РЕЗУЛЬТАТИВНОСТЬ ИНФОРМАЦИОННОГО ПОИСКА В ЭЛЕКТРОННЫХ БИБЛИОТЕЧНЫХ СИСТЕМАХ (ЭБС)

Габеева Е.И.¹, Раднаева В.Д.²

*Восточно-Сибирский государственный университет технологий и
управления (ВСГУТУ), г. Улан-Удэ, Россия*

*Определена результативность информационного поиска в различных
электронных библиотечных системах (ЭБС).*

Ключевые слова: *результативность, информационный поиск, ключевые
слова.*

EFFICIENCY OF INFORMATION SEARCH IN ELECTRONIC LIBRARY SYSTEMS (ELS)

E.I. Gabeeva, V.D. Radnaeva

*East Siberia State University of Technology and Managment,
Ulan-Ude, Russia*

*The effectiveness of information retrieval in various electronic library sys-
tems (ELS) has been determined.*

Keywords: *performance, information search, keywords.*

Решение любой научной проблемы начинается с информационно-го поиска литературных источников. В недавнем прошлом исследова-тель много времени уделял работе в библиотеках. Причем результат работы зависел от фонда той или иной библиотеки, поэтому часто ис-следователь не был удовлетворен и вынужден был обращаться к биб-лиотекам, расположенным в других городах, пользоваться услугами межбиблиотечного абонементa (МБА). Не все зарубежные источники, особенно периодические издания, были доступны. В настоящее время

¹ Габеева Елена Игоревна, студент 2 курса направления подготовки «Технология изделий легкой промышленности» факультета экологии, сервиса, технологий и дизайна, gabeevaelena8@gmail.com

Gabeeva E.I., 1st year student of the direction of training "Technology of Light Industry Products».

² Раднаева Вера Дашиевна, д.т.н., доц. каф. «Технология кожи, меха. Водные ресурсы и товароведение», e-mail: radnaevav@mail.ru

Radnaeva Vera D., Dr. of technical science, Ass. Prof. of Department «Technology of Leather, Fur. Water Resources and Commodity Research».

с развитием интернета процесс информационного поиска изменился. Исследователю стали доступны информационные источники не только ведущих библиотек России, но и других стран. Вместе с тем возникла необходимость выбора информационно-поисковых систем (ИПС), корректной формулировки ключевых слов, от которых зависит результативность информационного поиска. По данным [1] по популярности ИПС первое место занял Яндекс, который выбрали все опрошенные респонденты, на втором месте – Google 75% на третьем – Rambler 50 %, на четвёртом – Aport 45%, на 5 и 6 – AltaVista и Yahoo с 20 % каждый, на 7 и 8 – AllTheWeb и Opentext с 5% каждый. По средней оценке, качества поиска места распределились следующим образом Google - 8.6, Яндекс -8.05, Rambler – 7.2, Yahoo – 7, AltaVista – 6.75, Aport – 6.66, AllTheWeb- 5, и OpenText - 4.

Целью работы является оценка результативности различных информационно-поисковых систем в зависимости от набора ключевых слов.

Объекты и методы исследования

Основными объектами исследования являлись ИПС электронной библиотеки e-library, Российской государственной библиотеки (РГБ), Rucont, Лань, Яндекс, Google. Яндекс и Google были выбраны, как наиболее популярные ИПС. Руcont и Лань входят в информационную базу данных университета. Из библиотек выбраны РГБ и elibrary, в которых сосредоточен большой фонд технической литературы, включая периодические издания и материалы конференций.

Результативность ИПС определяли по релевантности. Релевантность – это обозначение субъективной степени соответствия чего-либо в моменте времени [2,3] и является в некоторой степени субъективным понятием. В критерии поиска, как правило, используются два основных показателя *recall* — полнота и *precision* — точность. В совокупности они обозначают значимость соответствия данного запроса, что и является понятием релевантности.

Проводили поиск по теме, посвященной изучению химического состава дыма. Дым является типичным аэрозолем, используемым в пищевой промышленности, в частности, для обработки рыбы и мясных продуктов. Причем известен так называемый «технологический» дым, обладающий определенными физико-химическими характеристиками [4]. В меховой промышленности известна традиционная обработка дымом меховых шкур с целью придания ей определенных

свойств: устойчивости к факторам внешней среды и окрашиванию кожной ткани.

Целью поиска являлись результаты исследований по определению химического состава дыма. В соответствии с темой работы изменяли набор ключевых слов, сформулированных на русском и английском языках, и выполняли поиск литературных источников в ИПС. Определяли релевантность по формуле:

$$R = \frac{m}{n} \cdot 100\% \quad (1)$$

где m – количество релевантных источников;
 n – количество результатов поиска.

Результаты и обсуждение

Для определения результативности информационно-поисковых систем по теме исследования изменяли ключевые слова, в которых использованы следующие слова: химический состав дыма, характеристика дыма, продукты, древесина, древесный дым, burning, chemical composition, products. Полноту и точность определяли на основе известной информации и предыдущих знаний.

Были проведены поиски в различных электронно-библиотечных системах, результаты которых приведены в таблице 1. Как видно из таблицы 1, количество результатов запросов в поисковые системы Яндекс и Google находится в пределах от 0,395 до 71,5 млн в зависимости от набора ключевых слов. Причем количество результатов запроса в Google существенно превышает этот показатель в Яндекс. Однако релевантных источников, удовлетворяющих запросу пользователя, очень мало – 1%. Релевантность Google и Яндекс рассчитывалась не на все результаты, а только на первые 100, так как физически невозможно просмотреть все. Для снижения количества результатов в этих ИПС необходимо увеличивать количество ключевых слов, конкретизирующих тему.

На третьем месте по количеству результатов запросов стоит Российская государственная библиотека и электронная библиотека elibrary – максимальное количество результатов равно соответственно 5961 и 4902. Данные таблицы 2 показывают, что по релевантности РГБ уступает электронной библиотеке – максимальная релевантность равна соответственно 0,13 и 12,5%.

Поиск в электронно-библиотечной системе Rucont не принес большой результативности. Электронно-библиотечная система Лань показала наименьшую результативность. Вероятно, это можно объяснить различным объемом фондов, а также качественным составом этих ИПС.

Таблица 1 – Влияние качественного и количественного состава ключевых слов на результаты информационного поиска ИПС Яндекс и Google

№	Запрос	Google		Яндекс	
		млн	Релевантность, %	млн	Релевантность, %
1	Химический состав, горение, продукты	0,36	0	9	0
2	Burning, chemical composition, products	71,50	1	5	1
3	Горение, продукты горения	3,26	0	5	0
4	Химический состав, продукты горения	2,71	0	10	0
5	Chemical composition of combustion products	45,1	0	4	0
6	Состав горения древесины	0,52	0	6	0
7	Состав дыма древесины	0,72	0	4	0
8	Характеристика дыма древесины	0,69	1	8	0
9	Химический состав древесного дыма	0,40	0	5	0

Таблица 2 – Релевантность информационно-поисковых систем в зависимости от состава ключевых слов

№	Ключевые слова	Релевантность, %					
		Elibrary	РГБ	Rucont	Лань	Google*	Яндекс*
1	Химический состав, горение, продукты	1,11	0,05	0,02	0,3	0	0
2	Burning, chemical composition, products	1,46	0	0	0	1,00	1,00
3	Горение, продукты горения	0,04	0,03	0	0	0	0
4	Химический состав, продукты горения	1,10	0,05	0,02	0,30	0	0
5	Chemical composition of combustion products	0,38	0	0,97	0	0	0
6	Состав горения древесины	2,20	0,13	0	1,33	0	0
7	Состав дыма древесины	12,50	0	0	0	0	0
8	Характеристика дыма древесины	25,00	0	0	0	1	0
9	Химический состав древесного дыма	25,00	0	0	0	0	0

Анализ данных таблицы 2 показал, что релевантность ИПС зависит от состава ключевых слов. Так, наибольшая релевантность 12,5% наблюдалась при наборе ключевых слов: химический состав дыма древесины (25%), характеристика дыма древесины (25%) и состав дыма древесина (12,5%). Поиск также показал значимость набора ключевых слов на разных языках. Так, при наборе ключевых слов на английском языке результаты поиска показали наличие требуемых источников на английском языке.

Заключение

Выполнена оценка результативности информационного поиска в различных электронно-библиотечных системах. Установлено, что наибольшей результативностью информационного поиска обладает электронно-библиотечная система eLibrary. в данной ЭБС предоставлен лучший расширенный поиск из представленных библиотечных систем. Показана значимость влияния языка набора ключевых слов и его состава на результаты информационного поиска.

Список использованных источников

1. М.В.Козлов, В.Я.Яцко. Метод оценки эффективности функционирования современных информационно-поисковых систем интернета. URL: <http://www.dialog-21.ru/digests/dialog2006/materials/html/Kozlov.htm> (дата обращения 02.10.2020).
2. Релевантность. URL: <https://ru.wikipedia.org/>
3. А.Н.Целых, М.Н.Котов. Оценка эффективности информационного поиска. URL: <https://cyberleninka.ru/article/n/otsenka-effektivnosti-informatsionnogo-poiska/viewer> (дата обращения 02.10.2020).
4. Химический состав дыма. URL: <https://mutnovka.livejournal.com/24189.html> (дата обращения 02.10.2020).

References

1. M.V. Kozlov, V.Ya. Yatsko. A Method for Assessing the Effectiveness of the Functioning of Modern Information Retrieval Systems of the Internet. URL: <http://www.dialog-21.ru/digests/dialog2006/materials/html/Kozlov.htm> (date of treatment 10/02/2020).
2. Relevance. URL: <https://ru.wikipedia.org/>
3. A. N. Tselykh, M. N. Kotov. Evaluation of the Effectiveness of Information retrieval. URL: <https://cyberleninka.ru/article/n/otsenka-effektivnosti-informatsionnogo-poiska/viewer> (date of treatment 10/02/2020).
4. The Chemical Composition of the Smoke. URL: <https://mutnovka.livejournal.com/24189.html> (date of treatment 10/02/2020).