International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 9, Number 22 (2014) pp. 15757-15776

http://www.ripublication.com

# Construction of the Publication and Patent Clusters Produced by the Arbitrary Terms with the Use of the Specialized Google Tools

**Vladimir M. Moskovkin[1], Sergey I. Chernyshev[2], Maria V. Moskovkina[2], Ruslan V. Lesovik[3], Konstantin I. Logachev[3], Vadim V. Shaptala[3]**

[1]Belgorod State University, Russia, 308015, Belgorod, Pobedy str., 85
Email: moskovkin@bsu.edu.ru
[2]Research and Technological Institution of Transcription, Translation and Replication , PJSC, Ukraine, 61166, Kharkov, Kolomenskaya Str., 3, P.O.B 352
[3]Belgorod State Technological University, Russia, 308011, Belgorod, Kostyukov str., 46

## Abstract

There has been developed the analytical technique of construction of the publication and patent clusters produced by the arbitrary terms with the use of the specialized Google tools. On the basis of indicators of the total amount of cluster articles and the total citation TOP-10 % of the most cited articles the integral metrics of the research papers cluster was constructed by means of the additive and multiplicative methods. Different names of types of the computer calculations and devices were selected as the scientific terms for testing with the use of Google Scholar, Google Books and Google Patents beginning with the words: Quantum, Bacterial, Cognitive, Cellular, Cloud, Ubiquitous.

Based on the results of experiments with these terms it was shown that the most powerful research papers and patent clusters were correlated with the terms beginning with the words quantum and ubiquitous. It was shown that the patent capacity of the consolidated research cluster produced by the terms beginning with the word Cellular exceeded the patent capacity of the other consolidated clusters of the primary research studies by 2-3 orders.

It was shown that the publication and patent clusters produced by the terms “quantum computing” and “cloud computing” appeared late 80’s of the 20th century.

The significant temporal variability of the links to the same most-cited works by the use of Google Scholar was observed.

This point to the fact that Google Scholar shall be very carefully used for citation analysis that is confirmed by the findings of the other studies as well.

**Keywords:** publication clusters, patent clusters, Google Scholar, Google Books, Google Patents, Quantum Computing, Bacterial Computing, Cognitive Computing, Cellular Computing, Cloud Computing, Ubiquitous Computing.

## INTRODUCTION

As a rule, the “publication cluster” and “patent cluster” terms are considered in terms of the co-citation analysis [1, 2], i.e., as a combination of publications and patents interrelated by the cross-references.

Within this study according to the research [3] we will mean by such clusters the combinations of publications and patents associated with any scientific terms. Whether an article, book or patent belongs to the relevant cluster produced by an arbitrary term will be determined by the occurrence of this term (at least once) in the article, book or paten under consideration. In order to construct such clusters within limited or unlimited periods of time it is very convenient to use the specialized Google tools.

The tools under consideration: Google Scholar, Google Books and Google Patents.

Google Scholar launched in November, 2004 is now widely used for scientometric analysis including comparative analysis of responses to queries on the various scientific terms [3-9]. There was found a small number of studies on the use of the standard Google Books (launched in December, 2004) options in such scientometric analysis [10, 11]. Thus, it is noted in the study [10] that nowadays there are no comprehensive research performed that would directly use the Google Books and Google Scholar by analyzing the books citedness in the various fields of knowledge. And it is well known that the vast majority of citations in the social sciences and art and humanity falls on the books. Within this research there was conducted the comparative analysis of citations based on the 1000 books represented within the frameworks of the 2008 UK Research Assessment Exercise with the use of Google Books, Google Scholar and Scopus. Google Patents launched in December, 2006 is almost not used for analysis of the patent clusters produced by any terms. There is a number of researches on the patentometrics analysis with the use of this tool for study of the patent activities of the leading universities and transnational companies [12, 13].

However, the situation relating to the comparative analysis of the responses to the queries on different terms with the use of the above-mentioned analytical-search tools may change rapidly. This situation may be easily monitored through testing the name of these tools in the search line “with exact phrase” Google Scholar.

The analysis of these responses allows to understand to which extent they represent the articles that are of interest to us. The above publications and some conclusions drawn from the analysis thereof were obtained as the result of the above-mentioned analysis of the responses to the queries “Google Scholar”, “Google Books” and “Google Patents” in the search engine Google Scholar.

Thus, if we test the names of these tools in the above-mentioned order in the line “with exact phrase” and “in the title of the article” then we will obtain 254, 91 and 1 relevant responses, respectively (26.10.2012). In this experiment the “at least summaries” option was used and all the 7 subject areas were tagged in order to reduce the information noise and improve the response relevance (in the new version of the Google Scholar web-site the “at least summaries” and “subject areas” options are absent).

By searching “anywhere in the article” a great number of the irrelevant responses to the query “Google Books” are received from the web-site book.google.com, and for “Google Patents” - from perflavory.com and thegoodscentscompany.com.

Within the specified experiments we meant by the relevant responses those publications describing the peculiarities of the use and operation of the above-mentioned Google tools. However, the number of publications on the frequency comparative analysis of the various scientific terms amongst such responses was quite small.

It should be mentioned that the research papers cluster identifiable through Google Scholar is somewhat conditional since, besides the articles, the algorithm of this search engine may return the books, thesis works, preprints and other documents. Unfortunately, besides exclusion from the patent search results, Google Scholar does not feature any options for excluding the other documents different from the research papers.

This research is generally aimed at improvement of the model of construction of the publication and patent clusters produced by arbitrary terms with the use of the specialized Google tools, introduction of the qualitative metrics on these clusters for their comparison, identification of their features (the most-cited and the earliest indexed works within the publication cluster) and performance of the frequency analysis of their features (occurrence rate among the names of organizations, journals and other kinds of publications), analysis of the temporal dynamics of publications and patents falling within the clusters constructed. Description of the approaches and methods used for solution of the specified tasks is presented in the section below.

**RESEARCH METHODS**

In this research we will use the notions of the publication and patent clusters produced by any of the scientific terms [3].

We will mean by the publication or patent cluster produced by any scientific term a set of publications or patents including at least once this term within a certain period of time or at a given moment.

In order to study the specified clusters we suggest using the specialized Google tools: Google Scholar, Google Books and Google Patents [3].

We will identify the clusters of publications (articles and books) and patents (applications, patents issued) produced by the particular terms by testing the scientific terms within the wide-range time periods as well as on the annual basis in order to investigate the dynamics of these clusters. We will understand by the wide-range time periods the unlimited intervals as well as 25-year interval divided by the five 5-year intervals corresponding to the beginning of the publication clusters origination (1986 – 1990) and ending with the last 5-year interval (2006 – 2010). Breakdown by the 5-year intervals is the standard procedure by performance of the empiric-statistical investigations.

By testing the terms with the use of the Google Scholar search engine one may enter the search conditions with “with exact phrase” throughout the entire document or in its title using the “include citation” and “at least summaries” options. In order to reduce the information noise and exclude the patents from the responses one should tag all the 7 subject areas in Advanced Scholar Search (as was mentioned above, in the new version of Google Scholar the subject area search is not available and the option of inclusion (exclusion) of patents by the search is introduced). In Google Scholar the option of the books exclusion from the responses is absent but the number of books is usually not large.

Google Books also enables advanced searching with the exact phrase. In contrast to Google Scholar, it does not feature the options “cited by”, “citation” and “subject areas”. Along with the books Google Books indexes the entire magazine volumes. It is possible to search with the exact phrase throughout the entire book and in the book titles. In the latter case the tested term shall be quoted.

The term testing with the use of Google Patents may be performed in two ways:

- advanced searching with the exact phrase for Applications and Issued patents;
- advanced searching for Applications and Issued patents in the line «patent titles».

In the first case the search is performed throughout the entire patent description, in the second one – in its title. To perform searching in the patent title with the exact phrase the term shall be quoted.

It is important to emphasize that in contrast to Google Scholar the Google Patents provides a number of relevant responses that is different from their total amount. In the works [3, 12] it was shown that the smaller time intervals are chosen for the specified time period the better accuracy is achieved by calculation of the aggregate responses within the entire period specified.

The above-mentioned clusters of the research results shall be interpreted both generally and properly. In the first case it is assumed that the specified term is used in the text of the scientific publications and patents, in the second one – in their titles [3].

It follows from the above said that the research results cluster is divided into three separate clusters: research paper cluster, research book cluster and patent cluster. We will mean by the primary research results cluster the whole of the research papers and patents. Regarding the primary research results cluster, we can calculate the suggested factor of research cluster patent capacity which equals to the ratio between the number of patents issued and the numbers of the research papers [3], by doing so the number of the research papers is determined by means of Google Scholar with the use of the “at least summaries” option (in the new version of Google Scholar it corresponds to deactivation of the “include citation” option).

In order to evaluate the quality of the research papers cluster we suggest considering the (N, C) vector where N is the number of research papers in the cluster, C is the aggregate citation according to the TOP-10% of the most-cited papers.

On the basis of this vector one can construct the integrated indicator (metrics) of the research papers cluster quality using the additive or multiplicative method:

$$I_1 = \gamma_1 \overline{N} + \gamma_2 \overline{C} \quad (1), \qquad I_2 = \overline{N}^{\gamma_1} \overline{C}^{\gamma_2}, \quad (2)$$

where $\gamma_1, \gamma_2$ are the weighted coefficients, $\gamma_2 > \gamma_1$, $\gamma_1 + \gamma_2 = 1$, $\overline{N}, \overline{C}$ - standardized variables N and C (the standardization operation is performed by dividing N and C by the maximum values based on the sample of the investigated clusters).

Comparison of different clusters on the basis of the above formulas is basically relevant to the clusters that originated at the same time. If this is not the case one can construct the publication clusters within the frameworks of the same time intervals or construct the time cluster samples, i.e., test different terms at the same time and then calculate the total citation rate according to the TOP-10% of the most-cited research papers.

In each research papers cluster the Google Scholar allows identifying the most-cited («cited by» option) and the earliest papers indexed (by means of the varying time search interval). The most-cited papers are to be found in the beginning of the returned list of documents and identification of titles of the earliest papers indexed requires visual inspection thereof since poorly structured meta data may cause occurrence of errors [14].

By testing different scientific terms one may also analyze the names of the organizations the article writers are employed at as well as titles of journals and other academic documents they were published in. Within this research we will perform such analysis for the first ten responses.

We selected different names of the kinds of computer machines and computer calculations as the scientific terms for testing by means of the specialized Google tools, such terms begin with the words: Quantum, Bacterial, Cognitive, Cellular, Cloud, Ubiquitous. The selection of such terms is related to the incipient and rapidly developing research fronts in the computer science areas. Regarding the total amount of the documents analyzed, for example for the research papers cluster, it is determined by means of testing the selected computer terms with the use of Google Scholar within an unlimited time interval. But if we want to evaluate a share of documents the titles of which contain a particular special term in the amount of the documents the titles of which contain a more comprehensive term it can be easily done by means of the same search engine. The examples of such evaluation will be presented in the next section.

## RESULTS AND DISCUSSION

The results of testing the selected computer terms with the use of Google Scholar within the unlimited time intervals are presented in the Table 1. We can see from this table that the most powerful research papers clusters considered generally (search throughout the paper, at least summaries) involve the terms: “ubiquitous computing”, “quantum computing”, “cloud computing”, “quantum computer”, “quantum computers”. Within the frameworks of the present experiment it makes sense to speak of the consolidated clusters including three terms, for example, “quantum computer”, “quantum computers”, “quantum computing”. The capacity of these clusters is indicated in the Total line. The most powerful among them are the clusters beginning with the words “quantum”, “ubiquitous” and “cloud”.

The share of the documents the titles of which contain the “quantum computing” term in the total amount of the documents the titles of which contain a more generalized term “computing” equals to: 2490 / 21400=0,012. For the terms “cloud computing” and “computing” this ratio makes: 4070 / 214000=0,019. Thus, the estimates are about 1 – 2 %. The numerators for these calculations were taken from the Table 1 and the denominators – from the Google Scholar requests as of 10.10.2013.

The results of testing the selected computer terms with the use of Google Patents are presented in the Table 2. We can see from this table that the most powerful clusters of the patents issued considered generally (searching throughout the entire patent description) involve the terms: “ubiquitous computing”, “quantum computing”, “quantum computer”, “quantum computers”. The capacity of these clusters is indicated in the Total line. The most powerful among them are the clusters beginning with the words “quantum” and “ubiquitous”.

Table 1. Testing the selected computer terms with the use of Google Scholar, 22.11.2011

| **Name of the term** | **Total amount of responses** | | | |
|---|---|---|---|---|
| | **Searching throughout the paper** | | **Searching in the title** | |
| | **Including** | **At least** | **Including** | **At least** |
| Quantum computer | 16500 | 16400 | 1190 | 872 |
| Quantum | 14500 | 13700 | 610 | 431 |
| Quantum | 21900 | 21700 | 2490 | 1720 |
| Total | 52900 | 51800 | 4290 | 3023 |
| Bacterial computer | 32 | 31 | 3 | 2 |
| Bacterial | 15 | 12 | 1 | 1 |
| Bacterial | 21 | 19 | 5 | 4 |
| Total | 68 | 62 | 9 | 7 |
| Cognitive | 1030 | 941 | 63 | 41 |
| Cognitive | 80 | 76 | 5 | 4 |
| Cognitive | 853 | 792 | 65 | 45 |
| Total | 1963 | 1809 | 133 | 90 |
| Cellular computer | 640 | 580 | 42 | 19 |
| Cellular computers | 217 | 184 | 16 | 9 |
| Cellular computing | 782 | 629 | 66 | 34 |
| Total | 1639 | 1393 | 124 | 62 |
| Cloud computer | 539 | 512 | 38 | 24 |
| Cloud computers | 115 | 112 | 2 | 2 |
| Cloud computing | 20100 | 18900 | 4070 | 2970 |
| Total | 20754 | 19524 | 4110 | 2996 |
| Ubiquitous | 1330 | 1230 | 33 | 23 |
| Ubiquitous | 714 | 684 | 11 | 9 |
| Ubiquitous | 42000 | 38300 | 4300 | 2840 |
| Total | 44044 | 40214 | 4344 | 2872 |

Table 2. Testing the selected computer terms with the use of Google Patents, 05.12.2011

| **Name of the term** | **Number of patents found by searching** | | | |
|---|---|---|---|---|
| | **with the exact phrase** | | **in the patent titles** | |
| | **patent applications** | **patents issued** | **patent applications** | **patents issued** |
| Quantum computer | 448 | 394 | 33 | 26 |
| Quantum computers | 452 | 358 | 2 | 4 |
| Quantum computing | 407 | 409 | 32 | 26 |
| Total | 1307 | 1161 | 67 | 56 |
| Bacterial computer | 0 | 0 | 0 | 0 |
| Bacterial computers | 0 | 0 | 0 | 0 |
| Bacterial computing | 1 | 1 | 0 | 0 |
| Total | 1 | 1 | 0 | 0 |
| Cognitive computer | 2 | 5 | 13 | 8 |
| Cognitive computers | 0 | 0 | 0 | 0 |
| Cognitive computing | 18 | 23 | 2 | 0 |
| Total | 20 | 28 | 15 | 8 |
| Cellular computer | 65 | 52 | 48 | 39 |
| Cellular computers | 4 | 11 | 3 | 2 |
| Cellular computing | 21 | 8 | 3 | 7 |
| Total | 90 | 71 | 54 | 48 |
| Cloud computer | 83 | 11 | 10 | 1 |
| Cloud computers | 4 | 1 | 0 | 0 |
| Cloud computing | 274 | 51 | 109 | 5 |
| Total | 366 | 63 | 119 | 6 |
| Ubiquitous computer | 55 | 74 | 6 | 1 |
| Ubiquitous computers | 12 | 28 | 0 | 0 |
| Ubiquitous computing | 533 | 540 | 13 | 10 |
| Total | 600 | 642 | 19 | 11 |

On the basis of the Tables 1 and 2 there were calculated the patent capacity factors of the consolidated primary research results clusters (Table 3). We can see from the Table 3 that the patent capacity of the consolidated research cluster produced by the terms beginning with the word “cellular” exceeds the patent capacity of the other consolidated primary research results clusters on the two-three order.

Table 3. The factor of patent capacity of the consolidated primary research results clusters produced by different terms, 05.12. 2011 г.

| **Name of the term** | **Searching throughout the research paper and patent issued** | **Searching in the titles of the research papers and patents issued** |
|---|---|---|
| Quantum computer (computers, computing) | 0. 022 | 0.019 |
| Bacterial computer (computers, computing) | 0.016 | 0.0 |
| Cognitive computer (computers, computing) | 0.016 | 0.089 |
| Cellular computer (computers, computing) | 0.051 | 0.774 |
| Cloud computer (computers, computing) | 0.003 | 0.002 |
| Ubiquitous computer (computers, computing) | 0.016 | 0.004 |

In the Table 4 by the example of the 5-year intervals there was presented the dynamics of the research papers, books, applications and patents issued the titles of which include the terms “quantum computing” and “cloud computing”. The 25-year period covered captures the time of appearance of the narrow publication and patent clusters under consideration (late 80’s of the $20^{th}$ century). This table represents some differences between searching through a wider (1986-2010) and smaller time intervals. Even more significant differences are to be observed by comparing the total responses (1986-2010) on the research papers and patents with the data indicated in the Tables 1 and 2. Generally, the data in the Table 4 when compared with the Table 1 and 2 are underestimated. In the first case it is related to the exclusion of the papers with poorly structured meta data concerning the year of the article issue, in the second one – to the features of the Google Patents algorithm (the smaller time intervals are chosen for the specified time period the better accuracy is achieved by calculation of the aggregate responses within the entire period specified).

Table 4. Dynamics of the research papers, books, applications and issued patents the titles of which contain the terms “quantum computing” and “cloud computing”, 28.04.2012

| **Time interval** | **Quantum computing** | | | | **Cloud computing** | | | |
|---|---|---|---|---|---|---|---|---|
| | **Articles** | **Books** | **Applications** | **Issued patents** | **Articles** | **Books** | **Applications** | **Issued patents** |
| 1986-1990 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1991-1995 | 21 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| 1996-2000 | 207 | 38 | 8 | 2 | 0 | 1 | 0 | 0 |
| 2001-2005 | 681 | 50 | 18 | 7 | 4 | 0 | 0 | 0 |
| 2006-2010 | 589 | 52 | 13 | 12 | 1740 | 113 | 143 | 1 |
| Total | 1501 | 141 | 39 | 21 | 1745 | 114 | 143 | 1 |
| 1986-2010 | 1550 | 96 | 35 | 21 | 1750 | 113 | 143 | 1 |

In order to represent origination and dynamics of the research papers and patent clusters in the Table 5 and 6 the terms “quantum computing” and “cloud computing” were tested within the annual intervals from the moment of appearance of the first responses to the articles (Table 5) and patents (Table 6). The narrow articles and patents clusters were considered (searching in the document titles), and the “at least summaries” was used by working with the Google Scholar.

We can see from the Table 5 that the first research paper the title of which contains the term “ubiquitous” was published in 1992. About three-fold jump in the publication growth took place in 1999 and 2001, the further growth was observed until 2006 following which the decay of the publishing activity took place (Table 5).

Table 5. Temporal dynamics of the research papers the titles of which contain the terms "ubiquitous computing" and "cloud computing", 04.05.2012

| **Years** | **Ubiquitous computing** | **Cloud computing** |
|---|---|---|
| 1991 | 0 | 0 |
| 1992 | 1 | 0 |
| 1993 | 4 | 1 |
| 1994 | 3 | 1 |
| 1995 | 4 | 0 |
| 1996 | 4 | 0 |
| 1997 | 5 | 0 |
| 1998 | 8 | 0 |
| 1999 | 26 | 0 |
| 2000 | 30 | 0 |
| 2001 | 108 | 2 |
| 2002 | 149 | 0 |
| 2003 | 223 | 1 |
| 2004 | 271 | 1 |
| 2005 | 361 | 0 |
| 2006 | 325 | 1 |
| 2007 | 268 | 7 |
| 2008 | 257 | 67 |
| 2009 | 209 | 483 |
| 2010 | 216 | 1180 |
| 2011 | 183 | 1860 |
| Total | 2655 | 3604 |

Origination of the research papers cluster for the term "cloud computing" can be referred to the year 2007 (Table 5). As we can see, the publications growth progressed until 2011. The Table 5 presents good examples of the slow- and rapidly growing research papers clusters.

As we can see from the Table 6, the rapidly-growing patent cluster in the area ubiquitous computing still has not been established.

The patent activity in this area is very low. At the same time in the cloud computing area the rapidly-growing patent cluster for applications originated in the year 2008 which resulted in development into a great number thereof (178) within a short 4-year period of time. For now such applications resulted in a small number of patents.

Table 6. Dynamics of applications and issued patents the titles of which contain the terms "ubiquitous computing" and "cloud computing", 25.04.2012

| **Years** | **Ubiquitous computing** | | **Cloud computing** | |
|---|---|---|---|---|
| | **Applications** | **Issued patents** | **Applications** | **Issued patents** |
| 1998 | 0 | 0 | 0 | 0 |
| 1999 | 1 | 0 | 0 | 0 |
| 2000 | 1 | 0 | 0 | 0 |
| 2001 | 0 | 0 | 0 | 0 |
| 2002 | 3 | 0 | 0 | 0 |
| 2003 | 0 | 0 | 0 | 0 |
| 2004 | 4 | 0 | 0 | 0 |
| 2005 | 5 | 1 | 0 | 0 |
| 2006 | 2 | 2 | 0 | 0 |
| 2007 | 1 | 1 | 0 | 0 |
| 2008 | 2 | 1 | 12 | 0 |
| 2009 | 0 | 3 | 53 | 0 |
| 2010 | 0 | 2 | 78 | 1 |
| 2011 | 0 | 2 | 35 | 6 |
| Total | 19 | 12 | 178 | 7 |

By comparing the Table 5 and 6 one may assume that in the ubiquitous computing research area no critical mass of the research papers was formed which could have resulted in origination of the rapidly growing patent cluster for applications in contrast to the cloud computing area. A slightly greater number of patents issued for the term "ubiquitous computing" is related exclusively to the fact that applications for such patents were issued in small amounts starting from the year 1999, i. e., 10 years earlier than those for cloud computing.

Now let's illustrate the calculation of quality metrics of the research papers clusters for the investigated terms according to the formulas (1, 2).

On the basis of the Table 1 let's select the terms with more than 1000 responses received by testing through the advanced Google Scholar searching with the exact phrase in the line «in the title of the article» («at least summaries» option). These are the terms: "quantum computing", "cloud computing", "ubiquitous computing". Let us take as $\gamma_1$ and $\gamma_2$ the following values: $\gamma_1 = \frac{1}{3}$, $\gamma_2 = \frac{2}{3}$. The results of the $I_1$, $I_2$ metrics calculations are presented in the Table 7.

Table 7. Calculation of the metrics of the research papers clusters quality for the selected terms within the 5-year time interval (2001-2005). Google Scholar, 05.11.2012

| **Terms** / **Indicators** | **Quantum computing** | **Cloud computing** | **Ubiquitous computing** |
|---|---|---|---|
| N | 667 | 7 | 1120 |
| C | 6225 | 39 | 17096 |
| $I_1$ | 0.44 | 0.0 | 1.0 |
| $I_2$ | 0.43 | 0.0 | 1.0 |

As we can see, calculations according to the formulas (1) and (2) delivered almost the same results. If we take the quality potential of the research papers cluster for the term “ubiquitous computing” for 100% then with respect to it the quality potentials of the research papers clusters produced by the terms “quantum computing” and “cloud computing” make 44% and 0% for the $I_1$ metrics and 43% and 0% - for the $I_2$ metrics, respectively.

In the Table 8 there were presented the most-cited and the earliest works indexed concerning the research papers clusters investigated by us. Please note than the same article can be published as preprint as well as in different sources – in a magazine, as a report, in the conference works, etc., that’s why by calculating the links to such an article we receiver the overestimated total number of the relevant links. Because of the different versions of the same article the data in the Table 8 are not stable. Thus, we have found 1275 links to the most-cited article (searching by the article titles) of the research papers cluster produced by the term “cloud computing” that was written by M. Armburst, et al. (2009) [15] as of the middle of November, 2011, and 3191 links – as of 26.10.2012.

Table 8. The most-cited and the earliest works indexed in the clusters produced by the selected terms, 08-15.11.2011.

| **Name of the term** | **The most-cited work** | | **The earliest work indexed** | |
|---|---|---|---|---|
| | **Searching throughout the article** | **Searching in the article titles** | **Searching throughout the article** | **Searching in the article titles** |
| **Quantum computer** | D Deutsch - Proceedings of the Royal Society of …, 1985 - rspa.royalsocietypublishing.org Quantum theory, the Church-Turing principle and the universal quantum computer Cited by 2909 | D Deutsch - Proceedings of the Royal Society of …, 1985 - rspa.royalsocietypublishing.org Quantum theory, the Church-Turing principle and the universal quantum computer Cited by 2909 | KK Likharev - International Journal of Theoretical Physics, 1982 – Springer Classical and quantum limitations on energy consumption in computation | D Deutsch - Proceedings of the Royal Society of …, 1985 - rspa.royalsocietypublishing.org Quantum theory, the Church-Turing principle and the universal quantum computer |
| **Quantum computers** | MA Nielsen, I Chuang… - American Journal of Physics, 2002 - link.aip.org Quantum computation and quantum information Cited by 13708 | J Preskill - Proceedings of the Royal Society of …, 1998 - rspa.royalsocietypublishing.org Reliable quantum computers Cited by 623 | RP Feynman - International journal of theoretical physics, 1982 – Springer Simulating physics with computers | A Peres - Physical Review A, 1985 – APS Reversible logic and quantum computers. |
| **Quantum computing** | T Leighton - sce.uhcl.edu Introduction to parallel algorithms and architectures, 1992 Cited by 3037 | A Steane - Reports on Progress in Physics, 1998 - iopscience.iop.org Quantum computing Cited by 865 | A Gerstenfeld… - Business Horizons, 1970 – Elsevier Why engineers transfer:: Survey pinpoints reasons for job changes | JD Brasher, CF Hester… - International journal of theoretical …, 1991 – Springer Virtually-deterministic quantum computing of nondeterministic polynomial problems |
| **Cloud computer** | W Dungan Jr - ACM SIGGRAPH Computer Graphics, 1979 - portal.acm.org A terrain and cloud computer image generation model Cited by 21 | W Dungan Jr - ACM SIGGRAPH Computer Graphics, 1979 - portal.acm.org A terrain and cloud computer image generation model Cited by 21 | 1979 W Dungan Jr - ACM SIGGRAPH Computer Graphics, 1979 - portal.acm.org A terrain and cloud computer image generation model | 1979 W Dungan Jr - ACM SIGGRAPH Computer Graphics, 1979 - portal.acm.org A terrain and cloud computer image generation model |
| **Cloud computers** | D Wentzlaff, C Gruenwald III, N Beckmann… - Proceedings of the …, 2010 - portal.acm.org An operating system for multicore and clouds: mechanisms and implementation Cited by 22 | C Moretti, K Steinhaeuser, D Thain… - 2008 Eighth IEEE …, 2008 - computer.org Scaling up classifiers to cloud computers Cited by 25 | RK Bagga - Defence Science Journal, 1993 - publications.drdo.gov.in [PDF] Defence Research & Development Laboratory, Hyderabad-500 258 | C Moretti, K Steinhaeuser, D Thain… - 2008 Eighth IEEE …, 2008 - computer.org Scaling up classifiers to cloud computers Cited by 25 |

| | | | | |
|---|---|---|---|---|
| **Cloud computing** | T Leighton - sce.uhcl.edu<br>Introduction to parallel algorithms and architectures<br>Cited by 3038 | M Armbrust, A Fox, R Griffith, AD Joseph… - … , University of California …, 2009 – Citeseer<br>[PDF] Above the clouds: A berkeley view of cloud computing<br>Cited by 1275 | G Lavendel - Chemical And Engineering News, 1977 - eric.ed.gov<br>Special Libraries. | Y Guoding… - Journal of Southeast University (Natural …, 2003 - en.cnki.com.cn<br>Cloud computing: a method to realize conceptual computing |
| **Ubiquitous computer** | K Birman - Worldwide Computing and Its Applications, 1997 – Springer<br>Building secure and reliable network applications<br>Cited by 476 | A Schmidt, M Kranz… - … of the 2005 joint conference on …, 2005 - portal.acm.org<br>Interacting with the ubiquitous computer: towards embedding interaction<br>Cited by 38 | 1964<br>AC Beer - Science, 1964 - sciencemag.org<br>Imperfections and Active Centres in Semiconductors. RG Rhodes. Pergamon, London; Macmillan, New York | L Sydiinheimo, M Salmimaa… - … , 1999. ICC'99. 1999 …, 1999 - ieeexplore.ieee.org<br>Wearable and ubiquitous computer aided service, maintenance and overhaul |
| **Ubiquitous computers** | M Weiser… - Xerox PARC. Retrieved July, 1996 – Citeseer<br>[PDF] The coming age of calm technology [1]<br>Cited by 578 | T Masui - Handheld and Ubiquitous Computing, 1999 – Springer<br>POBox: An efficient text input method for handheld and ubiquitous computers<br>Cited by 66 | KH Pribram - Daedalus, 1980 – JSTOR<br>The role of analogy in transcending limits in the brain sciences | T Masui - Handheld and Ubiquitous Computing, 1999 – Springer<br>POBox: An efficient text input method for handheld and ubiquitous computers |
| **Ubiquitous computing** | H Ishii… - Proceedings of the SIGCHI conference on Human …, 1997 - portal.acm.org<br>Tangible bits: towards seamless interfaces between people, bits and atoms<br>Cited by 2481 | J Hightower… - Computer, 2001 - ieeexplore.ieee.org<br>Location systems for ubiquitous computing<br>Cited by 2435 | J Fulk, J Schmitz… - Organizations and communication …, 1990 - mendeley.com<br>A social influence model of technology use | M Weiser - Scientific American, 1991 - ubiq.com<br>[PS] Ubiquitous computing |

If in the first case this article was published in Cite seer [15] then in the second case it was presented in the magazine "Communications of the ACM" (vol. 53, № 4) for the year 2010 [16] and generated already 3191 links. Also the title of the article has changed – "A view of cloud computing". But relating to the content in both cases we have to do with the same article.

During the annual period under consideration the number links to the most-cited works H. Ishii, et al (1997) [17] and J. Hightower, et al (2001) [18] for the research publications cluster produced by the term "ubiquitous computing" increased significantly and made 2844 and 2821, respectively, at the same time the first article ranked 32th in the list of the Google Scholar responses, and its place was filled by the article G. Abowd, et al.(1999) [19] with the best result of 2950 links.

As to the second article, there are links to different versions of this article in which even the sequence of the authors is changed. All of this points to the fact that Google Scholar shall be very carefully used for citation analysis that is confirmed by the findings of the studies [20] as well.

It may be concluded that it makes sense for the Google Scholar team to develop the statements (operators) that would allow searching the publications separately by the databases Web of Science and Scopus. In such case the importance of this tool would increase significantly.

In the Table 9 there was presented the frequency of occurrence of the names of organizations where researches were conducted, titles of journals and other publications in the first ten Google Scholar responses for the selected terms. The search was performed in the article titles with the use of the "at least summaries" option. In each response the same organization was indicated only once regardless the number of the authors from this organization.

Table 9. Frequency of occurrence of the names of organizations, titles of journals and other publications in the first ten responses to the Google Scholar requests on the selected terms, 08-09.11.2011

| **Terms** | **Organization** | **Magazines and other publications** |
|---|---|---|
| Quantum computing, Quantum computer, Quantum computers | University of Oxford, UK (10); AT & T Bell Laboratories, USA (4); University of California, Los Angeles, USA (3); Hewlett-Packard Laboratories, Bristol, UK (3); University of Innsbruck, Austria (2); Ludwig Maximilians University, Munich, Germany (2); Los Alomos National Laboratory, USA (2); IBM Thomas J. Watson Research Center, USA; Masaryk University, Brno, Czech Republic; University of Vienna, Austria; Imperial College, London, UK; Max Plank Institute for Quantum Optics, Garching, Germany; Erwin Schrödinger Institute for Mathematical Physics, Germany; University of Leeds, UK; Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, Austria; Cornell University, USA; National Institute of Informatics, Tokyo, Japan; University of Queensland, Australia; Louisiana State University, USA; Texas A&M University, USA; University of New South Wales, Australia; University of Michigan, USA; Massachusetts Institute of Technology, USA; California Institute of Technology, USA; University of British Columbia, Vancouver, Canada; Israel Institute of Technology, Haifa, Israel; Stanford University, USA; Microsoft Research, USA; Indiana University, USA; University of Barcelona, Spain; City University of New York, USA | Nature (7); Phys. Rev. A (5); Phys. Rev. Letters (4); Proc. of the Royal Soc. A (3); Science (2); Rep. Prog. Phys.; Rev. Mod. Phys.; J. of Chem. Phys.; Comm. in Math. Phys.; APS Journal; Nanotechnology; arxiv. org; Colloquim; Book |

| | | |
|---|---|---|
| Cloud computing,<br>Cloud computer,<br>Cloud computers | University of California, Santa Barbara, USA (2); Linkoping's University, Sweden (2); UC Berkeley Reliable Adaptive Distributed Systems Laboratory, USA (2); Intel Corporation, USA, Germany (2); Fert University, Sweden (2); University of Melbourne, Australia; Microsoft Pty Ltd, Melbourne, Australia; University of Chicago, USA; Vienna University of Technology, Austria; Institute of Sci. Comput., Karlsruhe, Germany; RWTH Aachen University, Germany; ETH Zurich, Switzerland; North Carolina State University, USA; Intel Technol., Bangalore, India; Chengdu University of Information Technology, China; University of Notre Dame, USA; Georgia State University, USA | Communications of the ACM (2); Lecture Notes in Computer Science (2); J. of Computing and Information Technology; Future Generation Computer Systems; Computer and Engineering Science; Journal of Solid-State circuits, IEEE; ACM SIGGRAPH Computer Graphics; ACM SIGOPS Operating Systems Review; Conference (4); Workshop (3); Symposium (2); Preprint (2); Manual, Book |
| Ubiquitous computing,<br>Ubiquitous computer,<br>Ubiquitous computers | National Central University, Taiwan (3); University of Munich, Germany, (2); Xerox PARC, Palo Alto, USA (2); Washington University, USA; Georgia Institute of Technology, USA; Stanford University, USA; Carnege Melon University, USA; University of Hawaiian Manoa, USA; San Jose State University, USA; University of Cambridge, UK; University of Birmingham, UK; Sheffield Hallam University, UK; Rank Xerox Euro PARC, Cambridge, UK; University of Burnaby, Canada; Fraunhofer IPSI, Germany; University of Aarhus, Denmark; University of Lapland, Finland; University of Jyvaskyla, Finland; Tampere University of Technology, Finland; TUCS, Turku, Finland; Parko Scientifico e Tecnologico, Italy; Sony Computer Science Laboratory, Japan; Keio University, Japan; Kobe University, Japan; National Institute Advanced Industrial Science and Technology, Japan; Kainan University, Taiwan; Polaris Company, India | Lecture Notes in Computer Science (4); Computer (3); Personal Communications, IEEE; Pervasive Computing, IEEE; MICRO, IEEE; Information Processing; Communications of the ACM; ACM Transactions on Computer-Human Interaction; Personal and Ubiquitous Computing; International Journal of Mobile Learning and Organization; Oyo Butsuri (Japan); Joho Shori (Japan); Conference (7); Book (2); Workshop; PhD Dissertation |

In the first cluster of publications relating to the quantum computing and computers the highly cited journals for natural sciences and physics prevailed while in the other two publication clusters – the computer journals did.

The frequency of occurrence of the organization names in the publication clusters depending on the countries where the authors work is presented in the Table 10. As we can see from it only the USA and Germany are presented in all the publication clusters. There are groups of countries that are represented in a single publication cluster only. Thus, Finland and Taiwan are significantly represented in the last publication cluster only.

Table 10. Frequency of the organization names occurrence for the selected terms by countries, 08-09.11.2011

| **Country** | **Quantum computing (computer, computers)** | | **Cloud computing (computer, computers)** | | **Ubiquitous computing (computer, computers)** | |
|---|---|---|---|---|---|---|
| | **Occurrence of the organization names in general** | **Occurrence of the different organization names** | **Occurrence of the organization names in general** | **Occurrence of the different organization names** | **Occurrence of the organization names in general** | **Occurrence of the different organization names** |
| USA | 35 | 14 | 8 | 6 | 8 | 7 |
| United Kingdom | 15 | 4 | - | - | 4 | 4 |
| Germany | 4 | 3 | 4 | 3 | 3 | 2 |
| Austria | 4 | 3 | 1 | 1 | - | - |
| Australia | 4 | 3 | 2 | 2 | - | - |
| Japan | 1 | 1 | - | - | 4 | 4 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Spain | 1 | 1 | - | - | - | - |
| Czech Republic | 1 | 1 | - | - | - | - |
| Canada | 1 | 1 | - | - | 1 | 1 |
| Israel | 1 | 1 | - | - | - | - |
| Sweden | - | - | 4 | 2 | - | - |
| Switzerland | - | - | 1 | 1 | - | - |
| India | - | - | 1 | 1 | 1 | 1 |
| China | - | - | 1 | 1 | - | - |
| Finland | - | - | - | - | 4 | 4 |
| Taiwan | - | - | - | - | 4 | 2 |
| Denmark | - | - | - | - | 1 | 1 |
| Italy | - | - | - | - | 1 | 1 |

**CONCLUSION**

Within the research we mean by the publication and patent clusters the combinations of publications and patents produced by random scientific terms. Such clusters in terms of research papers, books and patents can be effectively identified with the use of Google Scholar, Google Books and Google Patents, respectively. They can be constructed whether within a particular time interval or at the present time. The above-mentioned clusters are understood both generally and properly. In the first case it is assumed that the specified term is used in the research publications and patents texts, in the second one – in the titles. In the latter case we arrive at the most relevant publication and patent clusters produced by such terms. Consideration of the research papers and patent clusters produced by the same scientific terms allows introducing the concept of the factor of capacity of the primary research results clusters combining the research papers and patents on the specified limited field.

In order to evaluate the research papers cluster we consider the vector the coordinates off which are the number of articles in the cluster and total citation ratio TOP-10% of the most-cited articles. On the basis of this vector the additive and multiplicative metrics are constructed.

In each research papers cluster Google Scholar allows identifying the most-cited and the earliest articles indexed. For a certain number of the first responses one can determine the rate of occurrence of organizations the writers are employed at and magazines in which they are published.

We selected different names of the kinds and computer calculations as the scientific terms for testing by means of the specialized Google tools, such terms begin with the words: Quantum, Bacterial, Cognitive, Cellular, Cloud, Ubiquitous.

Based on the results of experiments with these terms it was shown that the most powerful (in terms of the number of articles and patents) research paper and patent clusters were correlated with the terms beginning with the words quantum and ubiquitous. It was shown that the patent capacity of the consolidated research cluster produced by the terms beginning with the word Cellular exceeded the patent capacity of the other consolidated clusters of the primary research studies by 2-3 orders.

Analysis of the temporal dynamics of the research papers, books, applications and issued patent clusters performed by searching by the document titles showed that at the end of the 80's of the 20$^{th}$ century there appeared the publication and patent clusters produced by the terms "quantum computing" and "cloud computing". It was also shown that in the ubiquitous computing research area no critical mass of the research papers was formed which could have resulted in origination of the rapidly growing patent cluster for applications in contrast to the cloud computing area.

The calculations of the additive and multiplicative metrics delivered the same results. If we take the quality potential of the research papers cluster for the term "ubiquitous computing" for 100% then with respect to it the quality potentials of the research papers clusters produced by the terms "quantum computing" and "cloud computing" make 43-44% for the first metric and 0% for the second one, respectively.

The significant temporal variability of the links to the most-cited works by the use of Google Scholar was observed.

This point to the fact that Google Scholar shall be very carefully used for citation analysis that is confirmed by the findings of the other studies as well.

It was concluded that it makes sense for the Google Scholar team to develop the statements (operators) that would allow searching the publications separately by the databases Web of Science and Scopus. Moreover, it is required to add options that would exclude from the search results the documents different from the research papers. For now such an option is available for the patent exclusion only.

By analyzing the first ten responses to the selected terms requests it was shown that in the cluster of publications relating to the quantum computing and computers the highly cited magazines for natural sciences and physics prevailed while in research papers clusters relating to the cloud and ubiquitous computing and computers – the computer magazines did.

It was shown that only the researchers from the USA and Germany were represented in all the three clusters of the most-cited publications and the scientists from Finland and Taiwan were significantly represented in a single cluster produced by the terms ubiquitous computing (computer, computers).

**REFERENCES**


1. Osareh, F., 1996, "Bibliometrics, Citation Analysis and Co-Citation Analysis*:* A Review of Literature I," Libri, 46 (3), pp.149-158.
2. White, H. D., and McCain, K.W., 1998, "Visualizing a discipline: An author co-citation analysis of information science, 1972-1995," Journal of the American Society for the Information Science, 49 (4), pp.327-355.
3. Moskovkin, V. M., 2012, " Construction of the research results clusters with the use of the Google tools," Scientific and Technical Information. Series 2, 8, pp. 9-13 (In Russian).
4. Noruzi, A., 2005, "Google Scholar: The New Generation of Citation Indexes," Libri, 55, pp.170-180.
5. Robinson, M. L., and Wusteman, J., 2007,"Putting Google Scholar to the test: a preliminary study," Program: electronic library and information systems,. 41 (1), pp.71 – 80.
6. Aalst, J. V., 2010, "Using Google Scholar to estimate the impact of journal articles in education," Educational Researcher, 39 (5), pp.387-400.
7. Mastrangelo, G. et al., 2010, "Literature search on risk factors for sarcoma: Pub. Med and Google Scholar may be complementary sources," BMC Research Notes.-2010,3, pp.131(http://www.biomecentral.com/1756-0500/3/131)

8. Walters, W. H., 2011, "Comparative recall and precision of simple and expert searches in Google Scholar and eight other databases," Portal: Libraries and the Academy, 11 (4), pp. 971-1006.
9. Moskovkin, V. M., 2009, "Simulation expert system for making students' college decisions," Automatic Documentation and Mathematical Linguistics, 43 (5), pp. 292-295. - Mode of access: (http://link.springer.com/article/10.3103/S0005105509050057).
10. Kousha, K., Thelwall, M., Rezaie, S., 2011, "Assessing the citation impact of books: The role of Google Books, Google Scholar, and Scopus," Journal of the American Society for Information Science and Technology, 62 (11), pp.2147-2164
11. Abdullah, A., and Thelwall, M., 2013, "Can the impact of non -Western Academic Books be measured? An investigation of Google Books and Google Scholar for Malaysia," Journal of the American Society for Information Science and Technology (in Press).
12. Moskovkin, V. M., 2011, "Open access to scientific knowledge and feudalism knowledge: Is there a connection?" Webology, 8 (1). - Mode of access: (http://www.webology.org/2011/v8n1/a83.html).
13. Aoki, R., and Schiff, A., 2008, "Promoting access to intellectual property: patent pools, copyright collectives, and clearinghouses," RgD Management, 38 (2), pp.189-204.
14. Moskovkin, V. M., Delux, T., Moskovkina, M.V., 2012, "Comparative Analysis of University Publication Activity by Google Scholar (On Example of Leading Czech and Germany Universities)," Cybermetrics, 16 (1), paper 2. Mode of access: (http://cybermetrics.cindoc.csic.es/vol16iss1.html).
15. Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R. H., Konwinski, A., Lee, G., Patterson, D. A., Rabkin, A., Stoica, I., Zaharia, M.2009, "Above the clouds: A Berkeley view of cloud computing," University of California, Berkeley.
16. Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R. H., Konwinski, A., Lee, G., Patterson, D. A., Rabkin, A., Stoica, I., Zaharia, M., 2010, "A view of cloud computing," Communications of the ACM, 53 (4), pp.50-58.
17. Ishii, H., and Ullmer, B., 1997, "Tangible bits: towards seamless interfaces between people, bits and atoms," Proceedings of the ACM SIGCHI Conference on Human factors in computing systems, pp.234-241.
18. Hightower, J., and Borriello, G., 2001, "Location systems for ubiquitous computing," Computer, 34 (8), pp.57-66.
19. Abowd, G. D., Dey, A. K., Brown, P. J., Davies, N., Smith, M., Steggles, P., 1999, "Towards a Better Understanding of Context and Context-Awareness," Lecture Notes in Computer Science, 1707, pp. 304-307.
20. López-Cózar, E. D., and Robinson-Garcia, N., 2012, "Repositories in Google Scholar metrics or what is this document type doing in place as such?," Cybermetrics, 16, (1),paper 4. Mode of access: (http://cybermetrics.cindoc.csic.es/articles/v16i1p4.html)