

## THE OPTIMAL SEARCH FOR INFORMATION ON THE INTERNET

Andreyev K.V., student of the group ITb-171, 1<sup>st</sup> course  
Kuran N. V., student of the group ITb-171, 1<sup>st</sup> course  
Scientific adviser: Lebedintsev S. V., senior teacher  
T.F. Gorbachev Kuzbas State Technical University  
Kemerovo

Every year the volume of the Internet is going up many times, as it has already placed several billion web pages and files which continue to grow every day, so the probability to find the necessary information you need is dramatically increasing. Nevertheless, when you go to the Internet the main problem is: not the lack of related information but the opportunity to find it. It is quite possible that you have already faced this problem to find and to extract quickly the necessary information.

Search for information is the process of identifying documents (texts) that are devoted to the specified topic (subject), meet a predetermined search condition (query) or contain the necessary facts, information and data. The search process includes an algorithm that collects, processes, and provides information. Searching for information involves four steps:

1. «definition or specification of information needs and formulation of a query;
2. determination of the complex potential holders of data (sources);
3. retrieving information from the identified information files;
4. acquaintance with the received information and an assessment of search results» [3].

Let's have a detailed look at some methods to solve this problem. The main subject of our research is to consider basic techniques of information retrieval on the Internet, notably to show:

- 1) the common features of search tools,
- 2) some structures of search request for the most popular both Russian-speaking and English-speaking search systems.

To extract quickly the necessary information *search technologies* and *search tools* are used.

*Search technologies* are the set of rules and procedures, as a result of which the user obtains the desired result. When searching on the Internet, it is recommended to pay attention to two components: completeness and accuracy. Web-technology *World Wide Web* (WWW) which includes web pages, electronic libraries and catalogs and so on is considered a special technology for preparing and placing documents on the Internet. With such an abundance of information, the question arises of how to navigate in such a huge and the same time large-scale in-

formation space. It turns out that it is not so difficult. To solve this problem some search tools are used [1].

*Search tools* belong to special software. The main purpose of this software is to provide with the most optimal and qualitative information search for Internet users. Search tools which are placed on special web servers operate on the basis of algorithm. Each web server performs a certain function which aim is:

- to analyse web pages,
- to enter analysis results to a certain level of the search engine database,
- to search for information on the user's request,
- to provide with convenient interface for searching information and viewing the search result [1].

Each of the search tools uses different criteria for ranking documents. So, if you want to specify a query of the same design in the search string for each search tool you can get different search results. And one can agree because there is the same difference when the reference to one or another defined site is in the top-10 search but not the top-20.

In order to set a retrieval request correctly first of all, one should increase a search efficiency by using logical operators «or», «and», «near», «not», mathematical and special symbols in queries. Using operators «and» / «or» symbols, the user connects the keywords in the required sequence to get the search result that is most adequate to the query. The operator «and» give you an opportunity to specify that all the keywords should be included in the content of the document. Nevertheless, the number of documents can still be large, and their viewing can take long. Therefore in some cases it is much more convenient to use the context operator «near» indicating that words should be located in the document in sufficient proximity. Using «near» reduces greatly the number of found documents [2].

The presence of the «\*» symbol in the search string means that the word will be searched for by its mask. For example, we get a list of documents containing words beginning with «org», if we write «org \*» in the search string. These can be the words *organize*, *organization*, *organizer*, *organic* etc. In order to exclude documents containing certain terms, use the «-» sign before each word. For example, if you need information about Tolstoy's works, with the exception of «War and Peace», then enter the query in the form «Tolstoy-War and Peace».

On the contrary, in order to include certain references in the search results use the symbol «+». For example, in order to find references to buying or to sale of books, you need to write a query «purchase + book» or «sale +book» respectively. One should use combinations of these symbols to increase the efficiency and accuracy of the search.

To get the most effective search one should pay attention to the following points in advance. You ought:

- to determine the subject of the request what exactly it should be found,
- to pay attention to language, grammar, morphology,
- to formulate and to enter keywords correctly.

Each reference being in the list of search results contains a snippet. Snippet is a small piece or brief extract – several lines from the found document, among which can be occurred as someone's keywords. Before clicking on the reference, evaluate the snippet relevance to the query topic. After clicking on the reference to the defined site, look through the main page carefully. As a rule it would be enough to view the first web page to understand if you have come to Uniform Resource Locator (it's called for short URL).

In conclusion it should be noted that it is very important in browsing:

1) to formulate exactly a search query that reflects accurately the information to be needed;

2) to use operators «and», «or», «near», and special symbols: «-», «+», «\*» etc.;

3) to pay attention to the snippet because of its description you can often understand whether there is a website on the right information.

### **References:**

1. Технологии поиска информации в сети Интернет. Смоленск : Изд-во Смоленск. гос. ун.-та. 2010. / <http://globuss24.ru/doc/tehnologii-poiska-informacii-v-seti-internet>.
2. Технология поиска информации в Интернете: информационные поисковые системы Интернет / <https://www.seonews.ru/masterclasses/tehnologiya-poiska-informatsii-v-internet-vidyi-poiskovyih-instrumentov-informatsionnyie-poiskovyie-sistemyi-interneta/>.
3. Jansen, B. J. and Rieh, S. The Seventeen Theoretical Constructs of Information Searching and Information Retrieval. Journal of the American Society for Information Sciences and Technology. 61(8), (2010). P. 1517-1534.