages of citation analysis over other methods of evaluation of scientific information, for example, faster results, greater availability, and low cost, this method cannot be called perfect for the following reasons. In studies of the sociology of citations, the subjectivity of the citing author, which depends on different motivations, has been repeatedly pointed out. These incentives can include, for example, the desire to remain silent about new results, show their irrelevance, or, on the contrary, a private interest in some authors can occur; a wish to refer to a review that contains a large number of results instead of citing a few original papers, etc. The language barrier is an important factor, as preference in citation is given to literature in the language understood by an author. These disadvantages have been recognized by both the original developers of the method, who indicated that a list based on citations should first of all contribute to a more adequate evaluation of journals by domain experts [47], and those who used the method several decades after its discovery [56].

1. One of the major disadvantages of journal-evaluation methods based on citation is that they do not take their evaluative side into account and recognize positive and negative citations as equivalent. At the same time, as noted by researchers, the proportion of negative evaluative citations in individual works can exceed 50% [77]. In a detailed study of the nature of citation, A.P. Derevyanko and Yu.P. Kholushkin identified 54 different types of citation [77]. This fractional division makes it possible to evaluate citation quality, which makes the evaluation of a cited work more accurate. Until recently, such studies could be carried out only in manual mode, which limited their number. In 2010, researchers from the University of Oxford presented software that, according to its developers,

makes it possible to distinguish between positive and negative citations [78]. Perhaps, further development will automate the evaluation of a larger number of types of citation and the integration of the software with the citation accounting systems will make it possible to correct for their semantics.

2. Self-citation is an important problem, which is regarded as a negative phenomenon because of its excessive unethical use. There are cases where journal editors either directly or indirectly forced their authors to refer to other papers of the journal, used the practice of mutual citation with other journals, or quoted their journal from works that were organized by the editors of conferences [26]. In response to this practice the citation database has the means to exclude self-citations from the calculation of many indicators: the researcher citation count, the Hirsch index, the impact factors of journals, etc.

3. The disadvantages of the journal evaluation based on citations also include the fact that works with restricted access are not cited, for example, if a library does not subscribe to a journal and user has only its abstract. In this sense, an interesting observation was made by M.A. Azarkina that often readers are content with the literature that is available to them. On this basis, Azarkina concludes that a librarian should not fully trust readers [18, 19]. This point of view is supported by data that according to a poll that was conducted by the Library of the Russian Academy of Sciences; it was found that in 89% of all cases readers require literature that is available in the library. Obviously, from this it should not be inferred that other journals should not be added to the library.

Let us note the fact that Russian authors do not always cite open-access journals, despite their high impact factors, i.e., despite their active citation by foreign colleagues.

4. In papers that use the citation-analysis method, it is assumed that the authors of papers carefully study citations. However, American researchers indicate that often authors cite literature that they did not read [79, 80]. This occurs in the cases of the "borrowing" of references from other authors. Researchers have proposed algorithms that make it possible to identify the proportion of such citations among all the sources that are cited by authors.

5. While some researchers argue that there is a significant correspondence between the usage of journals and their citation level [32], others point to the lack of a direct correlation, as well as to the complex nature of citation, which is not considered in the analysis of citations in order to optimize the acquisition process for library collections [65, 81]. In a series of papers, the American researcher C. Tenopir also pointed out that the reading range of researchers is much broader than the list of cited sources. This is associated with the fact that a wider range of reading purposes exists in addition to purely research applications [75]. For example, these objectives include the desire to be knowledgeable in related fields of research or the preparation for a course of lectures for students. According to the observations of experts, over the years from one-third to one-half of all papers that were read were not cited [82]. European researchers G. Halevi and H.F. Moed claim that a correlation exists between reading and subsequent citation of papers [83]. The following values are given: the number of downloads for reviews, papers, and abstracts is two orders of magnitude greater than that of citations, while for editorial columns this index is four orders of magnitude greater. At the same time, these researchers recognize the low dependence of citation on reading for a number of disciplines where the readership includes not only scientists but also representatives of other professions who do not cite journals because they themselves are not authors.

6. Another drawback of conclusions that are based on quoted sources is that in recent years the citation model has narrowed. If earlier researchers cited literature based on what they read, in the era of electronic texts the model has shifted to citation of hyperlinks [75]. Thus, reading authors are influenced by already cited sources; here, Matthew effect comes into force, according to which highly cited papers gain more citations. As noted by S.D. Khaitun, this can also include the exaggeration of the role of notable scientists and the underestimation of small researchers. "At an equal quality of papers, the works of scientists of the high rank <…> are cited more frequently than the ones with low rank" [84].

Thus, citation analysis of references in its application to the evaluation of scientific journals has several disadvantages. At the same time, this method is more objective in comparison with other methods of acquisition of library collections, including expert assessment ones. It is much more accessible, especially with the expansion of services accompanying bibliometric databases and does not require additional staff and resources.

It is obvious one should not entirely rely on this method and use it as the only means for the evaluation of journals. However, its use in combination with other bibliometric methods can create the most objective and comprehensive picture of the process of journal selection in the collection of scientific periodicals.

USE OF AUTOMATICALLY EXTRACTED KEYWORDS (CONTENT ANALYSIS)

Content analysis is a less common bibliometric method. According to our data, this method has not been used for the selection of journals in collections of research libraries. We offered it for the first time [73]. The feature of the proposed approach to finding and evaluating scientific information is that it is based on the creation of a complex query to the database based

on keywords that are automatically extracted from the papers of an organization. This approach refers to quantitative bibliometric methods and involves the use of a natural language. According to the classification of the domestic researcher Khaityn, it is a thesaurus method [84]. The prospects of such approaches were mentioned by Marshakova. She pointed out that the "task of quantification solved on the vocabulary basis seems to be more flexible, since the authors of papers will use the language vocabulary at their own discretion" [7].

In order to achieve greater objectivity, keywords should be retrieved automatically. It is possible to use three types of keywords:

(1) KeyWords Plus is a development of the creators of the Web of Science database. These words are derived from the titles of papers to which the author of the original paper refers.

(2) words from controlled domain thesauruses that are assigned to papers in thematic databases.

(3) words that are extracted using software from the full texts of papers.

We have successfully tested a request that was constructed of the words of the first category (KeyWords Plus). Their selection was made due to the following issues:

1. These keywords are a product that is ready for use in a scientific library only if a subscription to the Web of Science database is in place.

2. KeyWords Plus is the result of many years of research at the American Institute for Scientific Information; it makes it possible to accurately identify the thematic focus of papers. KeyWords Plus selection algorithms have been detailed in the papers written by the developers [85, 86]. Summarizing these papers, we can say that KeyWords Plus is a normalized ranked list of the most frequent single terms and phrases from paper titles in the references.

3. As in the case with the analysis of citations, in this case references are also used, but the work is carried out not with the journals themselves. Here, we deal directly with the subject of research that is expressed in the titles of the cited papers. The semantic field is analyzed, rather than scientific periodicals. It is important that even with the subjective preferences that the author gives to some references over others, the reference subject in most cases is preserved. This makes the use of keywords more objective and preferable.

4. Based on the fact that a sequence is ranked by the frequency of occurrence KeyWords Plus is essentially a "brief retelling" of the entire paper and its abstract; the search query by keywords reflects an objective research theme that was stated by authors.

Based on the extracted keywords, a complex query to the database of citations is formed by KeyWords Plus. In response to this query, papers with exactly the same sets of keywords as in the papers of the employees of an organization that is serviced by a library are provided. The subsequent analysis of these papers makes it possible to reveal the core of the journals that contain the maximum number of papers that meet the information needs of the employees of a particular organization.

The main advantage of the proposed method lies in its accessibility and high thematic search accuracy, the relevance of the results, the lack of information noise, automation of the extraction of keywords for a subsequent request, and the objective nature of the approach.

CONCLUSIONS

The best bibliometric methods for use in the acquisition process of a library will be those in which data sets will be obtained from several sources. It is important that these sets correlate with each other. Researchers have already attempted to combine citation analysis with other methods. Thus, in the work of the Indian scholar S.N. Dhawan a model for acquisitions of collections of journal periodicals is proposed that combines three approaches: citation analysis based on primary sources, citation analysis based on secondary sources, and the usage of journals by readers [87]. In this case, the author understands the use of citations as indirect usage and contrasts this approach to direct usage, which is now measured in the number of downloads for an item. As noted by the authors, none of these approaches should be used as a completely sufficient selection criteria for documents to a library collection because each of them has certain disadvantages.

In a previous study the authors of this review also combined two bibliometric methods, i.e., citation analysis and content analysis, for the subsequent comparison and consolidation of journal lists [73]. Both approaches are based on the use of evidence-base of papers of employees of organizations that are served by libraries; the resulting data supplement each other.

The need for the use of several approaches has been indicated both by domestic and foreign researchers. Thus, as noted by Vikhreva, the wider a set of varying attributes is, the more correct the decision that is made by an acquisitions librarian will be [9]. Employees of the research department at Elsevier suggest that information on the same problem with the use of two, three, or even more different indicators of the evidence-base will ensure that the data that are obtained at the top of a triangle will be the most reliable [88]. Thus, the combined use of several bibliometric methods for the selection of journals in the collections of a research library can be sufficient to generate the optimal subscription list.

REFERENCES

1. *Serials Price Projections for 2013*, Ipswich: EBSCO Information Services, 2012.

2. Meadows, A., A Librarians and Societies and Publishers—Oh My!, http://scholarlykitchen.sspnet.org/2013/04/03/librarians-and-societies-and-publishers-oh-my/. Cited October 29, 2014.

3. Slashcheva, N.A., Mokhnacheva, Yu.V., and Kharybina, T.N., Izuchenie informatsionnykh potrebnostei pol'zovatelei Pushchinskogo nauchnogo tsentra RAN v Tsentral'noi biblioteke Tsentra (department of the RAS LNS), in *Biblioteki natsional'nykh akademii nauk: problemy funktsionirovaniya, tendentsii razvitiya: Nauchno-prakticheskii i teoreticheskii sbornik*, (Library of the National Academies of Sciences, Functioning Problems, Development Trends. Research and Practice and Theoretical Collection), Kiev: NLU, 2008, pp. 247−264.

4. Jurski, D., *2013 Study of Subscription Prices for Scholarly; Society Journals Society Journal Pricing Trends and Industry Overview*, Lawrence: Allen Press, 2013.

5. Priem, J., Taraborelli, D., Groth, P., and Neylon, C., Altmetrics: a Manifesto, http://altmetrics.org/manifesto/ (Cited October 29, 2014).

6. Artamonova, G.V., Galeev, A.R., and Bazhenova, T.S., An alternative method of assessing the relevance of medical and biological research of Russian NRU, *Public Health Manager*, 2012, no. 12, pp. 35−41.

7. Marshakova, I.V., *Sistema tsitirovaniya nauchnoi literatury kak sredstvo slezheniya za razvitiem nauki* (System of Citation of Scientific Literature as a Means of Tracking the Progress of Science), Moscow: Nauka, 1988.

8. Moore, W.H., *Citation Versus Reputation: Assessing Political Science Journals*, Tallahassee: Florida State Univ., 2000.

9. Vikhreva, G.M., *Tsennostnye aspekty otbora dokumentov v fond nauchnoi biblioteki* (Value-Based Aspects of Document Selection for a Research Library Collection), Novosibirsk: SPSTL, 2004.

10. Vikhreva, G.M., On axiological nature of library selection, *Bibliotechnye Fondy: Problemy Resheniya*, 2004, no. 6, pp. 1−4.

11. Shilov, V.V., The core of the library fund, *Bibliotechnaya Zhizn' Kuzbassa,* 1998, no. 4, pp. 134−140.

12. Tsyganov, A.V. Brief description of scientometric indicators based on citation, *Naukometriya Ekspertiza Upravlenii Naukoi*, Moscow: IPU RAS, 2013.

13. Davis, P.M., Where to spend our e-journal money? Defining a university library's core collection through citation analysis, *Portal* (Baltimore), 2002, vol. 2(1), pp. 155−166.

14. Garfield, E., Citation analysis as a tool in journal evaluation, *Science*, 1972, vol. 178(4060), pp. 471−479.

15. Fedorets, O.V., The use of the training sample to determine the priority of criteria in the rating system of evaluation of scientific journals, *Control Probl.*, 2009, no. 1, pp. 59−65.

16. Hirst, G., Discipline impact factors: a method for determining core journal lists, *J. Am. Soc. Inf. Sci.*, 1978, vol. 29(4), pp. 171−172.

17. Cawkell, A.E., Evaluating scientific journals with journal citation reports − a case study in acoustics, *J. Am. Soc. Inf. Sci.,* 1978, vol. 29(1), pp. 41−46.

18. Azarkina, M.A., Organization of journal fund of the research library. Acquisition problems, *Bibliotechnoe Delo*, 2007, no. 5, pp. 17−20.

18. Azarkina, M.A., Organization of journal fund of the research library. Acquisition problems, *Bibliotechnoe Delo*, 2007, no. 6, pp. 41−42.

20. Pislyakov, V.V., *Analiz kontenta vedushchikh elektronnykh resursov aktual'noi zarubezhnoi periodiki. Kolichestvennyi analiz v ekonomike* (Content Analysis of Leading Electronic Resources of Relevant Foreign Periodicals. Quantitative Analysis in Economics), Moscow: GU-VShE, 2002.

21. Kirillova, O.V., Andronova, M.B., Divil'kovskaya, T.Yu., and Khachko, O.A., New approaches and results of the information center evaluating the Russian flow of scientific journals: criteria and presentation of the ranked data, *Educational Technology & Society*, 2006, vol. 9(3), pp. 321−334.

22. Amin, M. and Mabe, M.A., Impact factors: use and abuse, *Medicina* (Buenos Aires), 2003, vol. 63(4), pp. 347−354.

23. The impact factor game: It is time to find a better way to assess the scientific literature, *PLOS Medicine*, 2006, vol. 3(6), pp. 0707−0708.

24. Rossner, M., Van Epps, H., and Hill, E., Show me the data, *J. Cell Biol.,* 2007, vol. 179(6), pp. 1091−1092.

25. Adler, R., Ewing, J., and Taylor, P., A report from the international mathematical union (IMU) in cooperation with the international council of industrial and applied mathematics (ICIAM) and the institute of mathematical statistics (IMS), *Stat. Sci.*, 2009, vol. 24(1), pp. 1−14.

26. Arnold, D.N. and Fowler, K.K., Nefarious numbers, *Notices Amer. Math. Soc.*, 2011, vol. 58(3), pp. 434−437.

27. Campbell, P., Escape from the impact factor, *Ethics Sci. Environ. Polit.*, 2008, vol. 8(1), pp. 5−7.

28. Falagas, M.E. and Alexiou, V.G., The top-ten in journal impact factor manipulation, *Arch. Immunol. Ther. Ex.*, 2008, vol. 56(4), pp. 223−226.

29. González-Alcaide, G., Valderrama-Zurián, J.C., and Aleixandre-Benavent, R., The impact factor in non-English-speaking countries, *Scientometrics*, 2012, vol. 92(2), pp. 297−311.

30. Seglen, P.O., Why the impact factor of journals should not be used for evaluating research, *Br. Med. J.*, 1997, vol. 314(7079), pp. 498−502.

31. *San Francisco Declaration on Research Assessment Putting Science into the Assessment of Research*, San Francisco: United States Society for Cell Biology, 2013.

32. Duy, J. and Vaughan, L., Can electronic journal usage data replace citation data as a measure of journal use? An empirical examination, *J. Acad. Librariansh.*, 2006, vol. 32(5), pp. 512−517.

33. Chrzastowski, T.E., Journal collection cost-effectiveness in an academic chemistry library, *Collection Management*, 1991, vol. 14(1−2), pp. 85−98.

34. Pan, E., Journal citation as a predictor of journal usage in libraries, *Collection Management*, 1978, vol. 2(1), pp. 29–38.

35. Thelwall, M., Journal impact evaluation: a webometric perspective, *Scientometrics*, 2012, vol. 92(2), pp. 429–441.

36. Leydesdorff, L., Alternatives to the journal impact factor: I3 and the top-10% (or top-25%?) of the most-highly cited papers, *Scientometrics*, 2012, vol. 92(2), pp. 355–365.

37. Colledge, L., De Moya-Anegón, F., Guerrero-Bote, V., Lopez-Illescas, C., El Aisati, M., and Moed, H.F., SJR and SNIP: two new journal metrics in Elsevier's Scopus, *Serials*, 2010, vol. 23(3), pp. 215–221.

38. Moed, H.F., Measuring contextual citation impact of scientific journals, *J. Informetr.*, 2010, vol. 4(3), pp. 265–277.

39. González-Pereira, B., Guerrero-Bote, V.P., and Moya-Anegon, F., A new approach to the metric of journals' scientific prestige: the SJR indicator, *J. Informetr.*, 2010, vol. 4(3), pp. 379–391.

40. Haddawy, P. and Hassan, S.-U., A comparison of three prominent journal metrics with expert judgement of journal quality, *Context Counts: Pathways to Master Big and Little Data*, Leiden, 2014, pp. 238–240.

41. Tiers for the Australian Ranking of Journals. http://www.arc.gov.au/era/tiers_ranking.htm. Cited Octorber 29, 2014.

42. Kim, M.T., Ranking of journals in library and information-science—a comparison of perceptual and citation-based measures, *Coll. Res. Libr.,* 1991, vol. 52(1), pp. 24–37.

43. Lozano, G.A., Larivière, V., and Gingras, Y., The weakening relationship between the impact factor and papers' citations in the digital age, *J. Assoc. Inf. Sci. Technol.*, 2012, vol. 63(11), pp. 2140–2145.

44. Ellison, G., Is peer review in decline, *Econ. Enq.*, 2011, vol. 49(3), pp. 635–657.

45. Bensman, S.J., The impact factor: its place in Garfield's thought, in science evaluation, and in library collection management, *Scientometrics*, 2012, vol. 92(2), pp. 263–275.

46. Moed, H.F., Colledge, L., Reedijk, J., Moya-Anegon, F., Guerrero-Bote, V., Plume, A., and Amin, M., Citation-based metrics are appropriate tools in journal assessment provided that they are accurate and used in an informed way, *Scientometrics*, 2012, vol. 92(2), pp. 367–376.

47. Gross, P.L.K. and Gross, E.M., College libraries and chemical education, *Science*, 1927, vol. 66(1713), pp. 385–389.

48. Allen, E.S., Periodicals for mathematicians, *Science*, 1929, vol. 70(1825), pp. 592–594.

49. McNeely, J.K. and Crosno, C.D., Periodicals for electrical engineers, *Science*, 1930, vol. 72(1856), pp. 81–84.

50. Gross, P.L.K. and Woodford, A.O., Serial literature used by American geologists, *Science,* 1931, vol. 73(1903), pp. 660–664.

51. Hooker, R.H., A study of scientific periodicals, *Rev. Sci. Instrum.*, 1935, vol. 6(11), pp. 333–338.

52. Jenkins, R.L., Periodicals for medical libraries, *J. Am. Med. Assoc.,* 1931, vol. 97(9), pp. 608–610.

53. Gregory, J., An evaluation of periodical literature from the standpoint of endocrinology, *Endocrinol.*, 1935, vol. 19(2), pp. 213–215.

54. Hackh, I., The periodicals useful in the dental library, *Bull. Med. Libr. Assoc.*, 1936, vol. 25(1–2), pp. 109–112.

55. Via, B.J. and Schmidle, D.J., Investing wisely: citation rankings as a measure of quality in library and information science journals, *Portal Libr. Acad.*, 2007, vol. 7(3), pp. 333–373.

56. Sengupta, I.N., Impact of scientific serials on the advancement of medical knowledge: an objective method of analysis, *Int. Libr. Rev.*, 1972, vol. 4(2), pp. 169–195.

57. Sengupta, I.N., The growth of biophysical literature, *Scientometrics*, 1985, vol. 8(5–6), pp. 365–375.

58. Servi, P.N. and Griffith, B.C., A method for partitioning the journal literature, *J. Am. Soc. Inf. Sci.*, 1980, vol. 31(1), pp. 36–40.

59. McCain, K.W., Core journal networks and co-citation maps: new bibliometric tools for serials research and management, *Libr. Q.*, 1991, vol. 61(3), pp. 311–336.

60. Brown, C.H. and Krumm, R.V., *Scientific serials: characteristics and lists of most cited publications in mathematics, physics, chemistry, geology, physiology, botany, zoology, and entomology*, Chicago: Association of College and Reference Libraries, 1956.

61. Hockings, E.F., Selection of scientific periodicals in an industrial research library, *J. Am. Soc. Inf. Sci.*, 1974, vol. 25(2), pp. 131–132.

62. Brookes, B.C., Optimum p% library of scientific periodicals, *Nature*, 1971, vol. 232(5311), pp. 458–461.

63. Ash, J., Library use of public health materials: description and analysis, *Bull. Med. Libr. Assoc.*, 1974, vol. 62(2), pp. 95–104.

64. Chambers, G.R. and Healey, J.S., Journal citations in master's theses: one measurement of a journal collection, *J. Am. Soc. Inf. Sci.*, 1973, vol. 24(5), pp. 397–401.

65. McCain, K.W. and Bobick, J.E., Patterns of journal use in a departmental library—a citation analysis, *J. Am. Soc. Inf. Sci.*, 1981, vol. 32(4), pp. 257–267.

66. LaBorie, T. and Halperin, M., Citation patterns in library science dissertations, *J. Educ. Librarianship*, 1976, vol. 16(4), pp. 271–283.

67. Drott, C. and Griffith, B.C., Empirical examination of Bradford's law and the scattering of scientific literature, *J. Am. Soc. Inf. Sci.*, 1978, vol. 29(5), pp. 238–246.

68. Dubrov, A.P. and Krasikova, O.L., The use of the citation analysis for studying the formation of the fund of foreign journals in scientific academic libraries, *Nauchn. Tekh. Biblioteki*, 1998, no. 6, pp. 26–34.

69. Pislyakov, V.V. and Lyubushko, E.E., The analysis of scientific and information activities (reading, publications, and citations) of scientists of the Institute of Catalysis of the SB RAS, *Kataliz Prom-st.*, 2007, no. 3, pp. 55–63.

70. Mazov, N.A., The evaluation of the flow of scientific publications of the academic institution based on the

bibliometric analysis of citation, *Inform. Technol. Gumanitarnykh Issledovaniyakh*, 2011, no. 16, pp. 25−30.

71. Slashcheva, N.A. and Mokhnacheva, Yu.V., Electronic information in scientometric studies, *Nauchn.-Tekhn. Inform., Ser. 1. Organ. Metod. Inf. Rab.*, 2003, no. 5, pp. 21−27.

72. Mokhnacheva, Yu.V., Data support of research by academic libraries using bibliometric methods, *Cand. Sci. (Pedagogic) Dissertation*, Moscow, 2008.

73. Gureev, V.N. and Mazov, N.A., Themes of the publications of an organization as a basis for forming an objective and optimal repertoire of scientific periodicals, *Sci. Tech. Inf. Process.*, 2013, vol. 40, no. 4, pp. 195−204.

74. Gureyev, V.N. and Mazov, N.A., Detection of information requirements of researchers using bibliometric analyses to identify target journals, *Inf. Tech. Libr.*, 2013, vol. 32(4), pp. 66−77.

75. Tenopir, C. and King, D.W., Electronic journals and changes in scholarly article seeking and reading patterns, *D-Lib Magazine*, 2008, vol. 14(11-12), pp. 1−13.

76. Podkorytova, N.I., Bosina, L.V., and Lakizo, I.G., System of the centralized acquisition of the CLS SB RAS. Results and prospects, *Bibliosfera*, 2012, no. 5, pp. 54−57.

77. Derevyanko, A.P. and Kholyushkin, Yu.P., The problem of the qualitative analysis of archaeological publications, in *Metodologiya i metodika arkheologicheskikh rekonstruktsii: Sb. nauch. trudov* (Methodology and methods of archaeological reconstructions. Collection of scientific papers), Novosibirsk: SB RAS, 1994.

78. Shotton, D., Cito, The citation typing ontology, *J. Biomed. Semant.*, 2010, vol. 1, suppl. 1, p. S6.

79. Simkin, M.V. and Roychowdhury, V.P., Read before you cite, *Complex Syst.*, 2003, no. 14, pp. 262−274.

80. Simkin, M.V. and Roychowdhury, V.P., Do you sincerely want to be cited? Or: read before you cite, *Significance*, 2006, vol. 3(4), pp. 179−181.

81. Sandison, A., Library optimum, *Nature*, 1971, vol. 234(5328), pp. 368−369.

82. Tenopir, C., King, D.W., Boyce, P., Grayson, M., and Paulson, K.L., Relying on electronic journals: reading patterns of astronomers, *J. Assoc. Inf. Sci. Technol.*, 2005, vol. 56(8), pp. 786−802.

83. Halevi, G. and Moed, H.F., Usage patterns of scientific journals and their relationship with citations, *Context Counts: Pathways to Master Big and Little Data*, Leiden, 2014, pp. 241−251.

84. Khaitun, S.D., *Naukometriya. Sostoyanie i perspektivy* (Scientometrics. Status and Prospects), Moscow: Nauka, 1983.

85. Garfield, E., Keywords plus—ISI's breakthrough retrieval method. 1. Expanding your searching power on current-contents on diskette, *Current Contents*, 1990, vol. 32, pp. 5−9.

86. Garfield, E. and Sher, I.H., Keywords-plus(tm)—algorithmic derivative indexing, *J. Am. Soc. Inf. Sci.*, 1993, vol. 44(5), pp. 298−299.

87. Dhawan, S.M., Phull, S.K., and Jain, P., Documentation notes, *J. Doc.*, 1980, vol. 36(1), pp. 24−32.

88. Colledge, L. and Verlinde, R., *SciVal metrics guidebook*, Netherlands: Elsevier, 2014.

Translated by O. Pismenov