P.A. Boldyrev
Candidate of technical sciences,
Deputy Director for the automation of
bibliographic processes Research Library,
Orenburg State University,
Russia, Orenburg
library_oit@mail.osu.ru

SECTION 4. Computer science, computer engineering and automation.

THE APPLIANCE OF SOFTWARE PRODUCT "YANDEX. SERVER" FOR SEARCHING IN ELECTRONIC LIBRARY CATALOG "E-CATALOG"

Abstract: The stored in libraries enormous number of information sets a problem for the developers to manage fast and qualitative search. The solution is possible by using modern searching tools of Web technology. One of such tools is software program "Yandex. Server" that allows to organize the optimum search in the library e-catalog. The software "Yandex. Server" can query taking into account Russian and English Morphology and also with logical conditions. That provides effective and flexible search in the library e-catalog.
Key words: library e-catalog, searching tool, Yandex technologies.
Language: English
Citation: Boldyrev PA (2016) THE APPLIANCE OF SOFTWARE PRODUCT "YANDEX. SERVER" FOR SEARCHING IN ELECTRONIC LIBRARY CATALOG "E-CATALOG". ISJ Theoretical & Applied Science, 11 (43): 101-104.
Soi: http://s-o-i.org/1.1/TAS-11-43-18   Doi: crossref http://dx.doi.org/10.15863/TAS.2016.11.43.18

Introduction. E-catalog is the core and basis of information library system. The effectiveness of usage of information resource in libraries and the quality of library service whether in local level or online identify linguistic options of subject search in e-catalog. The system of effective search in e-catalog gives quick and qualitative access to information resources of the library. The agreed and proved choice of linguistic support is the problem without which the creation of local and distributed electronic catalogs is meaningless.

The usage in e-catalog of a wide choice of different in structure and lexical content linguistic tools is a consistent trend. Whereas different libraries demonstrate an outstanding variety of approaches to definition and justification of the content and functioning principles of linguistic support. It forces the library community to optimize the complex of retrieval languages thus to define the structure of linguistic support and parameters of each searching language that are necessary and sufficient for effective system functioning. The problem of optimization of linguistic support can be solved on the base of Internet searching tools.

The quick growth and dynamic development of information-communicative technologies (especially Web technologies) gives the opportunity to develop the brand-new approach to the search of libraries information resources. To find the necessary information in electronic resources of the library it's essential to find its address. There are special searching servers (index robots (searching systems), thematic e-catalogs, metasearch systems etc.). Web technology World Wide Web (WWW) is considered to be special technology in preparing and publishing documents in Internet. WWW consists of web pages, electronic libraries, catalogs etc. The stored in libraries enormous number of information sets a problem for the developers to manage fast and qualitative search. The solution is possible by using modern searching tools.

The usage of program "Yandex. Server" for searching in libraries e-catalogs. Searching tools is special software which aim is to provide the most optimal and qualitative information search for Internet users. One of such tools is "Yandex. Server". It is a product for searching on Web server that gives full text search taking into account Russian morphology.

"Yandex. Server" consists of following components: crawler and Indexer.

Crawler is a subsystem that reviews the user's request, looks for documents in index file prepared by indexer and makes report pages with search results. For each collection of documents there is its own searching engine.

Indexer is a subsystem that analyses documents and saves the information about them in special index files. For each collection of documents there is its own set of index files that is saved in certain directory and its own indexer.

Information search pattern is easy. The user types the key phrase, activates search and gets the selection of documents on the request given. The list of documents is ranked on specific criterions so that at the top of the list were that documents which are the most corresponded to the users request (picture 1).
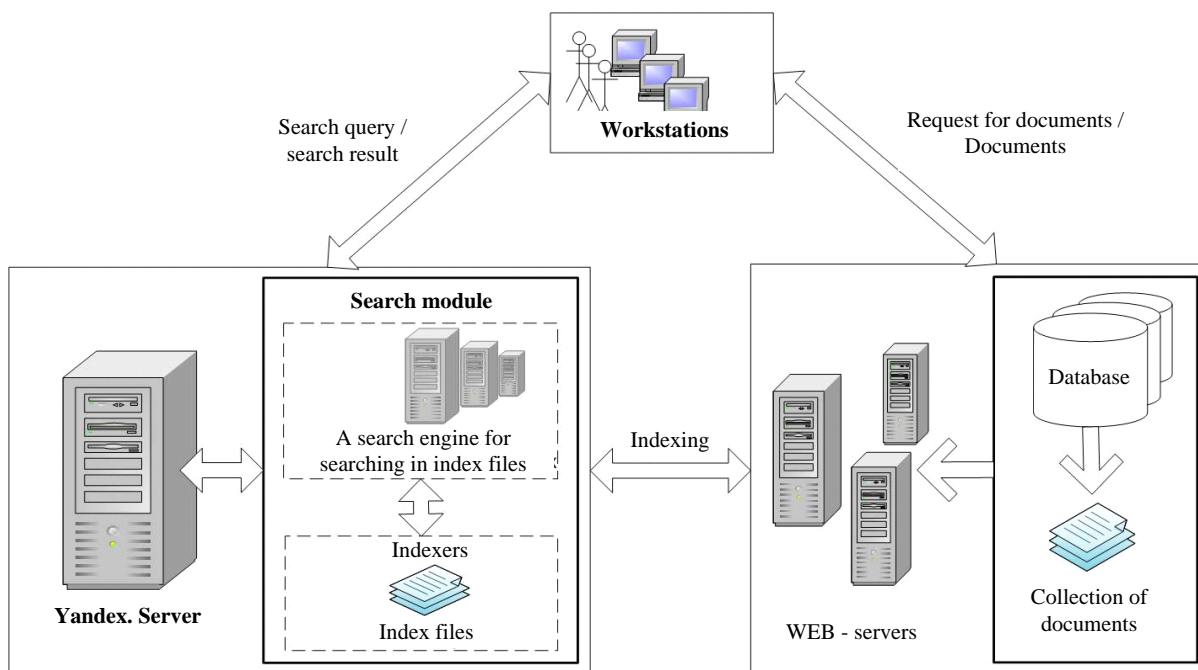
Picture 1 – The organization of searching system.

One installation of "Yandex. Server" allows to index and find necessary information on several Web servers at once unnecessary located physically on the same engine as "Yandex. Server". Consequently the differences in data base structure, working automated library and information systems and arrangement of Web servers don't interfere the combination of their e-catalogs.

"Yandex. Server" works with one or several collections of documents where the searching occurs. The searching in each collection occurs independently. Using "Yandex. Server" to search in e-catalog such documents collections become library bibliographic data bases. So, index and searching happens by e-catalogs independently from each other [3].

Each bibliographic recording in e-catalog contains a lot of data fields. The number of data fields can be changed depending on type of material on which bibliographic recording is created. Standard set of attributes is used for convenience and on the basis of consistency of e-catalogs. Standard set of attributes includes several dozens of attributes but practically used only part such as author; title; subject entry; year of publishing; place of publication; publishing house; key words; mode of publication.

On the base of attributes sets the search query. Each attribute agrees with one or several bibliographic recording fields. Such approach is convenient because it doesn't involve to point in search query bibliographic recording fields.

The search result is the list of bibliographic recordings arranged by pertinence. The pertinence considers not only the quantity of found documents but the contrast of words (popularity) and interval between words.

Flexible setter of "Yandex. Server" allows to select the mode of index without stopping of searching engine. Taking into account big collections of bibliographic recordings this opportunity is important because preindexing is made: new documents are added, changed documents are reindexed and unchanged stay the same. It saves a large part of resources including temporary. The indexing is made without stopping of searching service. There is no limitation on sixe and quantity of indexable documents.

The capabilities of search system "Yandex. Server". Data query language of "Yandex. Server" is

context free grammar which terminals are words of natural language.

The search works according to Russian and English language morphology the same as Internet searching. For example if you make a query "to go", the result will be references on documents connecting words "to go", "goes", "went" etc.

Choosing the word form it can be useful to take into account the word class. For example, if it is searched "computer" (a noun), the result will be bibliographic recordings with the words "computer" or "computers" not computerized (adjective).

If necessary to search the exact word form the morphology can be switched off. For that special symbols quotations and exclamation mark are used. For example, on query "!basis! analyses! electrical" recordings with words "basis", "analyses", "electrical" will be found and searching "!Solovyov" author's Solovyov recordings will be found.

To exclude bibliographic recordings with definite word also special symbols are used – twiddle (~~) or one minus (without space). For example, searching bibliographic recordings on English course books author Bonk you have to make query: "English course book~~Bonk".

"Yandex. Server" gives opportunity to make queries with different logical conditions for example queries with logical condition "or", with exceptions, queries with pointing the distance between words in a sentence etc. If it's exactly known not only the distance but the word order of query in which they go in the text of necessary bibliographic recordings, it's essential to point out between symbol «/» (slash) and the distance number symbol + (plus) direct word order or — (minus) for inversion. For example to find the basis of numerical or discrete mathematics etc. the query is "/+2mathematics". The recordings will be found where between words basis and mathematics there is one third word.

The indexing and searching on different bibliographic attributes allows full unveil the opportunities of advanced search if it's needed to form the query on the base of several searching attributes. Filling in several fields all specified conditions are taken into account together. Whereas for each searching field work the same query

language rules as for classical search forms. For example if it's necessary to find all the recordings with the word "history" without the phrase "from the earliest times" in title written by authors Solovyov and Radugin, in the field author it's set "Radugin|Solovyov" and in the field title – "The History~~the earliest". Advanced search allows to put filters on the searching process narrowing and specifying its area [1,2].

The result shows the list of bibliographic recordings ranged on relevancy. Apart from information about the author and the title bibliographic recordings include subject entries, key words; the blurb of the document if it's included into recording; additional access points; document location data.

Conclusion

Effective search in e-catalog allows to organize quick and qualitative access to library information resources.

The project created on the base of search engine "Yandex. Server". The usage of the search engine "Yandex. Server" allowed to organize "Central E-Catalogue of the libraries of the city Orenburg and Orenburg region" in the form of distributed data base. The purpose of Central E-Catalogue creation is the usage of libraries potential for the benefit of developing of science, culture and education and to support native informational space. Central E-Catalogue provides access for the citizens and libraries staff of Orenburg region, Russia and foreign countries to bibliographic recordings of publications that are stored in Orenburg and Orenburg region libraries and gives the opportunity to find bibliographic recordings of needed publications and the information on their location. The access to the Central E-Catalogue is possible at the following URL — http://svek56.ru/. The development of Central E-Catalogue allows the libraries of the region to move to a new level of professional cooperation and to create a new local informational resource.

References:

1. (2016) "Yandex. Server" - the procedure on your web server. - Available: http://company.yandex.ru/technology/server/. (Accessed: 10.11.2016).

2. (2016) "Yandex. Server" 2010. Manual for installation and operation. - Available: http://api.yandex.ru/server/doc/concepts/About.xml. (Accessed: 10.11.2016).

3. (2016) Yandex: official. blog. - Available: Http://blog.yandex.ru/. (Accessed: 10.11.2016).
4. (2016) Yandex.Server 2009.05: Installation and Operation Manual. - Available: Http://download.yandex.ru/yandex-server/download/yandex-server-manual.pdf. (Accessed: 10.11.2016).
5. Shevchenko V (2010) Information service: fast, efficient, reliable: to change the traditional scheme of service organizations // Library, 2010. - № 8. - pp. 51–53.
6. Marina M (2014) Search engine Yandex // Master's Herald, 2014. - № 4-1 (31). - pp. 82-84.
7. Shokin Y (2010) Search Problems information / Yu. Shokin, A. Fedotov V. Barakhnin. - Novosibirsk: Science, 2010. - 220 p.
8. (2016) "Content Runet" Yandex Newsletters. - Available: http://company.yandex.ru/facts/researches/. (Accessed: 10.11.2016).
9. (2016) Dictionary "Linguistic" Company "Yandex". - Available: http://lingvo.yandex.ru. (Accessed: 10.11.2016).
10. (2016) "Yandex" Ctimer company. - Available: http://company.yandex.ru/technology/mystem/. (Accessed: 10.11.2016).