ISSN 0147-6882, Scientific and Technical Information Processing, 2015, Vol. 42, No. 1, pp. 13–23. © Allerton Press, Inc., 2015.
Original Russian Text © A.B. Antopol'skii, 2015, published in Nauchno-Technicheskaya Informatsiya, Seriya 1, 2015, No. 1, pp. 14–24.

Algorithms and Methods that Measure the Level of Development of Information Resources at Scientific and Educational Organizations

A. B. Antopol'skii

Institute of Scientific and Educational Information, Russian Academy of Education, Moscow, Russia
e-mail: ale5695@yandex.ru

Received September 9, 2014

Abstract—This paper reviews the existing methods for monitoring and measuring electronic information resources, which are particularly applied in the fields of science and education. It also reviews a set of general indicators that characterize such resources (volume, scope, and composition), as well as indicators of use (popularity, citation, and usage, etc.). Possibilities for the integrated use of different instruments and methods that are applied to measure resources are discussed.

Keywords: scientific and educational information resources, monitoring of information resources, methods of measuring and evaluating information resources, development of resources, use of resources, internet resources

DOI: 10.3103/S0147688215010025

INTRODUCTION

Modern scientific and educational communications have become increasingly digital. Many traditional scientific publications (e.g., encyclopedia and reference materials) are almost completely digitized, while others (scientific periodicals, monographs, conference proceedings, and dissertations) either are in the process of digitalization, or coexist in traditional and digital forms. Similar trends can be observed in the educational environment. In Russia, the decision has recently been made to publish all high-school textbooks and teaching materials in both paper and electronic forms starting from 2015. Distance learning systems, which rely exclusively on the electronic format of learning and teaching, are actively developing at present.

A variety of new means of communication that were created due to the rise of new information technologies, particularly, the Internet, penetrate the scientific and educational landscape. Such means include websites, portals, social networks, wikis, databases, expert systems, interactive learning materials, and many other tools.

The creation of information resources, including in the electronic form, is the key measured outcome in the research area. Various information resources, which allow one to obtain quantitative assessments of such outcomes (publication activity, citation, impact factor, and h-index, etc.) have been developed and actively used in the research area over the recent years. For example, an overview of such tools was presented in the author's papers [1, 2]. However, the application of quantitative, including bibliometric, indicators for the assessment of scientific performance continues to be a widely discussed issue in the scientific community. For example, several interesting publications on this subject can be consulted on the website *The Reorganization of the Russian Academy of Sciences in 2013* [3].

The debates about the use of bibliometric methods are quite heated. Accounting for the results of information activities in science and education, which is enabled, among others means, by the use of quantitative indicators that allow one to objectively assess these results (i.e., scientific and educational resources) is undoubtedly important for many tasks, including scientific and educational management.

Clearly, the effective management of science and education relies on feedback, which makes it possible to assess management results. However, regarding information resources in the fields of science and education, such best practice has unfortunately not been reflected in the activities of public science and education regulation bodies, which is due to the inefficient spending of limited resources and the implementation of redundant projects that have no real value for research and education.

The implementation of nationwide information projects requires monitoring and quantitative accounting for digital information resources in the fields of science and education.

SCIENTIFIC AND EDUCATIONAL INFORMATION RESOURCES

One should note that scientific and educational information resources as an object of study have a number of significant features, as compared to other (both public and commercial) information resources.

First, they are primarily public. Restricted resources (e.g., the research results that are obtained in the defense sector or the personal data of scientists or students) are beyond the scope of this study.

Secondly, a significant number of such resources are created by using public funds, but not by public authorities. Therefore, the question of whether such resources are public has not yet been resolved. The author of this paper consistently advocates for assigning publicly funded resources to the public domain. However, Russian legislators do not share the author's point of view.

Thirdly, the open-access movement has emerged as a key alternative approach to the commercial use of information activities and is actively developing in the area of scientific information (for more information about this movement, see [4]).

Finally, due to the specific nature of the producers and consumers of scientific and educational information, the scientific and educational sector is the most susceptible to technological innovation. It is known that the spread of the Internet (after it was handed over by the American military to the civilian sector) occurs most rapidly in the academic environment. The same applies to other IT that is used to create and apply information resources.

In this context, one should note that the distinction between scientific and educational information resources has no particular sense, although separate resources certainly exist both in the research and education areas. However, their overlap is very significant. In our view, it extends to more than half of the entire scientific and educational information space.

One can conclude that it is appropriate to view the metrics of scientific and educational resources (i.e., indicators and methods of measurement) as being independent, although one should certainly consider the practices and methods of measurement of information resources in other areas. At the same time, it makes sense to present scientific and educational resources as a whole.

THE STATISTICAL PARAMETERS OF INFORMATION RESOURCES

It is appropriate to begin the quantitative analysis of information resources with an overview of the statistical data contained in the official Russian and international documents. One should note that the related possibilities are very limited. Specifically, the founding document "Strategy for the Development of the Information Society in the Russian Federation" [5] assigns only three indicators to information resources:

(1) the share of archival funds, including digitalized audio and video archives;

(2) the share of digitalized library collections in the total quantity of funds possessed by public libraries;

(3) the share of electronic catalogs in the total volume of catalogs possessed by the Museum Fund of the Russian Federation.

The international Knowledge Economy Index statistical project refers to a single indicator: "The number of publications in scientific and technical journals per 1 million inhabitants."

Overall, public statistics do not properly reflect the state of the national information resources. Thus, the most detailed statistical study on the development of the information society in Russia, which was published by the Higher School of Economics [6], does not refer to information resources. The only class of indicators that are relevant for information resources is a set of economic indicators that were calculated for organizations that operate in the content and media sector, including a single full-scale figure "The production of books, newspapers, and magazines." The *Indicators of Science* section contains a set of indicators that refer to publication and patent activities [7] (*Appendix A*). These indicators do not describe scientific information resources sufficiently, especially, taking digitalization of scientific resources into account.

The *Indicators of Education* section does not contain any indicators that refer to information resources in the field of education.

The above situation is puzzling. After all, the quintessence of the concept of an information society consists in the new quality of production and consumption of information, particularly that achieved through new forms of communication. At the same time, the main product and outcome of information activities, viz., information resources, de facto is not subject to public monitoring. Accordingly, there are no statistics on this outcome, which would provide feedback for public monitoring.

However, over the recent years, the government has been taking steps to intensify efforts in order to develop information resources, namely, through the support for "open government" and the activity of the Council on Public Data under the Government of the Russian Federation [8].

A new regulatory document that was issued by the Ministry of Education and Science of the Russian Federation regulates the methodology for assessing the effectiveness of scientific and educational organizations [9]. It also includes a set of indicators that reflect the development of information resources in the field of science and education:

(1) the number of publications of an organization, indexed in Russian and international information and analytical scientific-citation systems;

(2) all the citations that were obtained by the publications of an organization, indexed in Russian and international information and analytical scientific-citation systems;

(3) the cumulative impact factors of journals in which an organization's papers were published.

One can see that these figures rely on classical methods of bibliometrics. They do not take modern forms of digital scientific communications into account, especially those that are related to the quantity and quality of resources that are created by scientific organizations.

The only indicator that is directly related to the information resources of scientific organizations is the following:

(18) the number of visits (traffic) to official websites and (or) webpages of organizations hosted on the Internet information and telecommunications network (based on independent traffic counters).

Without doubt, the inclusion of this indicator in a regulatory document is positive. However, there are a variety of methods that allow one to measure traffic activity but result in highly different results.

The only sector in which information resources are reflected in statistics, at least to some extent, is library activities. The recently organized All-Russia Library census refers to the following indicators:

Digital Resources

Total number of private databases, thousands of records	of these	
	bibliographic databases	including the size of a digital directory

The Development and Use of Library Funds

Indicator title	including digital issues
Consisting of items by the end of the reporting year	
Items received in the reporting year	
Items lost in the reporting year	
Items delivered in the reporting year	

Information Library Services

The number of visits to a library's website

One could argue that these figures do not sufficiently describe the state of library information resources. Therefore, a more detailed set of indicators, which is presented in *Appendix B*, was developed in the framework of a study on the informatization of cultural institutions [10]. This set was experimentally tested; however, it did not achieve widespread use.

One should note that the monograph *Information Resources of Russia* [11] gave a detailed review of the statistics that are used to evaluate information resources in general and in different areas in particular. It concluded that the public statistical effort in this field is unsatisfactory. Although this monograph was published a decade ago, nothing has changed since that time.

The above-cited monograph suggests the following common base for the division of statistical indicators for information resources: indicators of resources (this is related to the formation and storage of resources) and indicators of the use of resources (which are called indicators of library services in librarianship). We will apply this division below.

One can refer to an interesting, although rather controversial, list of information-resource parameters, which is presented in the material posted on a website [12] that is dedicated to web design; unfortunately, it is anonymous. This list includes the following elements:

- content;
- scope;
- time;
- source;
- quality;
- compliance with requirements;
- method of fixing;
- language;
- cost.

Each parameter is analyzed in detail. Quantitative methods are discussed for some of the parameters.

However, this attempt to parameterize information resources is rather exceptional. It has not gained any official distribution or use.

BIBLIOMETRICS

Bibliometrcis remains as the basic discipline that studies problems related to the quantitative analysis of information resources. A detailed review of bibliometric approaches and methods is beyond the scope of this paper, since numerous publications have been devoted to this issue; for example, one of the most common works is the above-cited monograph by S.V. Bredikhina and A.Yu. Kuznetsov [2], which also contains a comprehensive list of bibliographic sources on the subject.

However, it is necessary to highlight one important limitation: the existing bibliometric methods almost exclusively rely on the measurement of scientific periodicals. However, top global bibliometric services (the Web of Knowledge and Scopus) have recently been considering monitoring other types of publication (e.g., monographs and conference proceedings). However, generally these services do not offer a sufficient coverage of scientific and educational resources, especially new types of resources that are widely available online. Generally, bibliometric methods, for all their undoubted merits, are open to sharp criticism. The key point of the criticism is the fact that according to Goodhart's law the use of formal indicators for science management leads to a distortion of the meaning and quality of scientific activity. Goodhart's law suggests that once a social or economic indicator becomes the target of a social or economic policy, it loses its credibility.

In recent years, there has been a movement against the use of bibliometric indicators. For example, the UK Government rejected the use of the Web of Science citation index in a number of newly adopted regulations that are applied for the assessment of scientists and scientific institutions [13]. At the same time, these regulations allow the use of tools such as Google Academia, at least for a number of scientific fields.

In Russia, one should note an important event that took place in the field of bibliometrics in 2013. This was the launch of a resource called the Map of the Russian Science, which was commissioned by the Ministry of Education and Science of the Russian Federation [14]. This project aims to integrate the most important Russian and international resources; therefore, its potential is very high. However, the current state of this project (it is currently on trial), which has been convincingly criticized, suggests that the actual use of the Map of Science [15] can only be expected after a significant upgrade. At the same time, it is unclear what the prospects of the Russian Science Citation Index (RISC) will be after the launch of the Map of Science and what the relationship between these two highly reputable resources will be.

Bibliometrics remains an important tool for the evaluation of scientific results; however, its limitations are quite apparent. To compensate for these limitations, which are primarily due to the recent development of new means of scientific and educational communication, new services focused on measuring web resources have emerged over the recent years.

TAKING ACCOUNT AND CLASSIFICATION OF INFORMATION SCIENTIFIC AND EDUCATIONAL RESOURCES

Taking an account of the web resources is an independent problem that is related to their measurement. Indeed, in order to measure resources it is necessary to define what scientific and educational resources are and to agree on the classification and definition of each type. Our experience shows that this is not an easy task.

Many catalogs of scientific and educational online resources exist. However, they differ greatly in their scope, completeness, accuracy, and other parameters; for example, the list of digital library catalogs is one of the most relevant and useful types of scientific and educational resources. A list of such catalogs, indicating the number of covered digital libraries, is given in *Appendix C*. This list shows that catalog authors interpret the concept of a digital library and the related concept of document collection in different ways.

At the same time, the inventory and assessment of electronic document collections in the scientific and educational area can be useful to identify the items that can be included in the National Digital Library [16] or the resources that can be included in the National System of Web Resources. The development of the latter system was foreseen under the Foundations of the State Cultural Policy project [17].

One of the important tasks that is associated with the selection of indicators for the assessment of digital information resources is their classification. Clearly, one should use different indicators for document collections, such as digital libraries and periodicals, than one uses for individual documents, even in those cases where a single document is comparable in terms of its size to an entire collection. Similarly, it is inappropriate to compare text and audio documents, educational, methodical, and scientific materials, etc. In this case, it might be useful to differentiate resources by discipline and level of education, for example, as is done on the Russian Education portal based on the RUSLOM metadata standard.

There may be other means of differentiation, for example, by novelty, language, software tools, and formats, etc. In any case, one should aim at assessing homogeneous groups of resources. However, the consistent application of an in-depth classification is absurd, unless all resources that are taken into account are comparable. Therefore, along with the differentiation, it is necessary to use mechanisms that allow one to aggregate resources based on various parameters.

MEASUREMENT OF WEB RESOURCES

We will next consider the services that are used for the quantitative assessment of web resources.

First, these are the services that allow one to assess the volume of web resources both in physical (e.g., in GB) and logical units (e.g., the number of pages, documents or files of a certain type that are available on a website). Various services exist that count the quantity of text resources that are expressed in words and graphics, including unique images and audiovisual resources (as measured in length).

A variety of thoroughly studied indicators that are associated with the use of online information resources exists. Such indicators are commonly used for commercial purposes: popular resources are more attractive to advertisers.

Finally, webometrics, which explores the visibility of websites and their connectivity based on website citation and hyperlinks, as well as altmetrics, which evaluates references to information items in social media, are two independent areas that have proliferated in recent years. These methods are discussed below.

When it comes to measuring the volume of resources on the Russian Internet, one of the most famous Russian internet companies, Yandex, has the priority in this area.

Yandex had been supporting a service called Numbers for several years, which calculates the following parameters:

Direct Values

- The number of unique servers, in pieces
- The number of unique documents, in pieces
- The amount of indexed information, GB
- Reciprocal Values
- The average size of one page (document), KB
- The average number of pages on one server, in pieces
- The average volume of one server, MB

This tool was later replaced by several new services that capture various statistics on the Russian Internet. In 2009, Yandex released an interesting report on the content statistics of the Russian Internet [18], which contained the following data:

- the total number of Russian websites;
- the number of pages on websites;
- the distribution of websites;
- the quantity of text and graphics, as well as audio and video information (NB: Yandex uses different units of measurement for different types of data);
- Internet language indicators (language, word count, spelling errors, and neologisms, etc.).

Currently, Yandex offers a number of services that enable users to identify relevant indicators of website use (Metrics), or indicators for the use of words in queries

(Wordstat), and to obtain the necessary data themselves.

WEBSITE TRAFFIC DATA

As noted above, services related to the use of information resources based on the analysis of website and webpage traffic data are actively developing on the Internet. These services are also referred to as web analytics.

Web analytics is applied for the following purposes:

- developing website functionality based on user-behavior patterns;
- evaluating the effectiveness of online advertising campaigns;
- identifying bottlenecks in website structure, navigation, and content.

Based on traffic statistics, the following can be measured:

- the number of webpages viewed;
- frequency of keywords applied by users to find a website through a search engine;
- locations of users;
- time spent on a webpage by a visitor;
- transfers between webpages;
- website audience (random or regular users, etc.).

Web analytics is a young industry not only in the CIS countries, but also worldwide. However, the Web Analytics Association has already launched common standards underpinning measurements and analysis in the field of web analytics.

Wikipedia [19] provides the following classification of web analytics tools:

Log Analysers:

WebTrends

Webalizer

AWStats.

Counters—ratings (The number of visitors per day, week, month, and for the entire history):

Rambler's Top100

Liveinternet

Rating@Mail.ru

OpenStat

HotLog.

Internet statistics systems (The aggregated data on visits calculated for a selected parameter set by a user):

Google Analytics

Piwik

Yandex.Metrica

Liveinternet

Rating@Mail.ru

OpenStat (ex. Spylog)

HotLog.

Internet Statistics Systems that Offer Detailed Information on Webpage Views (Information on webpage views during each visit is available in addition to the summary overview):

Woopra.

Web Analytic Systems that Offer Detailed Information about User Behavior on a Webpage (The maximum possible detail with the ability to monitor all user actions: mouse movements, clicks, keystrokes, etc.). Collected behavioral data are applied to create reports in the form of user webpage activity maps):

ClickTale

Yandex.Metrica

SpyBOX.

Citation indices (CIs) are one of the most popular tools for analysis and subsequent optimization of information resources applied for both websites and individual pages. Citation indices are generalized numerical indicators of link popularity (link citations). These indices are calculated by search engines and later used for result ranking algorithms. Generally, citation indices can serve as indicators of website "popularity," while information on individual ranks of webpages allows one to identify the strengths and weaknesses of a given information resource.

Some of the best-known citation indices are presented below.

Yandex tCI (CY) is a thematic citation index of the Yandex search engine calculated in relation to semantically similar resources.

Logarithmic Yandex tCI (LCY) is the *value* calculated based on the thematic citation index. The LCY scale can vary from 0 to 100, which currently corresponds to the *mCI* ranging from 10 to 250000. The LCY of the websites that have no thematic citation index is equal to −1 (© S. Kholod, 2007).

Google PageRank (PR) is an algorithm for calculating the reputation of webpages using the *Google* search engine. This algorithm applies a method for calculating the weight of a webpage by measuring the importance of references to this page. The PR scale can vary from 0 to 10. The PR of relatively known websites is equal to 4 to 5. The PR of highly visible websites is 6. Seven is a value that is almost unreachable for the majority of websites; however, it can still occur in some cases. Highly popular and important projects have the values 8, 9, and 10. The PR is a parameter that is related to each individual page, rather than an entire website. The same website can host webpages with different PR values. It is recommended to use the value – 1 for webpages that have not yet been assigned their PageRank.

Alexa Trafic Rank (ATR) defines the place of a domain on a list ordered by the frequency of domain visits. **Alexa Rich Rank** (ARR) defines the place of a domain on a list ordered by the accessibility of a domain to the public. The indicators of the *alexa* ranking system (a subsidiary of *Amazon*) are based on the calculation of visits and page views frequency. The traffic calculation algorithm of the Alexa Rank relies on the simple averaging of the number of page views on a particular website during 3 months.

Site Rank (SR) is a complex analog of a thematic citation index for the entire Internet, which was proposed by the *XAP* and *TNX* advertising systems and was recently released for beta testing. The SR is affected not only by backlinks, but also by the presence of indexed pages, thematic citation index, traffic, and Alexa Rank, etc. The developers expect that the Site Rank will become a standard. However, due to its multivalued complexity, one of the ranking's disadvantages is the long time that is necessary to obtain the results (from 5 to 120 seconds).

Several examples of statistical services are presented below.

Google Analytics [20] (GA) is a free service offered by *Google*, which allows one to collect detailed data on website traffic. Statistics are stored on the Google server, whereas users only place a JS code on their webpage.

The free version is limited to 10 million page views per month. Users with an active Google AdWords account can track an unlimited number of page views. One of the specific features of this tool is that the webmaster can optimize advertising and marketing campaigns via Google AdWords by analyzing the data on where visitors come from, how long they stay on a website, and where they are located geographically, which is obtained through the *Google Analytics* tool.

Users can see ad groups and the return on keywords in the reports. Additional features, including the division of users into groups, are also available. Service users can set goals and transition sequences. Such goals can be a display of a final sales page or specific pages, or file uploading. The use of this tool helps marketing experts to evaluate the success of advertising campaigns and to identify new sources of the target audience.

Google Analytics displays basic information "on a toolbar," while more detailed information can be provided in a report. At present, 80 types of customizable reports are available for use.

Yandex.Metrica [21] is a free service that is designed to evaluate *website* traffic, and analyze user behavior. The tool has been available since 2009.

The *Yandex.Metrica* counter operates based on the principle of a conventional visit counter: the webmaster installs the JS code; it collects data on each visit.

This tool is integrated with *Yandex.Direct* and Yandex.Market and allows one to group resource users by several parameters.

Yandex.Metrica measures the conversion of a website and online advertising. To calculate the conversion, the tool assesses the share of users that have reached the "goal," i.e.:

• reached the webpage; the visit can be regarded as the achievement of a result,

• checked a number of webpages, which represents the campaign success target,

• performed a necessary operation (clicked the button, downloaded the price list, etc.).

Up to 100 "goals" can be assigned for the Yandex.Metrica. The tool generates data for a current day. The reports are updated every 5 minutes. It is possible to monitor the accessibility of the website and to send SMS notifications when the website is not accessible to users.

A special Report Designer, which allows one to create detailed reports, is embedded into the tool.

To analyze the sequence of page and section views by users, *Yandex.Metrica* offers a Route Map for a website (the report presents the main directions of user navigation through the website in the form of a graph, whose vertices correspond to page or site sections, while the edges correspond to user routes).

Top 100 is a service designed by Rambler [22], which provides an aggregate "popularity index" for information resources (websites), selected on the basis of criteria set by users. At present, the tool considers the following characteristics:

- the number of hosts on the main page per day;
- the number of hosts on the entire website per day;
- the daily average number of views per host for the previous 7 days;
- the number of hosts on the main page for the previous 7 days;
- the number of hosts on the entire website in the previous 7 days;
- the average number of views of one page in the previous 7 days;
- the average number of views of three to five pages in the previous 7 days;
- the average number of visitors based in Russia in the previous 7 days.

The host is a unique IP address. The index is recalculated every hour. In addition, the service details the number of unique visitors that browsed the resource or its main page in a given period.

Google Scholar

The Google Scholar tool is worthy of a particular reference [23]. First, this service is specifically focused on scientific and educational resources that are available online. This service handles not only periodicals in the public domain, but also other categories of online publications, such as preprints placed in open archives. Furthermore, it allows one to calculate indicators that were used previously as part of the traditional bibliometric indices, such as publication activity and citation, as well as the h index. Thus, this tool makes an attempt to combine classical bibliometric indicators with the new types of scientific communications. This fact can be regarded as an undoubted advantage of the tool.

It should be noted that the above-cited new government regulations that underpin the assessment of academic institutions in the UK suggest using Google Scholar. However, the use of this tool is only allowed for certain branches of knowledge. Indeed, the reflection of scientific and educational resources by Google Scholar is more complete for a number of disciplines, for example, pedagogy, as compared with the traditional citation indices, such as the Web of Science, Scopus, or the Russian Index of Science Citation. However, the question of the kinds of algorithms that are applied for the selection of resources for indexing by this tool is not clear. Therefore, the results of indicator calculations can be questioned. In any case, this service represents one of the most adequate methods of calculation, at least for a number of indicators that are applied to measure scientific and educational resources.

Webometrics

We have written about webometrics and the creation of the Russian webometric index in the Institute of Scientific and Educational information of the Russian Academy of Education on several occasions [24–26]. Therefore, a detailed overview of this methodology is beyond the scope of this paper. One should only note that webometric indicators represent the only tool that is officially used to measure online scientific and educational resources, specifically, the websites of scientific and educational institutions. These indicators provide the basis for the Webometrics Ranking of World Universities. They are also used in most university rankings. However, one should recall that the classic webometric index is made of four indicators:

- the size of a website (the number of pages);
- the size of a website (the number of so-called "rich" files);
- the "visibility" of a website in the Google Academia search engine;
- the citation of a website (the number of references).

Lately, webometrics has been applied for other scientometric tasks. For example, A.A. Pechnikov and his colleagues from the Karelian Research Center of the Russian Academy of Sciences conduct research on webometric connectivity of scientific and educational websites in order to cluster the scientific and educational information space [27] and to construct a general model of the Russian university Web [28].

Altmetrics

Altmetrics (alternative metrics) is a new tool that is relatively unknown in Russia for the assessment of information resources, which specifically focuses on science and education. Altmetrics is described below in more detail based on [29].

We noted above that the use of bibliometric indicators based on citation, especially the impact factor, is perceived by many researchers as inadequate and highly dependent on commercial and market risks.

Against this background, and given the spread of social media in the scientific and educational environment, an alternative system has emerged for measuring scientific information resources, primarily papers and datasets. This system is called altmetrics (alternative metrics).

This service, which arrived only in 2013, quickly gained popularity among many academic institutions and publishers. It is an aggregator of the scientific content referenced in the media, social networks, blog posts, and other online sources. Essentially, data on a particular paper, rather than on the entire journal, are processed, which makes the assessment more accurate. This service allows one to build an aggregated estimate of a journal, research institution, or a research project. Altmetrics based estimates are suitable for drawing attention to specific scientific results that are not sufficiently reflected in traditional citation indexes, but attract attention online.

Altmetrics is a composite index, which is calculated as the sum of references in social media (Twitter, Facebook, etc.), as well as specialized scientific blogs, Google+ services, and Mendeleev, etc.

Altmetrics is particularly useful for publishers, who can appreciate attention to publications that appear in their journals. Open-access and open-archive publications have a great advantage from the viewpoint of altmetrics. The service can also be useful for the selection of the most effective online communication formats for dissemination of research results.

As of November 2013, the service had collected data on 1.7 million papers. This tool is oriented towards publishers, allowing them to select papers based on a variety of criteria, including time, identifier, ISSN, and journal title. Such data collection is fast and large scale. Every day the tool collects 15000 references to scientific results. Furthermore, the service captures 22000 unique articles on a weekly basis (data as of November 2013).

The further development of the service involves the creation of tools that enable the differentiation of indicators that characterize scientists, academic institutions, the results obtained by individual university laboratories, research groups, and inter-university projects.

The authors of [29] emphasize that altmetrics indicators differ significantly across different research areas and disciplines. For example, they are almost four times higher for biomedical sciences compared with social sciences. It still remains to clarify how this indicator can be applied for other subject areas.

CONCLUSIONS AND RECOMMENDATIONS

An analysis of the state of measuring scientific and educational resources shows that objective common methods do not exist as of this date.

First, one can state that public statistics indicators do not meet modern requirements of development management and use of information resources.

Bibliometrics methods based on the calculation of publication activity and citation, especially in the scientific literature, still play an important role for science administration. However, in the future, given the intensifying digitalization of scientific and educational communications, the role of online-based methods for the measurement and evaluation of scientific and educational resources is expected to grow.

At the same time, one can observe that different indicators such as volume, traffic rates, visibility, citation, and reference in social media are applied for the assessment and ranking of online information resources. All of these indicators apply a variety of tools, although the calculations of most indicators rely on the most common search engines such as Google (for services related to the assessment of global resources) and Yandex (for Russian services).

The central problem is associated with the need to establish the relationship between quantitative measures of information resources and their qualitative characteristics. This relationship is particularly important for scientific and educational information purposes. Indeed, the assessment of traffic and user behavior patterns of a resource can help to rather accurately estimate the commercial value of the resource. However, this is certainly not sufficient for research and educational purposes. The resource value depends on its novelty, completeness, and quality. These parameters are not always correlated with traffic. For example, the most frequently visited resources in the field of education are abstract databases, which cannot be regarded as educational resources of the best quality.

The accounting of scientific and educational resources is an independent but highly necessary task for the development of a database of their inherent characteristics. This task can only be performed based on a coherent and commonly accepted classification of resources, including standardized definitions of their types.

The overall conclusion is as follows: each studied method that is applied for obtaining quantitative data on information resources has its own advantages and an optimal scope of application. Therefore, the development of a common system of monitoring in this area requires the creation of a consolidated database in which the accounted scientific and educational resources should be described as a set of metadata that adequately reflects the specific characteristics of resources by features that are relevant for different evaluations.

Furthermore, each resource should have a set of indicators that are calculated by different methods: volume, web analytics, webometrics, and altmetrics, etc. The database must naturally be updated automatically based on the relevant online services. Clearly, Google tools have a high priority, but they should not necessarily be the exclusive ones.

The database should include the indicators that are currently calculated as part of the Russian webometric index of scientific and educational institutions. As already mentioned, this index is currently being developed at the Institute of Scientific and Pedagogical Information of the Russian Academy of Sciences.

If the proposed method is approved in principle by the scientific community and public authorities who are responsible for science and education, it will be possible to discuss a specific set of relevant indicators.

Such an integrated approach will allow one to build rankings of scientific and educational resources based on various sets of features. Such rankings can help to solve immediate practical problems that are associated with the selection of scientific and educational resources for acquisition by the National Digital Library and the National System for Storage of online resources.

APPENDICES

APPENDIX A. Indicators for information resources in science (based on data contained in the statistical compendia of the Higher School of Economics)

6. 1 Publications of Russian authors in scientific journals indexed in Scopus, by type of document

6.2. Publications of Russian authors in scientific journals indexed in the Web of Science, by type of document

6.3. Publications of Russian authors in scientific journals indexed in Scopus

6.4. Publications of Russian authors in scientific journals indexed in the Web of Science

6.5. Specific share of Russian authors in the global number of publications, including publications indexed in Scopus, by research area, in 2012

6.6. Structure of Russian publications in scientific journals indexed in the Web of Science, by research area, in 2008–2012

6.7. Submission of patent applications and patents granted for inventions

6.8. Indicators of patent activity

6.9. Russian Federation's patents for inventions by sections of the International Patent Classification

6.10. Distribution of Russian Federation's patents granted for inventions, by origin of applicants and by sections of the International Patent Classification in 2012

6.11. Submission of patent applications and patents granted for utility models

6.12. Russian Federation's patents for utility models, by sections of the International Patent Classification

6.13. Registration of intellectual property in the field of information

APPENDIX B. Indicators for library information resources (suggested by the author)

1. Digital library fund

1.1. The number of full-text electronic documents (titles)

1.2. In the total volume of the library fund, %

Including:

On portable media

On the local network

Available online

1.3. Digitized by oneself in the reporting year (titles)

1.4. Received externally, including electronic document delivery (titles)

....

3. Electronic catalog

3.1. The total number of records in the electronic catalog

3.2. In the total volume of the catalog, %

3.3. The number of entries entered by oneself in the reporting year

3.4. The number of entries received externally in the reporting year

3.5. Possibility of online access to the electronic catalog

4. Online services

4.1. Support of a website or a webpage

4.2. The number of distance requests for library e-services

4.3. The number of unique users of the website

4.4. Citation (in the Yandex catalog)

4.5. The number of Internet connections from the library building (sessions)

4.6. Total cost spent on the access to external databases (in thousands of rubles)

4.6. The number of documents scanned or uploaded (during the service session) from external databases

4.7. The number of information queries processed through the e-service, %

4.8. The number of documents provided via electronic delivery service (titles)

5. Local mode services

5.1. Electronic catalog queries (requests per year)

5.2. Electronic publications issued on portable media (pieces)

5.3. The number of documents delivered through the local e-library

APPENDIX C. List of digital library catalogs

1. http://yaca.yandex.ru/yca/cat/Culture/Literature/Online_Libraries. Yandex, abstracts, 118 links;

2. http://till.ru/library. Online Russian text resources compiled by Kuznetsov, abstracts, 282 + 126 digital libraries and websites;

3. http://allreferats.narod.ru/libraries.htm, http://allbest.ru/libraries.htm. 4558 links, abstracts;

4. http://ekzo.net/Cat/?razd=6. Abstracts, 433 links;

5. http://www.library.ru/2/catalogs/elibs. Russian State Library for the Youth, abstracts, 319 links;

6. http://window.edu.ru/catalog?p. Single window access to educational resources. Abstracts by categories.

Training materials, **15688**

Educational materials, 13 719

Reference materials, 928

Illustrative and demonstration materials, 52

Digital libraries, 480

Scientific materials, 4531

Educational websites, 8653

Regulatory documents, 1375.

7. http://www.kulichki.com/inkwell. Inkwell. A catalog of links to digital libraries, collections, and individual documents.

8. http://pro-spo.ru/biblio. A well-structured and annotated catalog; the volume is unclear.

9. http://www.nlr.ru/res/inv/ic_www/cat_show.php?ri. The online guide of the Russian Scientific Library with almost 200 Russian addresses, mostly of libraries.

10. http://hotuser.ru/biblio/1421-2009-05-07-07-28-33#anch01. The enriched catalog no. 8.

11. http://fo20gpntb.ya.ru/replies.xml?item_no=130. In total, 113 links to digital libraries.

12. http://www.liveinternet.ru/users/znayki/post27758148. Free digital library containing almost 600 Russian links.

13. http://www.ibpm.su/stud/biblio.php. Institute of Business and Law, nearly 120 links.

14. http://www.msu.ru/libraries. 16 digital libraries of the Moscow State University.

15. http://www.neumeka.ru/biblioteki_besplatnyh_knig.html. About 70 links.

16. http://hsscm.msu.ru/links/3. 43 links.

17. http://feb-web.ru/feb/feb/sites.htm. The catalog of links of the Fundamental Digital Library, 816 links for literature.

18. http://budichome.narod.ru/Biblio/biblio1.html. Free text and audio file digital library consisting of 15 items.

19. http://lib.uran.ru 12. The digital library of the Ural Branch of the Russian Academy of Sciences.

20. http://lib.uran.ru/libr.html. About 70 links.

21. http://www.hristianstvo.ru/internet/libraries. 601 digital library collections on orthodoxy.

22. http://4pda.ru/forum/index.php?showtopic=462&st=20. A structured catalog of links; the volume is unclear.

23. http://nedorazvmenie.livejournal.com/869551.html. 417 links to free digital libraries.

24. http://lib.kantiana.ru/cgi-bin/irbis64r_81/cgiirbis_64.exe. The catalog of online books.

25. http://lib.kantiana.ru/jirbis/index.php?option=com_weblinks&Itemid=1675. Baltic Federal University, 100 links, the volume is unclear.

26. http://library.gu-unpk.ru/bibl_INT.php. About 100 links.

27. http://supercook.ru/spr-01-library.html. About 450 links.

28. http://old.gnpbu.ru/web_resyrs/Katalog.htm. National Pedagogical Library n.a. K.D. Ushinskii, 1140 links to educational resources.

29. http://guide.aonb.ru/library.html. About 130 links.

30. http://old.russ.ru/krug/biblio/catalogue.html. About 150 links for 2005.

31. http://igorladov.com/katalog.htm. About 50 links to digital libraries and a directory of resources on philosophy and religion, etc.

32. http://www.orc.ru/~patrikey/liblib/liblib.htm. A digital library catalog, by section.

33. http://search.list.mail.ru. 475 links provided for a digital library query.

34. http://top100.rambler.ru/navi/?type=10. 2281 links provided for a digital library query.

35.http://www.dmoz.org/World/Russian/@Источники_информации/Библиотеки/Электронные/ Approx. 80 links.

36. http://www.kinder.ru/kinder.asp?r. 186 websites on literature.

37. http://www.promotion.su/search?q=электронные+библиотеки. 362 links.

38. http://hotlinks.ru/catalog/120/108. 313 web libraries.

39. http://ivan.susanin.com/thema.phtml?them. 215 libraries, literature.

40. http://intersib.ab.ru/ 126 links on Siberia and literature.

NB: Directories of the portals such as Russian Education and the Ushinskii National Pedagogical Library, as well as several others, including not only digital libraries, but also other types of educational resources.

REFERENCES

1. Antopol'skii, A.B., The use of information resources for the performance assessment of scientific research, *Mezhotr. Inf. Sluzhba*, 2011, no. 1, pp. 40–53.
2. Bredihin, S.V. and Kuznetsov, A.Yu., *Bibliometric Methods and the Electronic Scientific Periodicals Market,* Novosibirsk: IVMiMG SO RAN, NEIKON, 2012.
3. Reorganization of the Russian Academy of Sciences 2013. http://www.saveras.ru
4. Antopol'skii, A.B., Open Data in Science and Education, in *Khudozhestvennoe obrazovanie v informatsionnom obschestve* (Arts Education in the Information Society), Ìoscow: MGUKI, 2014, pp. 115–215.

5. The Strategy for the development of the information society in the Russian Federation. http://www.rg.ru/2008/02/16/informacia-strategia-dok.html
6. Science and Technology. Innovations. Information Society: Pocket Data Book, Moscow: Nats. Univ. Vys. Shkola Ec., 2012.
7. *Science and Technology Indicators 2014: Data Book*, Moscow: Nats. Univ. Vys. Shkola Ec., 2014.
8. Open government. http://open.gov.ru/opengov
9. The order of the Ministry of Education and Science of the Russian Federation no. 162 as of March 5, 2014. http://www.rg.ru/2014/05/14/minobrnauki2-dok.html
10. Antopol'skii, A.B., Gorushkina, S.N., and Shlykova, O.V., The state of digital culture in some regions of Russia: the results of complex study conducted in 2010, *Sprav. Rukov. Uchrezhd. Kult.*, 2011, no. 5, pp. 6–21.
11. Antopol'skii, A.B., *Russian Information Resources: Scientific Tutorial*, Moscow: Liberia, 2004.
12. The list of information resources parameters. http://elearn.oknemuan.ru/?p=6&id=117
13. Research Excellence Framework. http://www.ref.ac.uk
14. The map of the Russian Science. http://mapofscience.ru
15. Science maps. http://onr-russia.ru/content/карта-российской-науки
16. National Digital Library. http://mers.medart.tomsk.ru/news/194.html
17. The Project "Foundations of the State Culture Policy". http://www.rg.ru/2014/05/15/osnovi-dok.html
18. The report on the content statistics of the Russian Internet. http://company.yandex.ru/researches/reports/ya_content_09.xml
19. Wikipedia. http://ru.wikipedia.org/wiki/@Основные_показатели_посещаемости_сайта
20. Google Analytics. http://www.google.com/analytics/
21. "Yandex.Metrica". https://metrika.yandex.ru/list/
22. "Top 100". http://top100.rambler.ru/
23. Google Scholar. http://scholar.google.ru/scholar?a
24. Antipol'skii, A.B., Polyak, Yu.E., and Usanov, V.E., The Russian index of websites of scientific and educational organizations, *Inf. Res. Rossii*, 2012, no. 4, pp. 2–7.
25. Antopol'skii, A.B. and Polyak, Yu.E., The new website ranking of Russian universities, *Dist. i Virt. Obuch.*, 2013, no. 1, pp. 40–47.
26. Antopolskii, A. B. On the feasibility of the Russian national webometric index, *Sci. Tech. Inform. Process.*, 2014, vol. 41, no. 1, pp. 33–37.
27. Pechnikov, A.A., Lugovaya, N.B., and Chuyko, Yu.V., Links between numerous websites of RAS, *Probl. Contemp. Sci. Pract.*, 2009, no. 26, pp. 47–52.
28. Pechnikov, A.A., The university Web model, *Vestnik Nizhegorod. Univ. im. N. I. Lobachevskogo*, 2010, no. 6, pp. 208–214.
29. Liu, J. and Adie, E., Realising the potential of altmetrics within institutions, *Ariadne*, 2014, no. 72. http://www.ariadne.ac.uk/issue72/liu-adie

Translated by V. Kupriyanova-Ashina