ISSN 0147-6882, Scientific and Technical Information Processing, 2017, Vol. 44, No. 4, pp. 285–290. © Allerton Press, Inc., 2017.
Original Russian Text © L.M. Koroleva, E.V. Koltunova, 2017, published in Nauchno-Tekhnicheskaya Informatsiya, Seriya 1: Organizatsiya i Metodika Informatsionnoi Raboty, 2017, No. 11, pp. 43–49.

The Presentation of Chemical Information in Abstract Databases

L. M. Koroleva* and E. V. Koltunova**

All-Russia Institute of Scientific and Technical Information, Russian Academy of Sciences, Moscow, 125190 Russia

**e-mail: lkorol@viniti.ru*

***e-mail: chemistry711@viniti.ru*

Received August 29, 2017

Abstract—We discussed several issues related to the systematization of scientific publications in the field of chemistry and chemical technology that are contained in abstract databases such as the Web of Science Core Collection, Scopus, and the VINITI RAS Chemistry Database. Next we noted a rapid growth of interdisciplinary publications authored by chemists. We discussed the problems of information support in view of the intensifying interdisciplinary and multidisciplinary nature of research. Then we presented the results of an information search performed across several abstract databases for the main subject areas. In the end, we discussed an increase in the number of scientific publications in the field of chemistry and chemical technology, which has accelerated in recent years in comparison to other fields of science.

Keywords: scientific publication, chemical information, abstract databases, publication activity, information support, subject area, Web of Science Core Collection, Scopus, VINITI RAS Chemistry Database

DOI: 10.3103/S0147688217040104

Since 1901, the achievements in chemistry, as well as in two other research disciplines, have been distinguished with the highest award. For 117 years, the Swedish Royal Academy of Sciences and the Nobel Committees on Physics and Chemistry have awarded the Nobel Prize to outstanding scientists who "shall have made the most important discovery or improvement in the field of chemistry ...", as stated in Alfred Nobel's will written on November 27, 1895 in Paris. The outstanding Russian chemist Nikolai Semenov is one of the Chemistry Nobel Prize laureates.

This annual award is the most respected award in the world scientific community. For over a century it has been emphasizing the value of chemistry for society. Therefore, one cannot overestimate the importance of accumulating and processing chemical knowledge.

The role of interdisciplinary and multidisciplinary researches in science is growing. Interdisciplinary and multidisciplinary studies lie at the intersection of sciences and require the involvement of experts from different areas. Such types of research have been actively developing since the 1980s. The rapid growth of interdisciplinary projects led to the development of new independent scientific areas. In particular, several independent multidisciplinary research areas, such as computer chemistry, nanotechnology, chemoinformatics, biophysics, bioinformatics and biomedicine, have emerged in physical Sciences.

The importance and immanency of interdisciplinary and multidisciplinary approach in science was succinctly underlined by the 2006 Nobel Prize laureate in Chemistry, the American biochemist Roger Kornberg, who famously noted: "The development of modern drugs requires knowledge of quantum theory." Zhores Alferov, the Russian Nobel Prize laureate, repeatedly highlighted the growing role of interdisciplinary research and emphasized the importance of fundamentally new interdisciplinary approaches in education.

Both the institutes of the Russian Academy of Sciences and the top Russian universities such as the Moscow State University, National Research University MIPT (Moscow Institute of Physics and Technology), and St. Petersburg State University are engaged in interdisciplinary research. The *Human Genome Project (HGP)*, which was launched in the 1990s, is one of the best-known interdisciplinary projects; it brought together biologists, chemists, mathematicians, and physicists.

On the one hand, such studies require the appropriate interdisciplinary information support. On the other hand, they often result in scientific publications that are interdisciplinary in nature. While dealing with information in the field of chemistry and material sciences, a rapid growth of interdisciplinary publications authored by chemists can be observed. As a result, there is a need for a better systematization of chemical information.

Abstract, bibliographic, full-text, and structural information in the field of chemistry and material sciences is accumulated by and stored in various data-

Table 1. Subject clusters in the Web of Science databases

Cluster	The number of subject areas (research areas)
Life Sciences and Biomedicine	75
Physical Sciences	17
Technology	21
Arts & Humanities	14
Social Sciences	24

Table 2. Subject clusters in the Scopus database

Cluster	The number of subject areas (including areas of knowledge)
Health Sciences	103
Life Sciences	52
Physical Sciences	116
Social Sciences & Humanities	66

bases, particularly, *Chemical Abstract Service (CAS)*, which is the oldest global abstract service in the field of chemical information. Today, the world's biggest players in the information market, such as *Clarivate Analytics* (which engages more than 4000 scientists and experts in more than 100 countries) and *Elsevier* (which engages more than 7200 scientists and experts in 24 countries), have developed and maintain large information clusters in various fields, including chemical information. Presentation, classification, systematization, and indexation of chemical information underpins the effectiveness of information support for chemists and chemical experts. Thus, the polythematic abstract, bibliographic and scientometric database *Web of Science Core Collection (WoSCC)*, which operates on the Clarivate Analytics' Web of Science platform, provides five global clusters for the purpose of knowledge systematization. Three clusters are of relevance to our study: *Life Sciences and Biomedicine*, *Physical Sciences,* and *Technology*. Each global cluster is divided into subject areas, which are referred to as *Research Areas* in the *Web of Science*. The number of

Table 3. The coverage of the Chemistry subject area in the WoSCC database (fragment)

The Web of Science field: Research areas	Number of records	%
CHEMISTRY in total	918500	100.000
including:		
Materials science	167739	18.262
Physics	126966	13.823
Science technology other topics	89247	9.717
Biochemistry Molecular biology	63867	6.953
Engineering	57304	6.239
Electrochemistry	45009	4.900
Energy Fuels	39207	4.269
Pharmacology Pharmacy	34615	3.769
Food Science Technology	32447	3.533
Polymer Science	19809	2.157
Instruments Instrumentation	18214	1.983
Crystallography	18147	1.976
Metallurgy Metallurgical Engineering	17341	1.888
Spectroscopy	13709	1.493
Thermodynamics	12990	1.414
Biophysics	11250	1.225
Agriculture	10050	1.094
Nutrition Dietetics	8855	0.964
Biotechnology Applied Microbiology	8787	0.957
Nuclear Science Technology	7651	0.833
Environmental Sciences Ecology	7491	0.816
Computer Science	5399	0.588
Toxicology	3776	0.411
Mathematics	3462	0.377

Table 4. The coverage of the Chemistry subject area in the Scopus database (fragment)

Scopus fields: Subject area	Number of records	%
CHEMISTRY in total	1090185	100.000
including:		
Chemical Engineering	332867	30.533
Materials Science	327834	30.071
Physics and Astronomy	248946	22.835
Biochemistry, Genetics and Molecular Biology	207586	19.041
Engineering	138539	12.708
Medicine	72501	6.650
Energy	66081	6.061
Pharmacology, Toxicology and Pharmaceutics	64573	5.923
Environmental Science	54047	4.958
Agricultural and Biological Sciences	34920	3.203
Computer Science	26633	2.443
Mathematics	20611	1.891
Social Sciences	8216	0.754
Earth and Planetary Sciences	8179	0.750
Immunology and Microbiology	3602	0.330
Business, Management and Accounting	2527	0.232
Health Professions	1949	0.179
Arts and Humanities	955	0.088
Economics, Econometrics and Finance	194	0.018
Dentistry	159	0.015
Psychology	104	0.010
Neuroscience	53	0.005
Nursing	25	0.002
Veterinary	17	0.002

subject areas (research areas) in subject clusters is presented in Table 1. All databases hosted on the *Web of Science* platform support information searches according to these subject (research) areas.

The *Scopus (Elsevier)* scientific information aggregator is a competitor of the *Web of Science* databases. This largest abstract and scientific citation database was founded in 2004. Knowledge is systematized in Scopus as follows. Four global subject clusters (*SubjectArea*) are used in this database to index publications. Chemical information is essentially covered by three clusters: *Health Sciences*, *Life Sciences*, and *Physical Sciences*. Global subject clusters are divided into 27 *major subject areas*, formed on the basis of ASJC codes (*All Science Journal Classification*), and 300+ so-called minor subject areas. The information search in the *Scopus* database is organized according to 27 major subject sections, which are referred to as *the most frequent Subject Area categories* in Scopus. The number of subject areas in the subject clusters is presented in Table 2.

The *Advanced Search* was performed for the *Chemistry* subject area covering the period between 2013 and 2017 in the *WoSCC* and *Scopus* databases. The retrieved documents classified by research areas (*WoSCC*) and areas of knowledge (*Scopus*) are presented in Tables 3 and 4.

This data shows that more than half of the documents in the Chemistry subject area are assigned to

Table 5. The ranking of countries by the number of documents reflected in the Chemistry subject area in the WoSCC database (fragment)

Field: Countries/territories	Number of records	%
In total records	918500	100.000
including:		
Peoples R China	248234	27.026
USA	170975	18.615
India	61228	6.666
Germany	58487	6.368
Japan	53308	5.804
South Korea	39715	4.324
France	39078	4.255
England	34463	3.752
Spain	33619	3.660
Russia	29267	3.186
Italy	28338	3.085
Iran	24604	2.679
Canada	21160	2.304
Poland	19563	2.130
Australia	18551	2.020
Brazil	17183	1.871
Taiwan	14569	1.586
Switzerland	13366	1.455
Netherlands	11235	1.223
Saudi Arabia	11008	1.198
Turkey	10985	1.196
Sweden	10089	1.098
Singapore	9954	1.084
Belgium	9321	1.015
Czech Republic	9115	0.992

Table 6. The ranking of countries by the number of documents reflected in the Chemistry subject area in the Scopus database (fragment)

Field: Country	Number of records	%
In total records	1090185	100.000
including:		
China	296071	27.158
United States	172187	15.794
India	75240	6.902
Germany	72484	6.649
Japan	63971	5.868
United Kingdom	47128	4.323
France	46416	4.258
South Korea	45609	4.184
Russian Federation	40205	3.688
Undefined	38508	3.532
Spain	37199	3.412
Italy	32423	2.974
Iran	29508	2.707
Canada	27205	2.495
Australia	23293	2.137
Brazil	22609	2.074
Poland	22264	2.042
Taiwan	16809	1.542
Switzerland	15491	1.421
Saudi Arabia	13847	1.270
Turkey	13502	1.239
Netherlands	13430	1.232
Sweden	11589	1.063
Egypt	11553	1.060
Belgium	11210	1.028

applied sciences such as biophysics, molecular biology and computers sciences, etc.

The comparison of the Research Areas and Knowledge Area categories that are used in the two databases shows that although these databases have some similarities and even overlap on several subject categories, they still differ in their approach to the classification of information in the field of natural sciences.

Furthermore, we analysed the coverage of publication activity pursued by scientists from different countries in the Chemistry subject field. Tables 5 and 6 present the results of the search for the period between 2013 and 2017.

The comparison of the samples obtained for the Country/Territory field from the two databases shows a good match of the results, particularly for the top five countries on the lists.

The analysis of changes in the publication activity of scientists from different countries for the period from 2008 to 2016 (compiled based on the *Scopus* data) (Fig. 1) shows that China firmly holds the position of the top global scientific leader in the field of chemistry

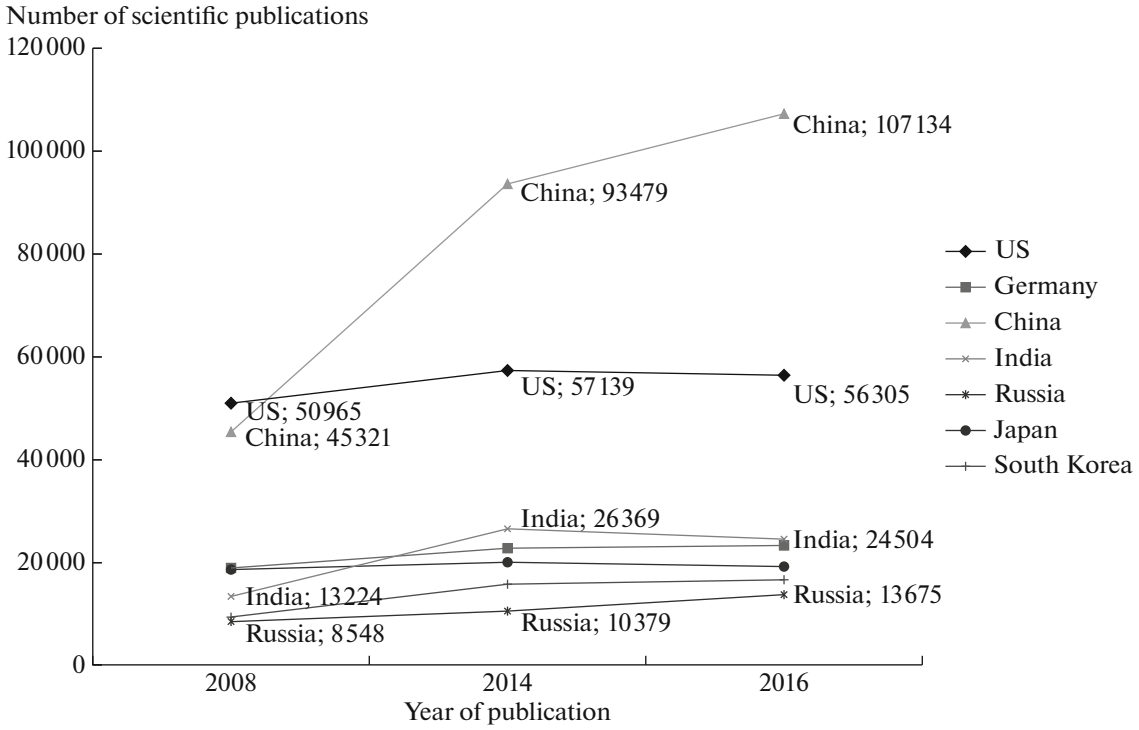

Fig.1. Dynamic publication activities of scientists from selected countries in the Chemistry and Chemical Technology subject area (based on *Scopus* data)

and chemical technology and has significantly overtaken the United States.

It should be noted that English, which is the main language of scientific communications, has started to lose ground due to a growing spread of scientific publications in Chinese. This trend is also fuelled by the development of specialized machine translation systems.

According to *Scopus* data, the share of scientific publications in the field of chemistry and chemical technology in the overall flow of scientific publications is 11.8%.

In recent years, there has been a steady increase in the number of scientific publications in the field of chemistry and chemical technology, which grew by 56.3% from 216406 to 338311 publications between 2008 and 2016 (Fig. 2).

The subject area Chemistry and Chemical Technology is in third place after Energy (90.2%) and Environment (67.4%) among all physical and technical Sciences in terms of the growth rates of scientific publications in the studied period. In terms of the total number of publications in 2016, the subject area Chemistry and Chemical Technology moved to fourth place, overtaking Biochemistry, Genetics and Molecular Biology, and Physics and Astronomy.

VINITI RAS has been issuing the national abstract database in the field of natural, exact, and technical sciences in print since 1953 and both in print and electronically since 1981. The Russian National Information Center VINITI RAS has approximately 65 years of experience in searching, collection, processing, use and dissemination of data. It has made significant achievements in the field of systematization and classification of scientific and technical information. The VINITI classifier, which was developed by detailing the Russian State STI Rubricator (GRNTI), supports the systematization of scientific and technical literature to the ninth level. The subject areas are defined in accordance with GRNTI. The VINITI Classifier is a set of classifiers applied for various knowledge areas, which describe the subjects of scientific and technical literature reflected in the VINITI database and in the Abstract Journal of VINITI. The database includes 30 thematic fragments (subject areas), which consist of 217 sections. The Chemistry thematic fragment currently includes 19 subject sections, which correspond to the main areas of chemical science. Each subject section consists of third-level subject headings. The total number of subject headings is 97. The share of chemical information in the VINITI's scientific and technical literature flow in 2016 reached 25.8% of the total number of processed documents. The Abstract database on chemistry and chemical technology, which is the largest thematic fragment of the VINITI's polythematic database, supports the systematization of chemical information in accordance with the principle of modern differentiation of sciences.

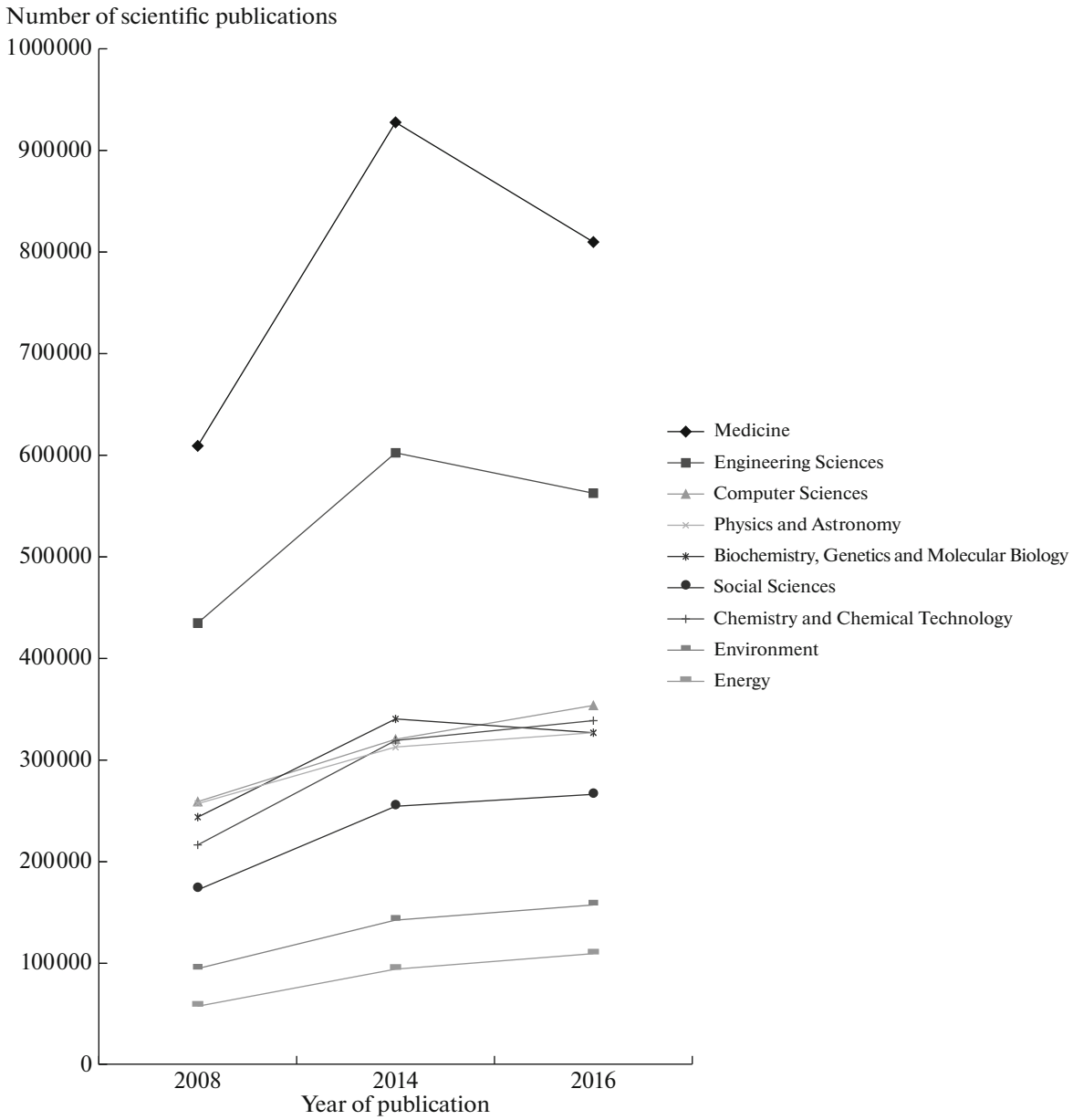

Fig. 2. The dynamic changes in the number of scientific publications published in various subject areas in the period between 2008 and 2016 (based on *Scopus* data)

REFERENCES

1. Web of Science Core Collection. https://clarivate.com/products/web-of-science/web-science-form/web-science-core-collection/.
2. Scopus. http://www.elsevierscience.ru/products/scopus/.
3. Database of VINITI RAS. http://bd.viniti.ru/index.php?option=com_content&task=view&id=238&Itemid=101.
4. VINITI Rubricator (current version, 2018). http://scs.viniti.ru/rubtree/main.aspx?tree=RV.
5. *Referativnyi zhurnal "19.Khimiya." Svodnyi tom* (19.Khimiya Abstract Journal. Cumulated Volume), Moscow: VINITI RAN. 1953, pp. 1953–2017.

Translated by V. Kupriyanova-Ashina