Study of Information Search Systems of the Internet

Yuri Ramsky and Olga Rezina

Computer Science Department,
M.Drahomanov National Pedagogical University,
9 Pirogov Str., Kyiv, 01030 Ukraine
rama@ukr.net, rezina_olga@mail.ru

Abstract. Qualified search for information in the internet is considered to be a fundamental capability of pupils to be educated for the information society. The theory of planned formation of intellectual actions is assumed as a basis of learning of internet search systems. The paper reports on the computer teaching program "Poshuk-META" (Retrieval Aim), a system to get acquainted with searching the internet even without an internet-access.

1 Introduction

Information and knowledge are determinant factors of the development of modern society and a more important national resource. Information retrieval is the issue of the day in the information society. The retrieval activity is a basis in the educational process of a secondary school, an academy and a professional activity for most of people today.

The problem of information retrieval comes up before pupils at the time of self-dependent and research work, preparation for reports and papers, selection of additional materials for different subjects. This problem can be solved in the process of learning of Informatics. It's necessary to form:

- Skills to define possible sources of information and the strategy of its retrieval;
- Skills to do the semantic processing of data;
- Skills to analyse and interpret findings;
- Skills to appreciate information from the point of view of its authenticity, accuracy, adequacy for the solution of the concrete information task.

The impetuous development and active use of modern information and communication technologies have made it possible to begin the wide-ranging transformation of information to the electronic form and creating of a huge quantity of information resources. The electronic form of the representation of information makes it possible to organize the processes of preservation, processing, spreading, and retrieval of data in a qualitatively new level.

The intensive disposal of data in the internet is the fundamental direction of the technological development of information resources. The quantity of scientific, educational, cultural, social internet-resources increases all the time. They are electronic libraries and encyclopedias, electronic journals and newspapers, materials

R.T. Mittermeir (Ed.): ISSEP 2005, LNCS 3422, pp. 84 – 91, 2005.

about educational institutions, school and students' unions and so on. The educational potential of the internet can be effectively used in studies as well as during the solution of didactical and pedagogical problems. A teacher of Informatics should help pupils organize their work in the internet, make information retrieval easier for them, shorten unproductive expenditures of time. Pupils must know that search systems of the internet are effective means of information retrieval and exploration.

2 Psychological Grounds of Studying the Theme "Search Systems of the Internet"

L.Vygotsky's cultural-historic theory of development of higher psychic functions is one of the leading in psycho-pedagogical science in Ukraine. It is widely used during the discussion of problems such as person and culture, studies and development, influence of information technologies on the development of children's psychics. According to the basic provisions of cultural-historic theory, signs and sign systems, which regularly get more complicated, contribute to the development and transformation of higher psychic functions [1].

The internet can be viewed as socio-technical system, which, together with modern digital technologies, is based on traditional sign systems and contributes to their quantitative complication and qualitative transformation. According to L.Vygotsky's views, such a socio-technical system, which regularly gets more complicated, promotes the development and transformation of higher psychic functions. That's why viewing the internet as modern means of activity which stimulates psychic development of a person, is quite grounded.

In spite of the variety of internet users' activity, researchers single out three main types of activity mediated by the Web: cognitive, communicative and playing [2]. We consider the information-retrieval activity as a constituent of a cognitive activity. Considering the internet as a sign system, working with it, and notably retrieving information from it contributes to the psychic and intellectual development of a child. Hence, we suppose it necessary to study internet search systems in the school course of Informatics.

Studying of internet search systems may take place on the basis of the theory of planned formation of intellectual actions, which is grounded by scientific works of P.Galperin and N.Talyzina. The theory of planned formation of intellectual actions supposes some successive stages.

At the first stage the basic motivation of the action is formed. Researchers have come to the conclusion that internet users' activity is poly-motivistic. Among the main motives are: cognitive, businesslike, communicative, recreational and playing, affiliative (the need to feel oneself as a member of some group), as well as motives of cooperation, self-fulfillment and self-assertion [3].

We think that the task of a Informatics teacher is to form pupils' cognitive motives of activity in the internet and to compare them with social ones. To form positive cognitive motivation of studying search systems of the internet, the teacher should inform pupils about the following facts:

- A qualified internet user may get access to scientific and literary sources of the leading libraries of the world;
- The majority of scientific, popular scientific and literary journals have electronic versions available via the internet, and some journals have only electronic version.;
- The number of different electronic encyclopedias, reference-books, dictionaries, thematic sites, present in the internet, regularly increases;
- Work with such types of resources is characterized by the availability of convenient and effective retrieval means.

Social motivation of the retrieval activity in the internet may be raised or formed by adducing the following facts:

- The retrieval activity is one of the leading while studying in the academy or institute of higher education, and for many in professional activity as well;
- The majority of institutes of higher educations in Ukraine and in other country have their own internet-sites, which offer data about faculties, specializations, terms of admission and so on;
- Some sites inform of possibilities and terms of studying the finding work abroad, studying the foreign language in its native country;
- Due to the information offered by the internet-sites, one may solve the problems of rest organizing, getting goods and so on.

At the next stage of planned formation of intellectual actions there the formation of the reference scheme of the action happens, i.e. the system of reference points and instructions. Considering of which is necessary to perform the action mastered with required qualities. The reference scheme of the action of information retrieval in the internet may be like the one in Figure 1.

The reference scheme of the action presented to pupils is polished up on the basis of training tasks.

Task 1. What is the height of Eiffel Tower? (The answer should be accompanied by a reference to the internet page).

To the query *Eiffel Tower* the search engine in the first results gives out documents that contain the answer. Three first operations of the reference scheme of the action are carried out.

Task 2. Which space project aims at studying the processes of the active Sun?

Solving of this task suggests introducing a sequence of query-specifications (for example, *space project* + *active Sun*) with the analysis of the results obtained. Four operations of the reference scheme of the action are carried out.

Task 3. To find documents which inform of the results of the regional phases of the all-Ukrainian contest “Young economy”. Take into account that the regional phase may be called the district one.

Solving of this task demands introducing the query *all-Ukrainian contest*, query-specification *Young economy* with the usage of the boolean operators *(regional | district) phase*. It’s necessary to show pupils that simple subsequent introducing of queries *regional phase* and *district phase* will lead to the exception from the area of retrieval of documents, which contain the words *district phase*, but don’t contain the words *regional phase*, which contradicts the condition of the task. All the operations of the reference scheme of the action are carried out.

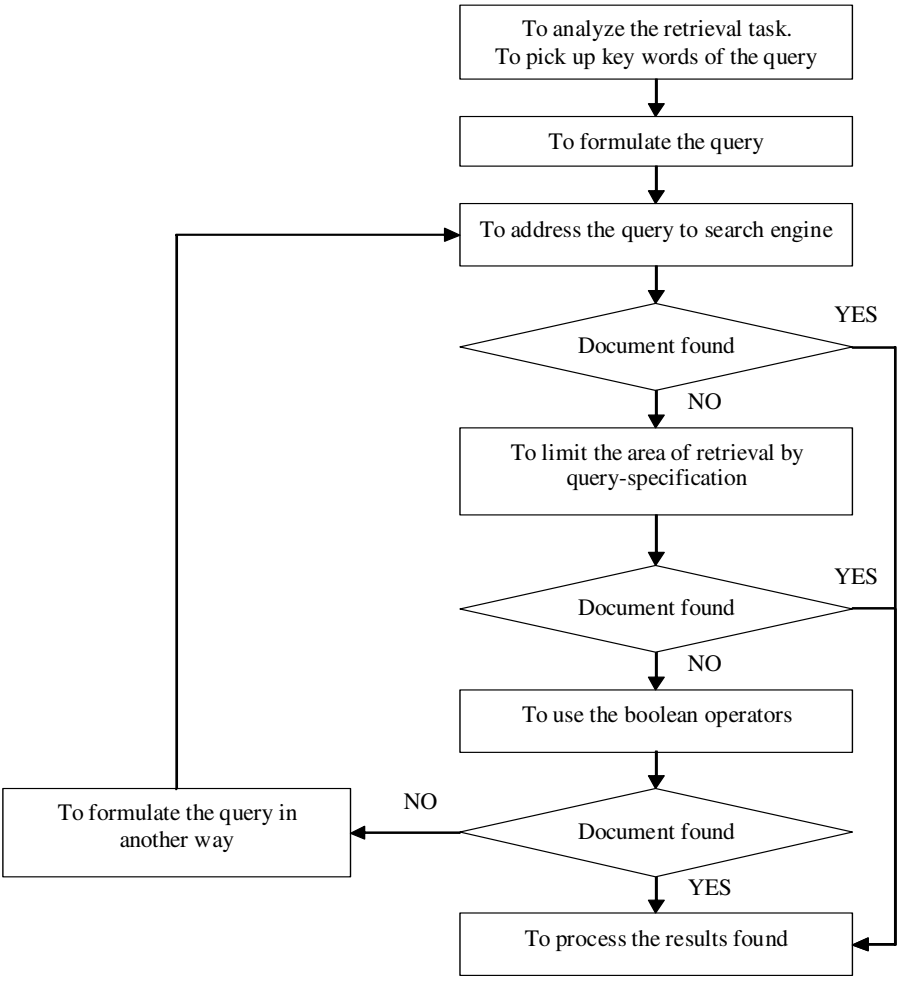

Fig. 1. The reference scheme of the action of retrieval in the internet

The criteria of formation of the action of retrieval in the internet is pupil's ability to find the information needed or to make sure that there is no such information in the Web.

3 Defining the Contents of "Search Systems of the Internet". Computer Teaching Program "Poshuk-META" (Retrieval Aim)

We suggest examining questions such as:

- Types of internet search systems (Subject Directories and Search Engines), their peculiar properties;
- The principles of functioning of search engines (the processes of scanning, indexing and ranging);

- The notions of information retrieval, relevance and boolean operators;
- The search engine on the example of which practical realization of the issues considered is carried out.

It's advisable to focus on mastering of possibilities of some search system: to define what information space it covers, to learn to formalize information needs by picking up key words and determining semantic links between them, to master boolean queries of the search system to analyze the results obtained. For Ukrainian pupils the system <META> may become such an object of studying.

Search system <META> (http://meta.ua) covers the whole Ukrainian segment of the internet, realizes retrieval according to the list of Ukrainian sites, gives a possibility of access to European information resources. <META> combines in itself the functions of subject directory and search engine, is characterized by high accuracy and speed of retrieval, processing documents by any European language with taking into account of morphology of Ukrainian and Russian. Studying of internet search systems is complicated by the fact that not all Ukrainian schools have an access to the internet. The authors have set a task to create a computer program of studying search systems without an internet-access. This was necessary :

- To present the theoretical material of the course studied in the hypertext-structure;
- To prepare examples for solving;
- To form the base of retrieval in accordance with pedagogical and educational objectives, to pick up documents, the theme of which would be interesting and urgent for pupils;
- To organize in this base a full-text retrieval of documents with the support of the boolean operators;
- To work out a system of retrieval task.

Such a program was created and it got the name "Poshuk-META" (Retrieval-Aim). Its interconnected constituents are: an electronic schoolbook, a system of examples for demonstration, a program-trainer.

The electronic school-book contains the theoretical material presented in the hypertext-structure. The systems Yahoo!, Google, Yandex (Russia) and <META> (Ukraine) are the objects of demonstration.

For creating the program-trainer from the internet there have been picked up web-documents (500 Mb), which have scientific, educational, cognitive direction. These documents were indexed by the employees of <META>, on their basis was organized a full-text retrieval with the support of the boolean operators of <META>. The program processes queries in Ukrainian, Russian, English. All the program is contained on CD.

For use of the program "Poshuk-META" there was worked out a system of laboratory works. Here are their themes and some tasks suggested for them.

Laboratory Work №1. Simple Retrieval.
Examples of retrieval tasks:

1. What is the aim of the project Space Interferometry mission?
2. What is the firm Bolt Beranek and Newman famous for?
3. What are the peculiarities of the car Honda Civic Hybrid?

The results of the retrieval must be entered into the following table:

Number of question	Formulation of the query	URL-address of the site where the answer is found	The answer to the question

Drawing up of the main results of laboratory works is done in the form of tables, which enable to summarize the things studied.

Laboratory Work №2. Boolean Operators of <META>.

Examples of retrieval tasks – to find documents, which contain certain key words, using boolean operators:

1. To find documents containing the word *school*.
2. To find documents containing the words *school* and *lyceum* at the same time.
3. To find documents containing at least one of the words *school* or *lyceum*.
4. To find documents which contain the word *lyceum* but don't contain the word *school*.

The results of the retrieval must be entered into the following table:

Number of question	Formulation of the query	Number of documents found

The analysis of the results obtained enables to follow the change in the number of the documents found with the usage of boolean operators.

Laboratory Work №3. Operators of Context Retrieval. Operators of Distance Between Words.

Task 1. With the help of the program "Poshuk-META", to find out who is the author of the given quotations:

1. The best mistake is the one made while studying.
2. Mathematics should be studied at least for the fact that it puts mind in order.

Task 2. To enter the suggested queries:

1. *physical culture, "physical culture";*
2. *computer games, "computer games".*

To analyze the numbers of documents found.

Task 3.

1. To find documents containing the words *internet, technology, access*.
2. To find documents containing the words *internet, technology, access*. The text fragment, where the key words occur, must not exceed 10 words.
3. To find documents containing the words *internet, technology, access*. The text fragment, where the key words occur, must not exceed 5 words.

To analyze the numbers of documents found.

Laboratory Work №4. Operators Limiting the Area of Retrieval by Certain Fields of Web-documents.

To find with the help of the program "Poshuk-META":

1. documents the titles of which contain the word *enactment*;
2. documents the titles of chapters of which contain the word *legislature*;
3. documents the titles of chapters of which contain the word *law*.

Concluding tasks:

1. To check if there is information of your native town in the retrieval base of the program "Poshuk-META"?
2. To check if there is information of the educational institution you'd like to study at after leaving school in the retrieval base of the program "Poshuk-META"?
3. To check if there is information about your favorite singer (music group, actor, football-player) in the retrieval base of the program "Poshuk-META"?

We suggest such time plan of work with this program: 2-3 lessons – for learning of theoretical material, and 4-7 lessons – for carrying out of the laboratory works. A number of lessons depends on the level of pupils' efficiency (for example, demons trations may be considered or not).

4 Results of the Pedagogical Experiment

The program "Poshuk-META" was introduced into the process of studies at some secondary schools in Ukraine. The introduction was accompanied by questioning the 284 pupils took part in the experiment.

The study has shown that: the majority of the people questioned considers the internet as an additional source of information (87%), but most respondents prefer entertaining kinds of resources (84%). Among other materials, pupils mentioned database of reports as absolute leaders in the list of interest. But electronic libraries are used only by 8% of the respondents. An important result of the survey is the fact that almost 98% of pupils think that the skills to realize information retrieval in the internet will be necessary for them in future studies and professional activity.

The question of search systems the pupils use was important for carrying out the experiment. Figure 2 shows the distribution of the rating of popularity of search systems before and after learning the theme "Studying of Information Search Systems of the internet".

Before studying the theme 29% of the pupils questioned couldn't name any information search system. After studying only 2% of pupils were ignorant. It should also be mentioned that after studying the theme the number of pupils who know and use in their work not one but some search systems got increased. It is underlined by the decrease of the percentage of popularity of Rambler and increase of the rating of search systems studied in the theme: <META> – from 5 to 22%, Yahoo! – from 6 to 9%, Yandex – from 11 to 12%, Google – from 9 to 16%. The data given show that most pupils prefer Russian search systems and Russian-language internet-resource.

The reasons for this are insufficient knowledge of English and underestimation of the development of the Ukrainian-language content of the internet. The authors set before them a task to attract pupils' attention to the Ukrainian-language segment of the internet

and to show the great potential of the English-language resource. As the data given show they managed to do it to a certain extent.

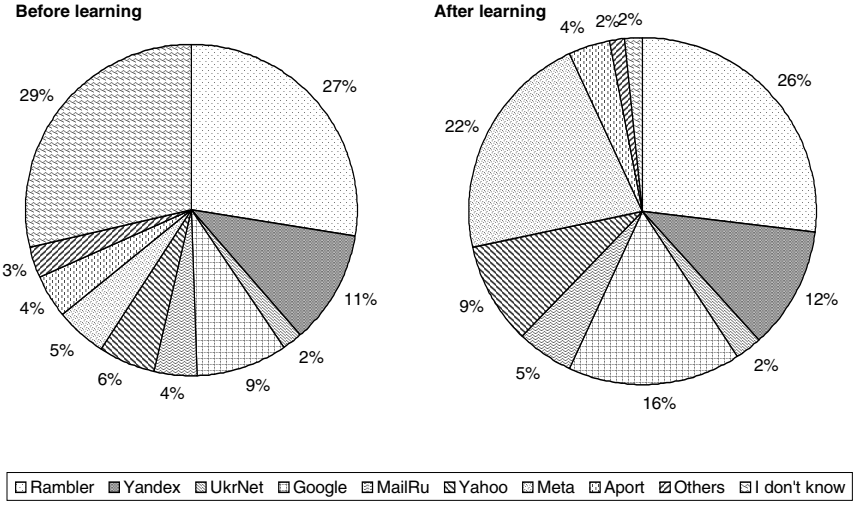

Fig. 2. Distribution of the rating of popularity of search systems

Application of the computer program "Poshuk-META" in the teaching process was effective and expedient.

5 Conclusion

The skills to appreciate information from the point of view of its authenticity, accuracy, adequacy for the solution of the concrete information task are formed either in the process of learning of internet search systems or in the process of the retrieval activity under the teacher's supervision. The skills to carry out information retrieval in the internet intensify pupils' cognitive activity, raise the level of their information culture.

References

1. Выготский Л.С. Собрание сочинений: В 6-ти т. – Т.3. – М.: Педагогика, 1983. – 368 с. (Vygotsky L.S. Complete works: In 6 vol. – Vol.3. – Moscow: Pedagogics, 1983. – 368 p.).
2. Бабаева Ю.Д., Войскунский А.Е., Смыслова О.В. Интернет: воздействие на личность // Гуманитарные исследования в Интернете / Под ред. А.Е. Войскунского. – М.: Можайск-Терра, 2000. – С. 11-39. (Babaeva Y.D., Voiskounsky A.E., Smyslova O.V. The internet: influence on a person // Humanitarian researches in the internet / Edited by A.E. Voiskounsky. – Moscow: Mozhaisk-Terra, 2000. – pp. 11-39.).
3. Арестова О.Н., Бабанин Л.Н., Войскунский А.Е. Мотивация пользователей Интернета // Гуманитарные исследования в Интернете / Под ред. А.Е. Войскунского. – М.: Можайск-Терра, 2000. - С. 55-76. (Arestova O.N., Babanin L.N., Voiskounsky A.E. Internet users' motivation // Humanitarian researches in the internet / Edited by A.E. Voiskounsky. – Moscow: Mozhaisk-Terra, 2000. – pp. 55-76.).