

ЭФФЕКТИВНОСТЬ ПОИСКА В ИНТЕРНЕТЕ УЧЕБНОГО МАТЕРИАЛА ПО МУЗЫКАЛЬНОЙ ИНФОРМАТИКЕ¹

EFFECTIVENESS OF SEARCHING FOR MUSIC INFORMATICS EDUCATIONAL MATERIALS ON THE INTERNET

Белоозеров В.Н.

Зав. сектором ВИНТИ РАН,
доцент Московского государственного
университета культуры и искусств,
кандидат филологических наук
E-mail: informling@narod.ru

Аннотация. В статье описан эксперимент по оценке эффективности поиска в Интернете учебного материала по курсу «Введение в информатику» студентами музыкальных специальностей Московского государственного университета культуры и искусств. В качестве инструмента поиска взяты поисковые системы Яндекс и Гугл. Их эффективность оценивалась классическим показателем точности T и вновь введённым показателем достаточности поиска D. Показатель D определялся по степени удовлетворенности пользователя результатами поиска. Эксперимент свидетельствует о пригодности показателей T и D для численной оценки и сравнения эффективности поисковых систем Интернета.

Beloozerov V.N.

Head of the VINITI sector of the Russian Academy of Sciences, Associate Professor at the Moscow State University for Culture and Arts, Candidate of science (Philology).
E-mail: informling@narod.ru

Annotation. The article describes an experiment to assess the effectiveness of Internet search for «Introduction to information science» course training material by students of musical specialties at the Moscow State University for Culture and Arts. Yandex and Google search engines were taken as the search tools. Their effectiveness was assessed by the classic accuracy rate T and the newly introduced adequacy search ratio D. Figure D was determined by the degree of user satisfaction of the search results. The experiment demonstrates the suitability of T and D indicators for numerical evaluation of Internet search engines efficiency.

¹ В обработке экспериментальных данных принимали участие студенты Московского государственного университета культуры и искусств: Мария Баранова (Институт музыки МГУКИ), Полина Терехова и Арина Овечкина (Институт социально-культурной деятельности МГУКИ).

Ключевые слова: эффективность информационного поиска, поисковые системы Интернета, точность поиска, достаточность поиска, Яндекс, Гугл, введение в информатику, музыкальная информатика

Keywords: information search effectiveness, Internet search engines, search accuracy, search sufficiency, Yandex, Google, introduction to informatics, music informatics.

Проникновение сетевых информационных технологий в образовательный процесс ставит на повестку дня задачу оценки эффективности и целесообразности использования тех или иных инструментов. В частности, среди студентов распространена тенденция искать сведения, необходимые для выполнения зачётных заданий, не в учебниках, а в Интернете с помощью той или иной общепользовательской поисковой системы. При этом зачастую необходимые сведения удаётся получить быстрее и легче, чем путём изучения учебных пособий в книжной форме или освоения специализированных сетевых ресурсов. Для объективной оценки и сравнения различных методов нахождения учебных материалов требуется разработать методику такой оценки. В настоящей работе предлагается способ определения эффективности поисковых систем Интернета на основе показателей **точности** и **достаточности** поиска.

Согласно классической теории информационного поиска [1, 2, 3] эффективность поисковых систем определяется двумя показателями – **точностью** (T) и **полнотой** (Π). Определение этих величин основано на усреднении результатов поиска в массиве документов по представительному набору типичных запросов при определённых условиях. Показатель точности равен отношению числа найденных релевантных документов (соответствующих запросу) (Nr) к общему числу выданных документов ($Nout$):

$$T = Nr/Nout.$$

Показатель полноты равен отношению числа найденных релевантных документов (Nr) к общему числу релевантных документов, имеющихся в поисковом массиве (Mr):

$$\Pi = Nr/Mr.$$

Достаточно удачным признаком хорошей работы поисковой системы оказался показатель суммарной эффективности $\mathcal{E} = T + \Pi$, который должен быть больше или не на много меньше единицы ($\mathcal{E} \geq 1$), чтобы систему поиска можно было признать достаточно эффективной.

Данные критерии доказали свою полезность для систем, работающих на контролируемых массивах документов, какими являются книжные коллекции библиотек и фонды информационных центров. В названных условиях применение этих показателей не сопряжено с принципиальными сложностями. При этом многие закономерности, приемы анализа и поиска информации, установленные для библиотечных информационных массивов,

с успехом могут быть применены и для совокупности информационных ресурсов, доступных в Интернете (см., напр., [4]). Однако определение показателя полноты P для поиска в Интернете требует оценки общего количества доступных ресурсов и общего количества имеющихся в них пертинентных документов (удовлетворяющих информационную потребность), что связано с очевидными методическими и практическими затруднениями.

Имеется и другое основание не включать полноту P в число основных показателей эффективности при анализе поисковых систем Интернета. Использование этого показателя подразумевает, что идеальным результатом поиска была бы выдача ВСЕХ пертинентных документов ($Nr = Mr$). Но при поиске в Интернете такая задача никогда не ставится (в отличие от поиска в библиотечных коллекциях). Что же касается показателя точности T , то его смысл и способ определения при поиске в Интернете сохраняется в том же виде, как при поиске в библиотечных массивах. Однако одного этого показателя для оценки эффективности недостаточно. Может существовать поисковая система с очень низкой точностью, которая сразу выдаёт документ, полностью удовлетворяющий пользователя. Возможны также ситуации, когда выдача системы будет содержать только релевантные документы ($T = 1 = 100\%$ – максимальное значение), которые при этом дают крайне мало необходимых сведений.

Поиск в Интернете идёт до тех пор, пока пользователь не получит ДОСТАТОЧНУЮ для удовлетворения его потребности информацию либо пока не убедится, что это невозможно. Качество системы определяется тем, как быстро пользователь найдёт нужные ему сведения. Это соображение послужило основанием введения в работе [5] оценки качества систем по показателю глубины пользовательского поиска, который определялся как сумма номера первого пертинентного документа в выдаче поисковой системы и числа переходов со страницы на страницу внутри документа. Окончательная оценка эффективности вычислялась как сумма баллов глубины первых десяти ссылок в выдаче системы с учётом ряда условий поиска нужных сведений. На наш взгляд этот метод имеет право на использование, но он излишне усложнён, а поставленный в работе эксперимент также усложнён формальными процедурами, не свойственными естественному поиску сведений.

Для учёта именно удовлетворённости пользователя работой системы мы предлагаем использовать показатель **достаточность** поиска (D).

Достаточность определяется как доля (выраженная в процентах) от того объёма сведений, который мы надеялись получить при обращении к системе. Например, когда при запросе «Расписание самолётов в аэропортах Москвы» мы находим расписание по аэропорту Шереметьево, наша потребность удовлетворена только на треть и достаточность такого поиска составляет 33%. Но когда мы при дальнейшем просмотре выдачи находим также расписания по аэропортам Домодедово и Внуково, удовлетворённость становится полной и $D = 100\%$. Таким образом, достаточность зависит не только от характеристик системы поиска, но также от психологической установки клиента. Поэтому может показаться, что этот критерий (в отличие от точности T и полноты P) менее объективен. Однако нужно вспомнить, что и классические показатели T и P определяются в зависимости от оценки клиентом полезности (пертинентности) того или иного документа в выдаче. Во всех случаях оценка эффективности поиска информации зависит от степени психологической удов-

летворённости пользователя системы. В этом отношении вводимый критерий D не имеет принципиальных отличий от критериев T и P .

По примеру классического набора показателей мы будем рассматривать также суммарный показатель эффективности $\Sigma = T + D$.

Чтобы исследовать практическую применимость критериев T и D для оценки эффективности поиска в Интернете, в течение 2011–2012 гг. был проведён эксперимент с участием в качестве информантов студентов IV курса Института музыки Московского государственного университета культуры и искусств (МГУКИ), проходивших обучение по курсу «Введение в информатику». Всем студентам предлагалась типовая домашняя контрольная работа «Исследование эффективности поиска в Интернете сведений по теме...». Тема поиска из области музыкальной информатики у каждого была своя и назначалась по согласованию преподавателя со студентом. По выбранной теме студент составлял простой словесный запрос к поисковым системам Интернета в обычной форме названия темы, например «Форматы музыкальных электронных изданий», «Аналоговая запись музыки», «Электронные музыкальные издания» и т. п.

Эта формулировка подавалась для поиска в две поисковые системы по выбору студента (Яндекс и Рамблер, АльтаВиста и Гугл, Рамблер и АльтаВиста и т. п.)². Студент последовательно просматривал подряд все документы, выданные машинами на первых страницах выдачи, и выписывал адреса тех документов, которые, на его взгляд, содержат сведения, полезные при сдаче экзамена по заданной теме. Для каждой поисковой системы подсчитывалось число полезных документов в каждом из четырёх первых десятков выдач. Результаты подсчётов сводили в таблицу (см. таблицу 1 на с. 150).

В столбцы таблицы «Число полезных» заносилось количество полезных документов в каждом десятке для каждой поисковой системы.

В столбцы «Точность поиска» записывалась доля найденных полезных документов в данном десятке плюс все предыдущие.

В столбцы «Достаточность поиска» заносили оценку того, на сколько процентов студент был удовлетворен найденными сведениями в соответствии с его представлением о требованиях к отличному ответу на экзамене по данному вопросу и с учетом всех найденных к данному моменту сведений – на этой и предыдущих страницах.

В столбцы «Сумма $T + D$ » заносят сумму процентов точности (T) и достаточности (D).

Данные студентов обсуждались с преподавателем. При этом студент делал доклад по существу найденных сведений, а в его письменном отчёте исправлялись арифметические и логические ошибки.

Результаты одного такого эксперимента показаны в таблице 1. Данные получены по выдачам поисковых систем Яндекс и Гугл в ответ на запрос «*обработка музыкальной информации*». Из таблицы следует, что поиск в системе Яндекс в целом оказался более эффективен, поскольку она выдала больше полезных (пертинентных) ссылок и более полно

² В настоящей статье мы рассматриваем только результаты, полученные по системам Яндекс и Гугл. Статистика по другим системам недостаточно представительна.

Таблица 1

Показатели поиска в зависимости от глубины просмотра выдачи

| Количество просмотренных ссылок | Показатели эффективности, % | | | | | | | |
|---------------------------------|-----------------------------|----|----|-------|----------------|----|----|-------|
| | Яндекс | | | | Google | | | |
| | Число полезных | Т | Д | Т + Д | Число полезных | Т | Д | Т + Д |
| 10 | 1 | 10 | 10 | 20 | 5 | 50 | 50 | 100 |
| 20 | 4 | 40 | 60 | 100 | 2 | 35 | 65 | 100 |
| 30 | 4 | 30 | 70 | 100 | 0 | 23 | 65 | 88 |
| 40 | 1 | 25 | 80 | 105 | 0 | 18 | 65 | 83 |
| Всего полезных ссылок | 10 | | | | 7 | | | |
| Из них совпадают | 1 | | | | | | | |

удовлетворила информационную потребность. Хотя в пределах первой страницы выдач преимущество было на стороне Гугла, который позволил получить 50% необходимой информации быстрее.

Полезные сведения были обнаружены по следующим ссылкам поисковых систем, приведённым в таблицах 2 и 3 (см. с. 151). Всего обеими системами было найдено 16 полезных ссылок, из которых одна ссылка оказалась общей. Незначительное число пересечений между выдачами разных систем свидетельствует о том, что они пользуются различными критериями поиска.

Для статистического анализа эффективности системы Яндекс при поиске по разным запросам составлены сводные таблицы результатов, полученных студентами Института музыки МГУКИ в разные годы. При этом вычислены средние значения показателей эффективности по десяткам выдач и сделаны выводы о возможности оценки эффективности поиска в Интернете по предлагаемым критериям. Все поисковые запросы относились к области музыкальной информатики.

На таблицах 4 и 5 (см. с. 152) представлены данные по эффективности поисков, проведённых в 2012 и 2011 гг. На таблице 6 (см. с. 154) представлены данные по поискам в системе Гугл, проведённым в 2012 г. одновременно с поисками в Яндексе. Поиск проводили те же студенты по тем же запросам.

В таблицах в каждой строке приведены результаты поиска одного студента. Первая клетка строки содержит условный идентификатор студента, вторая – формулировку запроса, по которому проводился поиск. Далее приведены показатели эффективности по четырём вариантам глубины просмотра выдачи: анализ одной страницы (10 ссылок), двух страниц (20 ссылок), трёх страниц (30 ссылок) и четырёх страниц (40 ссылок). В каждой клетке указывается значение показателя (точность, достаточность и их сумма) и че-

Таблица 2

Полезные ссылки, выданные системой Яндекс

| Адрес ссылки | Порядковый номер в выдаче | Д, % |
|---|---------------------------|------|
| http://www.superinf.ru/view_helpstud.php?id=4140 | 3 | 10 |
| http://weldingshop.3dn.ru/publ/obrazovatelnye/studentam/tekhnologii_i_sredstva_obrabotki_zvukovoj_informacii/8-1-0-101 | 11 | |
| http://www.bibliofond.ru/view.aspx?id=528900 | 18 | |
| http://the-programmer.ru/publ/informatika/informatika/osnovnye_trebovanija_k_kachestvu_informacii/13-1-0-31 | 19 | |
| http://www.0zd.ru/programmirovanie_kompyutery_i/programmy_dlya_obrabotki_zvukovoj.html | 20 | 60 |
| http://library.by/portalus/modules/culture/referat_readme.php?subaction=showfull&id=1168426343&archive=&start_from=&ucat=& | 22 | |
| http://mmagaa.narod.ru/vop/4.html | 23 | |
| http://rudocs.exdat.com/docs/index-568736.html?page=5 | 25 | |
| http://reformaobrazovaniya.blogspot.ru/2012/10/blog-post_28.html | 30 | 70 |
| http://www.master-skills.ru/articles/mixing/734-amb | 40 | 80 |

Таблица 3

Полезные ссылки, выданные системой Гугл

| Адрес ссылки | Порядковый номер в выдаче | Д, % |
|---|---------------------------|------|
| http://rudocs.exdat.com/docs/index-374548.html?page=14 | 1 | |
| http://ru.wikibooks.org/wiki/%D0%92%D0%B8%D0%B4%D1%8B_%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D0%B8_%D0%B8_%D0%B5%D1%91_%D1%81%D0%B2%D0%BE%D0%B9%D1%81%D1%82%D0%B2%D0%B0 | 3 | |
| http://distant.ioso.ru/do/curse/informat/inf10.htm | 4 | |
| http://lib.podelise.ru/docs/50940/index-2019.html | 5 | |
| http://weldingshop.3dn.ru/publ/obrazovatelnye/studentam/tekhnologii_i_sredstva_obrabotki_zvukovoj_informacii/8-1-0-101 | 7 | 50 |
| http://ru.wikipedia.org/wiki/%D0%9E%D0%B1%D1%80%D0%B0%D0%B1%D0%BE%D1%82%D0%BA%D0%B0_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85 | 18 | |
| http://daxnow.narod.ru/index/0-51 | 20 | 65 |

Таблица 4

Статистика поиска в системе Яндекс за 2012 г.

| Студент | Запрос (формулировка) | Показатели эффективности по десяткам ссылок в выдаче | | | | | | | | | | | | | | | | | |
|---|--|--|----|--------------------|-----|--------------------|----|----------------------|-----|--------------------|-----|--------------------|-----|----|----|----|----|------------------|--|
| | | Точность (Т), % | | | | | | Достаточность (Д), % | | | | | | | | | | | |
| | | Просмотрено ссылок | | Просмотрено ссылок | | Просмотрено ссылок | | Просмотрено ссылок | | Просмотрено ссылок | | Просмотрено ссылок | | | | | | | |
| | | 10 | 20 | 30 | 40 | 10 | 20 | 30 | 40 | 10 | 20 | 30 | 40 | 10 | 20 | 30 | 40 | $\Sigma = T + Д$ | |
| A12н | Форматы музыкальных электронных изданий | 60 | 50 | 45 | 37 | 70 | 73 | 75 | 80 | 130 | 123 | 120 | 117 | | | | | | |
| B12 | Полнота и точность поиска музыкальных публикаций | 60 | 50 | 40 | 30 | 35 | 60 | 75 | 80 | 95 | 110 | 115 | 118 | | | | | | |
| Д12 | Компьютерные технологии музыки | 70 | 70 | 77 | 67 | 70 | 70 | 90 | 95 | 140 | 149 | 167 | 162 | | | | | | |
| К12 | Законопроект о культуре | 90 | 75 | 63 | 55 | 70 | 80 | 85 | 90 | 160 | 155 | 148 | 145 | | | | | | |
| Ла12 | Международный стандартный музыкальный номер | 60 | - | - | - | 60 | - | - | - | 120 | - | - | - | | | | | | |
| Лю12 | Аналоговая запись музыки | 80 | 60 | 53 | 43 | 90 | 95 | 98 | 98 | 170 | 155 | 151 | 141 | | | | | | |
| Лы12 | Цифровая запись музыки | 40 | 50 | 47 | 38 | 40 | 60 | 65 | 100 | 80 | 110 | 112 | 138 | | | | | | |
| М12 | Цифровая запись музыки | 90 | 65 | 53 | 43 | 70 | 90 | 100 | 100 | 160 | 155 | 153 | 143 | | | | | | |
| Пана12 | Аналоговая запись музыки | 20 | 25 | 27 | 23 | 30 | 80 | 80 | 80 | 50 | 105 | 107 | 103 | | | | | | |
| Панч12 | Нотный текст как источник информации | 20 | 25 | 23 | 23 | 25 | 45 | 55 | 65 | 45 | 70 | 78 | 87 | | | | | | |
| Пас12 | Обработка музыкальной информации | 20 | 10 | 10 | 7,5 | 50 | 50 | 70 | 70 | 70 | 60 | 80 | 78 | | | | | | |
| X12 | Форматы музыкальных записей | 40 | 25 | 17 | 15 | 80 | 85 | 85 | 85 | 120 | 110 | 101 | 100 | | | | | | |
| Шабал12 | Аналоговая запись музыки | 20 | 25 | 23 | 22 | 25 | 45 | 55 | 65 | 45 | 70 | 78 | 87 | | | | | | |
| Шабан12 | Средства тематического поиска музыкальной информации | 20 | 15 | 17 | 13 | 20 | 50 | 70 | 70 | 40 | 55 | 87 | 83 | | | | | | |
| Ше12 | Электронные музыкальные издания | 40 | 30 | 20 | 10 | 80 | 87 | 90 | 99 | 120 | 117 | 110 | 109 | | | | | | |
| Средние показатели | | 49 | 41 | 37 | 30 | 54 | 69 | 78 | 84 | 103 | 110 | 115 | 115 | | | | | | |
| Среднее отклонение от среднего значения | | 23 | 19 | 17 | 14 | 21 | 14 | 12 | 11 | 39 | 28 | 24 | 21 | | | | | | |

Таблица 5

Статистика поиска в системе Яндекс за 2011 г.

| Студент | Запрос (формулировка) | Показатели эффективности по десяткам ссылок в выдаче | | | | | | | | | | | | | | | | | |
|---|---|--|----|----|------|-----|----------------------|-----|-----|-----|-----|--------------------|------|------|----|-----|-----|-----|------|
| | | Точность (Т), % | | | | | Достаточность (Д), % | | | | | $\Sigma = Т + Д$ | | | | | | | |
| | | Просмотрено ссылок | | | | | Просмотрено ссылок | | | | | Просмотрено ссылок | | | | | | | |
| 10 | 20 | 30 | 40 | 50 | 10 | 20 | 30 | 40 | 50 | 10 | 20 | 30 | 40 | 50 | 10 | 20 | 30 | 40 | |
| Ба11 | Форматы аудиофайлов | 50 | 55 | 46 | 42 | 50 | 60 | 70 | 80 | 100 | 100 | 115 | 116 | 122 | | 100 | 115 | 116 | 122 |
| Бе11 | Информационный аспект музыки | 10 | 15 | 10 | 7,5 | 30 | 45 | 45 | 45 | 40 | 60 | 60 | 55 | 52,5 | | 40 | 60 | 55 | 52,5 |
| Г11 | | 40 | 30 | 20 | 15 | 70 | 80 | 80 | 80 | 110 | 110 | 110 | 100 | 95 | | 110 | 110 | 100 | 95 |
| Ж11 | Цифровая запись музыки | 70 | 45 | 47 | 35 | 100 | 100 | 100 | 100 | 170 | 145 | 147 | 135 | | | 170 | 145 | 147 | 135 |
| И11 | Аналоговая запись музыки | 90 | 65 | 53 | 43 | 70 | 90 | 100 | 100 | 160 | 155 | 153 | 143 | | | 160 | 155 | 153 | 143 |
| К11 | Виды музыкальных изданий | 10 | 10 | 10 | 7,5 | 20 | 20 | 20 | 20 | 30 | 30 | 30 | 27,5 | | | 30 | 30 | 30 | 27,5 |
| М11 | Соотношение понятий релевантность и пертинентность музыкального документа | 80 | 75 | 73 | 65 | 100 | 100 | 100 | 100 | 180 | 175 | 173 | 165 | | | 180 | 175 | 173 | 165 |
| Па11 | Соотношение полноты и точности информационного поиска в музыке | 30 | 20 | 13 | 12,5 | 30 | 60 | 60 | 70 | 60 | 80 | 73 | 82,5 | | | 60 | 80 | 73 | 82,5 |
| Пл11 | Аудиоформаты | 70 | 45 | 37 | 30 | - | - | - | - | - | - | - | - | | | - | - | - | - |
| Р11 | Компьютерные вирусы в музыкальной информации | 0 | 0 | 13 | 18 | - | - | - | - | - | - | - | - | | | - | - | - | - |
| Х11 | Виды музыкальных произведений печати | 10 | 10 | 10 | - | 30 | 60 | 70 | - | 40 | 70 | 80 | - | | | 40 | 70 | 80 | - |
| Средние показатели | | 42 | 34 | 30 | 28 | 56 | 68 | 72 | 74 | 99 | 104 | 103 | 103 | | | 99 | 104 | 103 | 103 |
| Среднее отклонение от среднего значения | | 27 | 21 | 19 | 15 | 26 | 21 | 21 | 22 | 50 | 40 | 39 | 38 | | | 50 | 40 | 39 | 38 |

Таблица 6

Статистика поиска в системе Гугл за 2012 г.

| Студент | Запрос (формулировка) | Показатели эффективности по десяткам ссылок в выдаче | | | | | | | | | | | | | |
|---|--|--|------|----|----|----------------------|----|-----|-----|--------------------|------|-----|-----|----|----|
| | | Точность (Т), % | | | | Достигаемость (Д), % | | | | $\Sigma = T + Д$ | | | | | |
| | | Просмотрено ссылок | | | 10 | Просмотрено ссылок | | | 10 | Просмотрено ссылок | | | 10 | | |
| В12 | Полнота и точность поиска музыкальных публикаций | 20 | 30 | 40 | 10 | 20 | 30 | 40 | 10 | 20 | 30 | 10 | 20 | 30 | 40 |
| Д12 | | 40 | 37 | 33 | 25 | 45 | 55 | 75 | 751 | 85 | 93 | 108 | | | |
| Ла12 | Международный стандартный музыкальный номер | 60 | 46 | 40 | 60 | 60 | 65 | 70 | 120 | 120 | 111 | 110 | | | |
| Ло12 | | 90 | - | - | 60 | - | - | 150 | - | - | - | - | | | |
| Ль12 | Аналоговая запись музыки | 40 | 20 | 13 | 10 | 40 | 40 | 40 | 80 | 60 | 53 | 50 | | | |
| Шабал12 | | 50 | 30 | 23 | 20 | 50 | 60 | 70 | 100 | 90 | 93 | 100 | | | |
| Шабан12 | Средства тематического поиска музыкальной информации | 30 | 40 | 40 | 35 | 65 | 75 | 85 | 65 | 105 | 115 | 120 | | | |
| Средние показатели | | 40 | 35 | 33 | 30 | 60 | 80 | 90 | 70 | 95 | 113 | 120 | | | |
| Среднее отклонение от среднего значения | | 51 | 37,5 | 32 | 28 | 43 | 55 | 64 | 73 | 94 | 92,5 | 96 | 101 | | |
| | | 13 | 9 | 9 | 9 | 16 | 8 | 11 | 13 | 24 | 14 | 17 | 16 | | |

Таблица 7

Средние значения показателей эффективности

| Запрос (формулировка) | Показатели эффективности по десяткам ссылок в выдаче | | | | | | | | | | | |
|-----------------------|--|----------------|----------------|----------------|----------------------|----------------|----------------|----------------|--------------------|----------------|----------------|----------------|
| | Точность (Т), % | | | | Достигаемость (Д), % | | | | $\Sigma = T + Д$ | | | |
| | Просмотрено ссылок | | | 10 | Просмотрено ссылок | | | 10 | Просмотрено ссылок | | | 10 |
| Яндекс 2011 г. | 42/27 | 34/21 | 30/19 | 28/15 | 56/26 | 68/21 | 72/21 | 74/22 | 99/50 | 104/40 | 103/39 | 103/38 |
| Яндекс 2012 г. | 48/22 | 39/20 | 35/17 | 29/16 | 55/20 | 69/14 | 77/11 | 83/11 | 102/38 | 108/27 | 113/23 | 111/21 |
| Гугл 2012 г. | 51/13 | 37,5/9 | 32/9 | 28/9 | 43/16 | 55/8 | 64/11 | 73/13 | 94/24 | 92,5/14 | 96/17 | 101/16 |
| Преимущество имеет: | Гугл 2012 г. | Яндекс 2012 г. | Яндекс 2012 г. | Яндекс 2012 г. | Яндекс 2011 г. | Яндекс 2012 г. | Яндекс 2012 г. | Яндекс 2012 г. | Яндекс 2012 г. | Яндекс 2012 г. | Яндекс 2012 г. | Яндекс 2012 г. |

рез дробь – отклонение от среднего значения, указанного в последней строке таблицы. Там же (в последней строке) через дробь указывается и среднее отклонение.

Обращает на себя внимание то, что поиск разных студентов по одному и тому же запросу даёт разные результаты. Отчасти это обусловлено тем, что выдача систем поиска изменяется с течением времени, но основная причина – разница в оценках студентами полезности полученной информации. Это различие вполне естественно и не свидетельствует о противоречивости данных. В целом данные согласуются с ожиданиями теории информационного поиска: наиболее релевантная информация (полезные ссылки) находится в начале выдачи; с увеличением полноты поиска (количества просмотренных ссылок) точность T убывает, а степень удовлетворённости (D) возрастает. Суммарный показатель эффективности $\Sigma = T + D$, как правило, достигает определённого максимума на одной из начальных страниц выдачи и затем закономерно снижается вслед за снижением точности T и стабилизацией показателя достаточности D поиска около 100%.

В таблице 7 (см. с. 154) сведены средние значения показателей эффективности для всех трёх случаев поиска. По этим данным видно, что показатели системы Яндекс, вычисленные по двум разным независимым совокупностям экспериментов, достаточно близки. Их различие гораздо меньше среднего разброса показателей для отдельных поисков. Это свидетельствует об устойчивости выбранных параметров – точности T , достаточности D и их суммы Σ – и о возможности использовать их в качестве показателей эффективности поисковых систем. Система Гугл оказалась несколько менее эффективна, несмотря на то, что показатель точности у неё выше, чем у Яндекса. Параметр достаточности у Гугла оказался закономерно ниже при всех значениях глубины поиска. Соответственно, меньшая эффективность Гугла выражается в менее высоком показателе суммы точности и достаточности (Σ). Предварительно можно выдвинуть гипотезу, что Гугл позволяет быстрее найти основные сведения, но Яндекс при углублённом поиске более полно удовлетворяет информационную потребность.

Видимо, по причине меньшей информативности выдач Гугл пользуется у студентов МГУКИ меньшей популярностью, чем Яндекс. Гугл был выбран для исследования меньшим количеством студентов. В силу этого статистика для Гугла не очень представительна. Окончательный вывод об относительной эффективности этих двух авторитетных поисковиков можно будет сделать на основе привлечения дополнительных данных. Кроме того, следует изучить, как влияет на эффективность поиска общий характер запросов, различия в специализации студентов и возможность совершенствования систем с течением времени.

Список литературы:

1. Основы информатики / А. И. Михайлов, А. И. Чёрный, Р. С. Гиляревский. – М., 1968. – 756 с.
2. *Гиляревский, Р. С.* Основы информатики : курс лекций / Р. С. Гиляревский. – М.: Экзамен, 2004. – 320 с.
3. *Белоозеров, В. Н.* Социальная теория информатики : курс лекций / В. Н. Белоозеров. – М.: МГУКИ, 2013. – 160 с.

4. *Ефимов, А. Н.* Интернет как информационный массив: Применимы ли общие информационные законы к Интернету? / А. Н. Ефимов, М. В. Шойхер // Научно-техническая информация. Сер. 1. – 1999. – № 7. – С. 34–38.
5. *Козлов, М. В.* Метод оценки эффективности функционирования современных информационно-поисковых систем Интернета / М. В. Козлов, В. А. Яцко // Компьютерная лингвистика и интеллектуальные технологии : сб. тр. междуна-родн. конф. по компьютерной лингвистике «Диалог» [Электронный ресурс]. – 2006. – Режим доступа: <http://www.dialog-21.ru/digests/dialog2006/materials/html/Kozlov.htm>.

Spisok literatury:

1. *Osnovy informatiki / A. I. Mikhaïlov, A. I. Chërnyĭ, R. S. Giliarevskii.* – М., 1968. – 756 с.
2. *Giliarevskii, R. S. Osnovy informatiki : kurs lektsii / R. S. Giliarevskii.* – М.: Èkzamen, 2004. – 320 с.
3. *Beloozerov, V. N. Sotsial'naia teoriia informatiki : kurs lektsii / V. N. Beloozerov.* – М.: MGUKI, 2013. – 160 с.
4. *Efimov, A. N. Internet kak informatsionnyĭ massiv: Primenimy li obshchie informatsionnye zakony k Internetu? / A. N. Efimov, M. V. Shoïkher // Nauchno-tekhnicheskaiia informatsiia. Ser. 1. – 1999. – № 7. – S. 34–38.*
5. *Kozlov, M. V. Metod otsenki éffektivnosti funktsionirovaniia sovremennykh informatsionno-poiskovykh sistem Interneta / M. V. Kozlov, V. A. Iatsko // Komp'iuternaia lingvistika i intellektual'nye tekhnologii : sb. tr. mezhdunarodn. konf. po komp'iuternoĭ lingvistike «Dialog» [Élektronnyĭ resurs]. – 2006. – Rezhim dostupa: <http://www.dialog-21.ru/digests/dialog2006/materials/html/Kozlov.htm>.*

Интернет-журнал
«Проблемы современного образования»
2014, № 3