

Лабораторная работа 1 Поиск информации в сети Интернет

| | |
|---|----|
| Требования к оформлению отчета: | 1 |
| Требования к оформлению письма: | 1 |
| Поисковые системы | 2 |
| Поисковые каталоги | 2 |
| Поисковые индексы | 3 |
| Как работает поисковой индекс? | 3 |
| Поисковая система Yandex | 3 |
| Поиск по рубрикатуру поисковой системы..... | 5 |
| Задание 1..... | 7 |
| Поиск по ключевым словам | 8 |
| Задание 2..... | 10 |
| Правила формирования поисковых запросов в системе Yandex..... | 11 |
| Задание 3..... | 13 |
| Задание 4..... | 13 |
| Задание 5*..... | 14 |
| Задание 6..... | 15 |
| Контрольные вопросы: | 15 |

В данной лабораторной работе требуется оформить отчет по результатам выполнения заданий.

Требования к оформлению отчета:

Отчет представляет собой текстовый документ, выполненный в редакторе MS Word.

В текстовом документе записано: группа, № варианта, номер и название лабораторной работы. Каждое задание предваряется строкой с формулировкой задания.

Текстовый файл прислать вложением на электронный адрес преподавателя (предварительно уточните адрес).

Требования к оформлению письма:

В теме письма указывается группа (в русской литерации), фамилия и имя, номер работы: Например, так: ПС-101 Иванова Мария работа 2

К сведению: За несоблюдение правил оформления оценка будет снижаться, с формулировкой «*несоблюдение требований к оформлению и сетевого этикета*».

Поисковые системы

Сеть Интернет представляет собой базу данных, в которой хранится огромное количество самой разнообразной информации. Очень важно суметь найти во всем этом многообразии именно ту информацию, которая окажется действительно полезной и поможет вам решить поставленную перед вами задачу.

Искать информацию можно с помощью поисковой системы. В Интернете имеется достаточное количество подобных систем: yandex.ru, google.ru, nigma.ru, yahoo.com и др. В каждой поисковой системе имеются свои механизмы поиска. Некоторые системы имеют дополнительные средства фильтрации уже отобранных данных, например nigma.ru.

Поисковая система — веб-сайт, предоставляющий возможность поиска информации в Интернете.

Большинство поисковых систем ищут информацию на сайтах Всемирной паутины, но существуют также системы, способные искать файлы на ftp-серверах, товары в интернет-магазинах, а также информацию в группах новостей Usenet.

По принципу действия поисковые системы делятся на два типа: *поисковые каталоги* и *поисковые индексы*.

Поисковые каталоги

Поисковые каталоги служат для тематического поиска.

Информация на этих серверах структурирована по темам и подтемам. Имея намерение осветить какую-то узкую тему, нетрудно найти список веб-страниц, ей посвященных.

Каталог ресурсов в Интернете или каталог интернет-ресурсов или просто интернет-каталог — структурированный набор ссылок на сайты с кратким их описанием.

Каталог в котором ссылки на сайты внутри категорий сортируются по популярности сайтов называется рейтинг (или топ).

Поисковые индексы

Поисковые индексы работают как алфавитные указатели. Клиент задает слово или группу слов, характеризующих его область поиска, — и получает список ссылок на веб-страницы, содержащие указанные термины.

Первой поисковой системой для Всемирной паутины был «Wandex», уже не существующий индекс, разработанный Мэтью Грэйем из Массачусетского технологического института в 1993.

Как работает поисковой индекс?

Поисковые индексы автоматически, при помощи специальных программ (веб-пауков), сканируют страницы Интернета и индексируют их, то есть заносят в свою огромную базу данных.

Поисковый робот («веб-паук») — программа, являющаяся составной частью поисковой системы и предназначенная для обхода страниц Интернета с целью занесения информации о них (ключевые слова) в базу поисковика. По своей сути паук больше всего напоминает обычный браузер. Он сканирует содержимое страницы, забрасывает его на сервер поисковой системы, которой принадлежит и отправляется по ссылкам на следующие страницы.

В ответ на запрос, где найти нужную информацию, поисковый сервер возвращает список гиперссылок, ведущих веб-страницам, на которых нужная информация имеется или упоминается. Обширность списка может быть любой, в зависимости от содержания запроса.

Поисковая система Yandex

Yandex является пожалуй наилучшей поисковой системой в российском Интернете. Эта база данных содержит около 200 000 серверов и до 30 миллионов документов, которые система просматривает в течение нескольких секунд. На примере этой системы покажем как осуществляется поиск информации.

Поиск информации задается введением ключевого слова в специальную рамку и нажатием кнопки «Найти», справа от рамки (рис. 1).

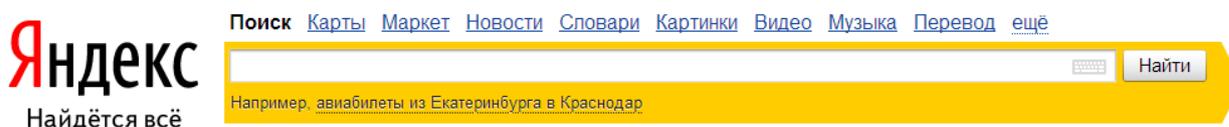


Рисунок 1 – Строка поиска Yandex

Результаты поиска появляются в течение нескольких секунд, причем ранжированные по значимости – наиболее важные документы размещаются в начале списка. При этом ранг найденного документа определяется тем, в каком месте документа находится ключевое слово (в заголовке документа важнее, чем в любом другом месте) и числом упоминаний ключевого слова (чем больше упоминаний, тем ранг выше) (рис. 2).

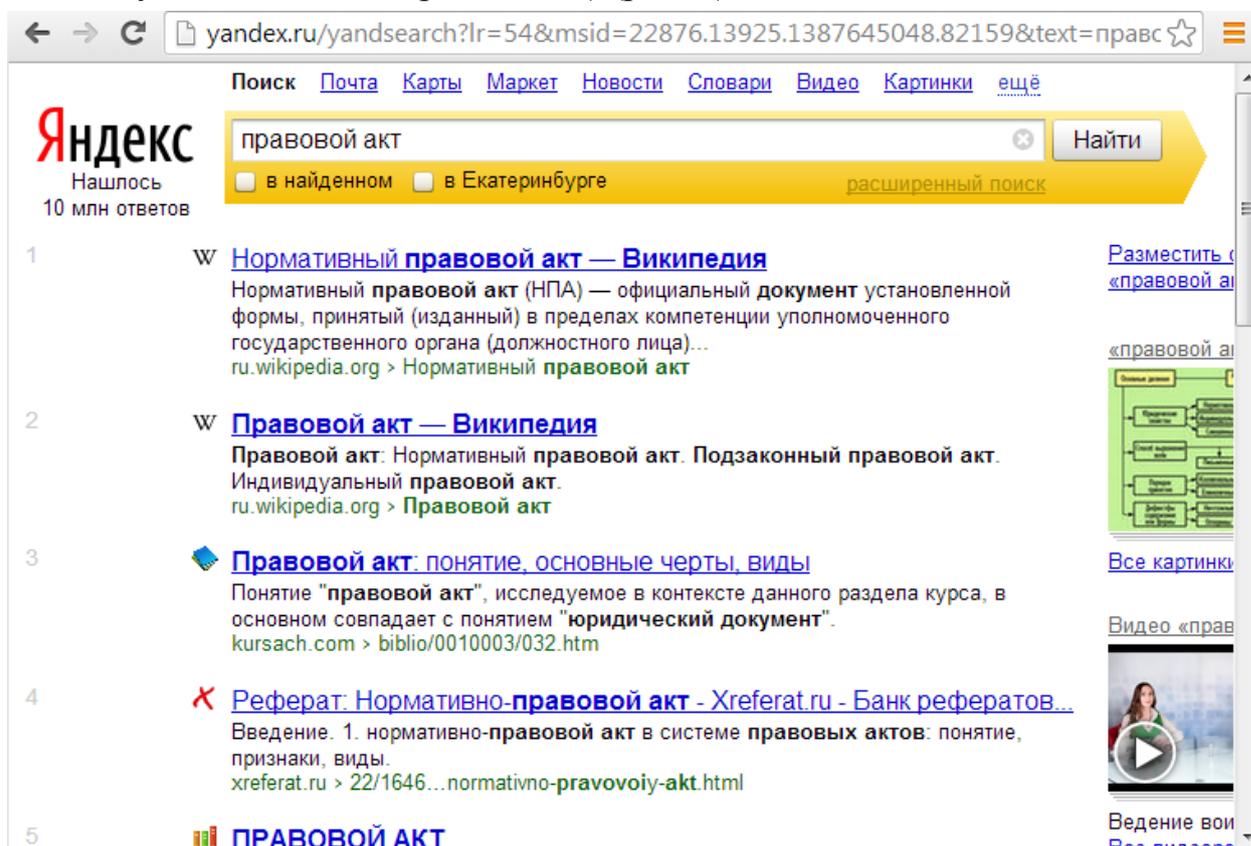


Рисунок 2 – результаты поиска по запросу

Таким образом, сайты, расположенные на первых местах в списке, являются ведущими не с содержательной точки зрения, а практически, по отношению к частоте упоминания ключевого слова. В связи с этим, не следует ограничиваться просмотром первого десятка предложенных поисковой системой сайтов.

Содержательную часть сайта можно косвенно оценить по краткому его описанию, размещаемому поисковой системой под адресом сайта. Некоторые недобросовестные авторы сайтов, для того, чтобы повысить вероятность появления своей веб-страницы на первых местах поисковой системы, умышленно включают в документ бессмысленные повторы ключевого слова. Но как только поисковая система обнаруживает такой «замусоренный» документ, она автоматически исключает его из своей базы данных.

Даже ранжированный список документов, предлагаемый поисковой системой в ответ на ключевую фразу или слово, может оказаться практически необозримым. В связи с этим в Yandex (как и других мощных Поисковых Системах) предоставлена возможность в рамках первого списка, выбрать документы, которые точнее отражают цель поиска, то есть уточнить или улучшить результаты поиска. Для этого можно воспользоваться флажком **в найденном**, который расположен под строкой поиска (см. рис. 2). Данным параметром можно пользоваться не один раз, с каждым разом вводя в строку поиска уточняющие ключевые слова (рис. 3).

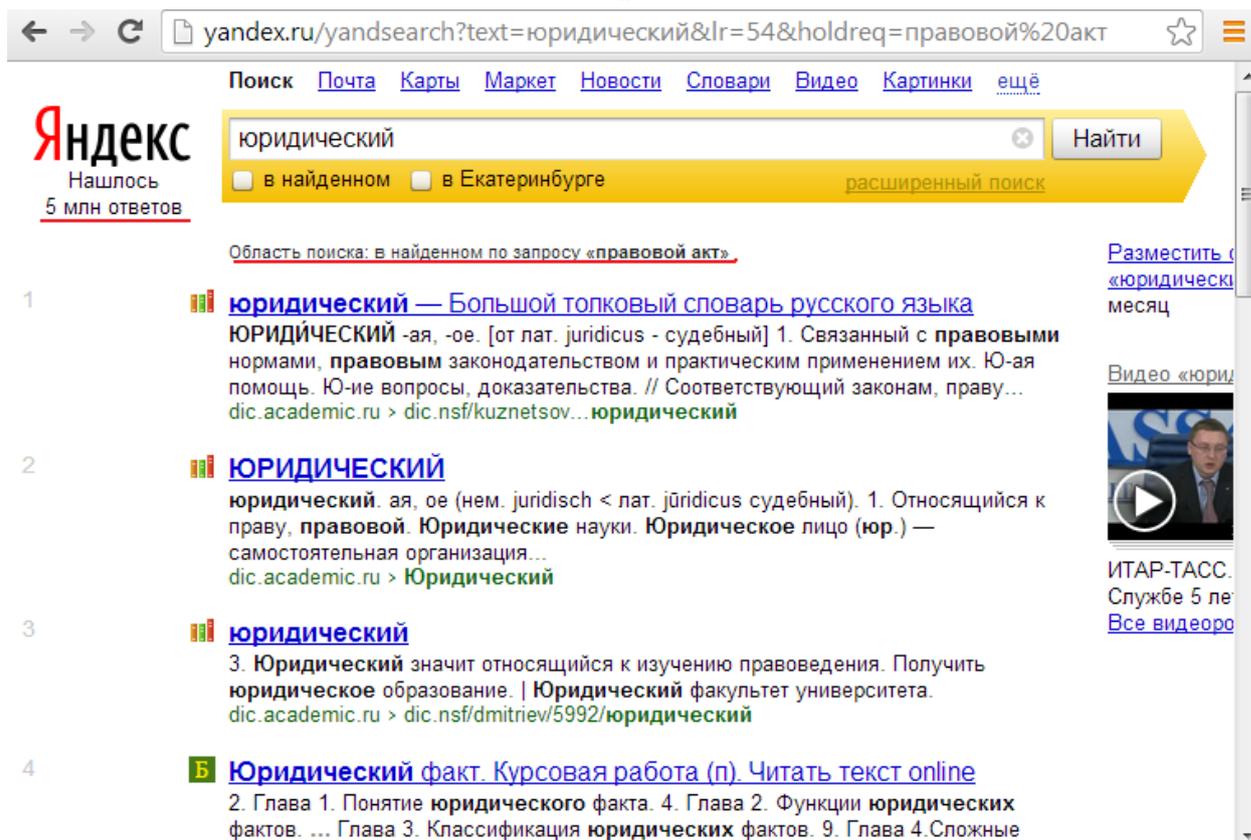


Рисунок 3 – улучшение результатов поиска с помощью флажка **в найденном**

Поиск по рубрикатору поисковой системы

Поисковые каталоги представляют собой систематизированную коллекцию (подборку) ссылок на другие ресурсы Интернета. Ссылки организованы в виде тематического рубрикатора, представляющего собой иерархическую структуру, перемещаясь по которой, можно найти нужную информацию.

Приведем в качестве примера структуру поискового интернет-каталога Yandex.

В верхней строке меню, которая расположена над строкой поиска в пункте **ещё** (рис. 4) необходимо выбрать команду **Каталог**. В результате окно браузера примет следующий вид (рис. 5).

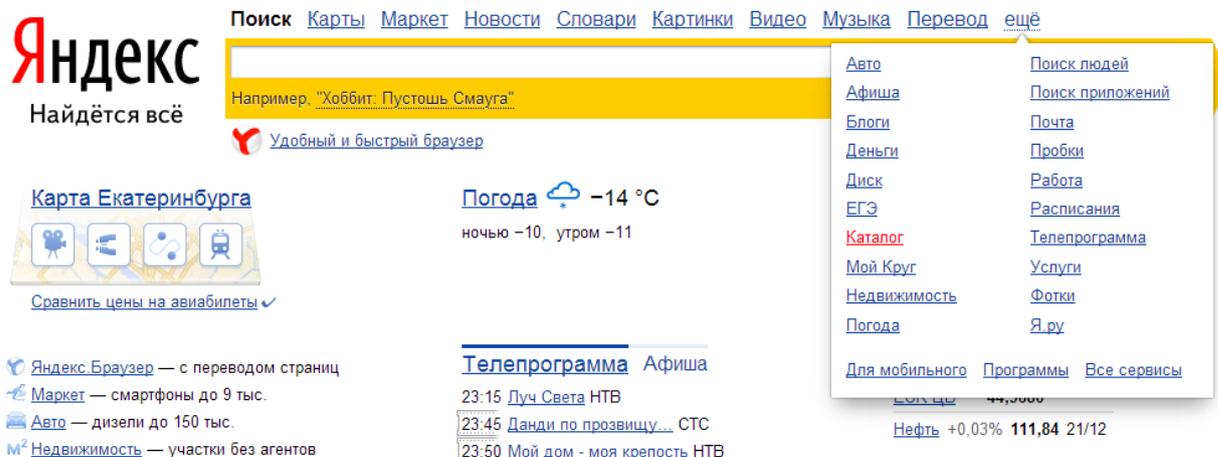


Рисунок 4 – Выбор Интернет-каталога Yandex

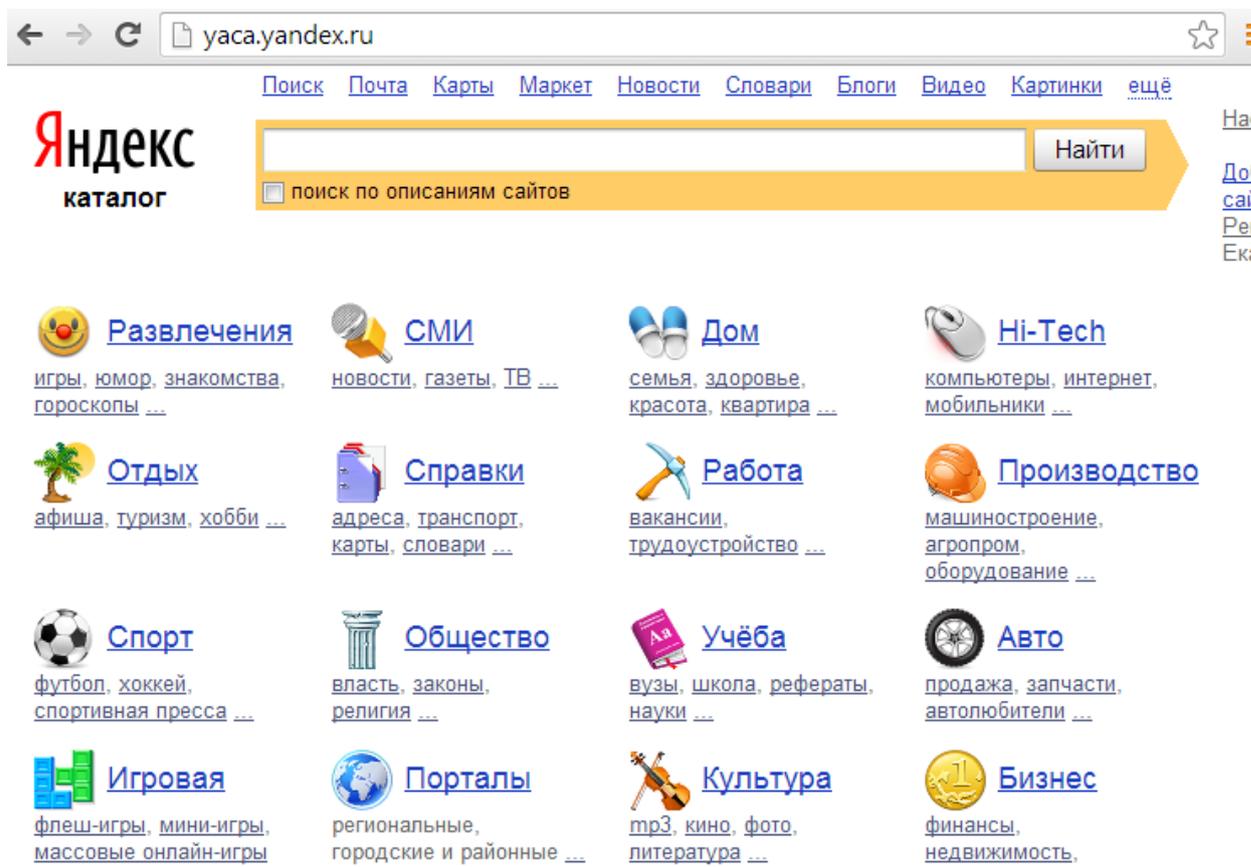


Рисунок 5 – внешний вид каталога Yandex

Это каталог общего назначения, так как в нем представлены ссылки на ресурсы Интернета практически по всем возможным направлениям: Развлечения, СМИ, Отдых, Дом, Культура, Учеба и т.д.

Каждая тема включает множество подразделов, а они, в свою очередь, содержат рубрики и т. д.

Предположим, вы готовите мероприятие ко Дню победы и хотите найти в Интернете слова известной военной песни Булата Окуджавы «Вы слышите, грохочут сапоги». Поиск можно организовать следующим образом:

Yandex > Каталог > Культура > Музыка > Авторская песня

Такой способ поиска является достаточно быстрым и эффективным. В конце вам предлагается достаточно ограниченное число ссылок, среди которых есть ссылки на сайты с песнями известных бардов. Остается только найти на сайте архив с текстами песен Б. Окуджавы и выбрать в нем нужный текст. Для ускорения поиска вы можете воспользоваться строкой поиска и флажком *только в этой рубрике* (рис. 6).

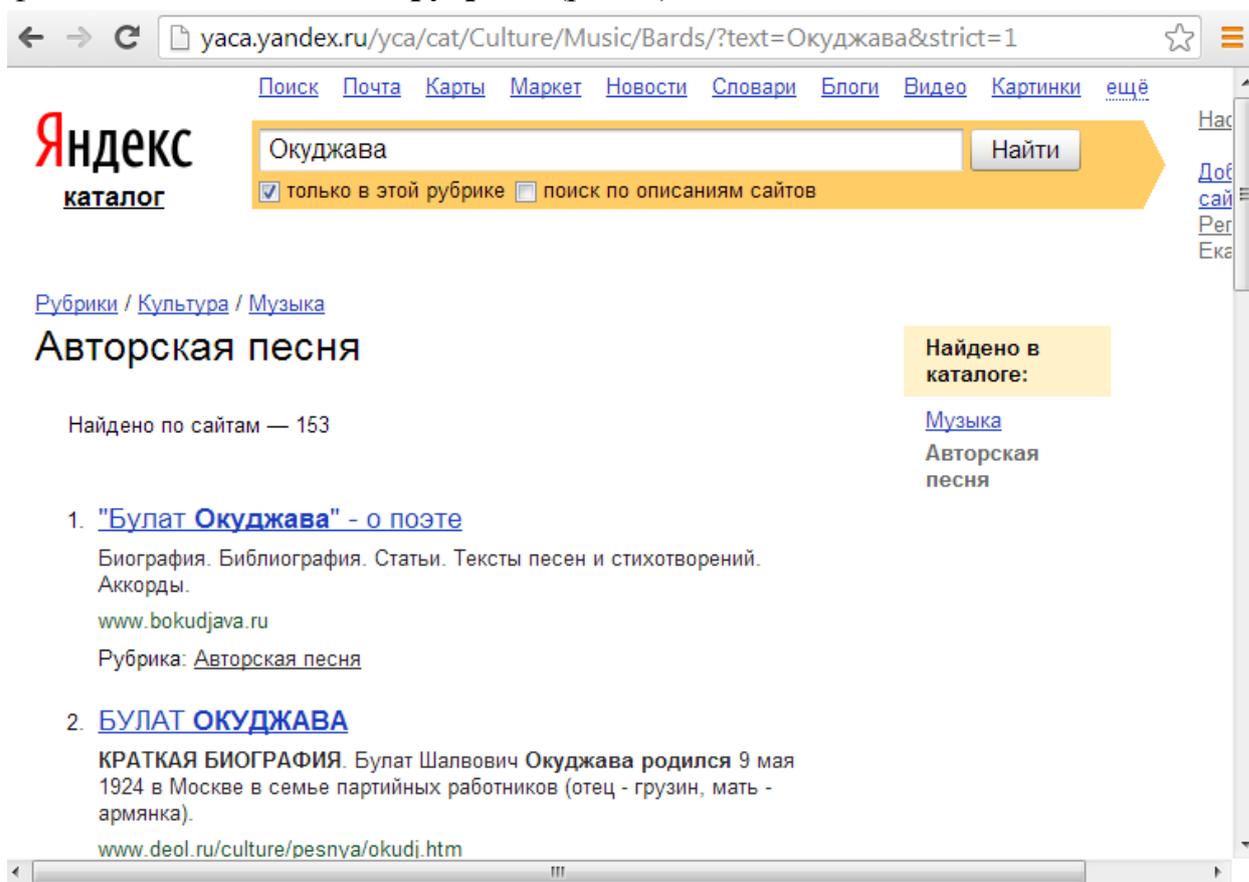


Рисунок 6 – Поиск по каталогу

Задание 1

Запустите браузер и в строке адреса укажите адрес поисковой системы yandex.ru.

С помощью каталога, данной поисковой системы, найдите следующую информацию (уточните вариант у преподавателя):

1. Текст песни популярной музыкальной группы (по вашему выбору);
2. Репертуар Мариинского театра на текущую неделю;

3. Характеристики последней модели мобильного телефона известной фирмы (по вашему выбору);
4. Рецепт приготовления украинского борща с галушками;
5. Долгосрочный прогноз погоды в вашем регионе (не менее чем на 10 дней);
6. Фотография любимого исполнителя современной песни;
7. Примерная стоимость мультимедийного компьютера (прайс);
8. Информация о вакансиях на должность секретаря в вашем регионе или городе;
9. Гороскоп своего знака зодиака на текущий день.
10. Статистику хоккейного клуба «Локомотив».
11. Сайты судов Урала, Свердловской области и Екатеринбурга.
12. Текст конституции Российской Федерации.

По результатам поиска составьте письменный отчет в Word: представьте в документе найденный, скопированный и отформатированный материал. Вначале выполненного задания укажите текст задания. Сохраните отчет в вашу папку.

Поиск по ключевым словам

Большинство поисковых систем имеют возможность поиска по ключевым словам. Это один из самых распространенных видов поиска.

Для поиска по ключевым словам необходимо ввести в специальном окне слово или несколько слов, которые следует искать, и щелкнуть на кнопке Найти.

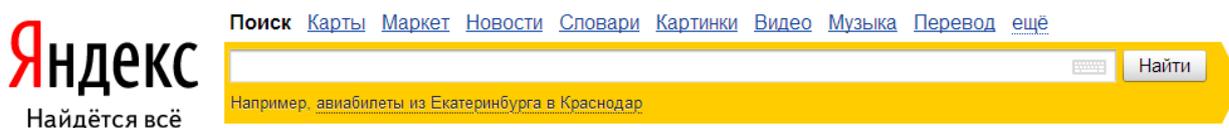


Рисунок 7 – Строка поиска

Поисковая система найдет в своей базе и покажет документы, содержащие эти слова. Таких документов может оказаться множество, но много в данном случае не обязательно означает хорошо.

Проведем несколько экспериментов с любой из поисковых систем. Предположим, что мы решили завести аквариум и нас интересует любая информация по данной теме. На первый взгляд самое простое — это поиск по слову «аквариум». Проверим это, например, в поисковой системе Yandex. Результатом поиска будет огромное количество страниц - огромное количество ссылок. Причем, если посмотреть внимательнее, среди них окажутся сайты,

упоминающие группу Бориса Гребенщикова «Аквариум», торговые центры и неформальные объединения с таким же названием, и многое другое, не имеющее отношения к аквариумным рыбкам.

Нетрудно догадаться, что такой поиск не может удовлетворить даже непритязательного пользователя. Слишком много времени придется потратить на то, чтобы отобрать среди всех предложенных документов те, что касаются нужного нам предмета, и уж тем более на то, чтобы ознакомиться с их содержанием.

Можно сразу сделать вывод, что вести поиск по одному слову, как правило, нецелесообразно, ведь по одному слову очень сложно определить тему, которой посвящен документ, веб-страница или сайт. Исключение составляют редкие слова и термины, которые практически никогда не используются вне своей тематической области.

Имея определенный набор наиболее употребительных терминов в нужной области, можно использовать расширенный поиск. На рис. 8 показано окно расширенного поиска в поисковой системе Yandex. В этом режиме возможности языка запросов реализованы в виде формы. Подобный сервис, включающий словарные фильтры, предлагается почти всеми поисковыми системами.

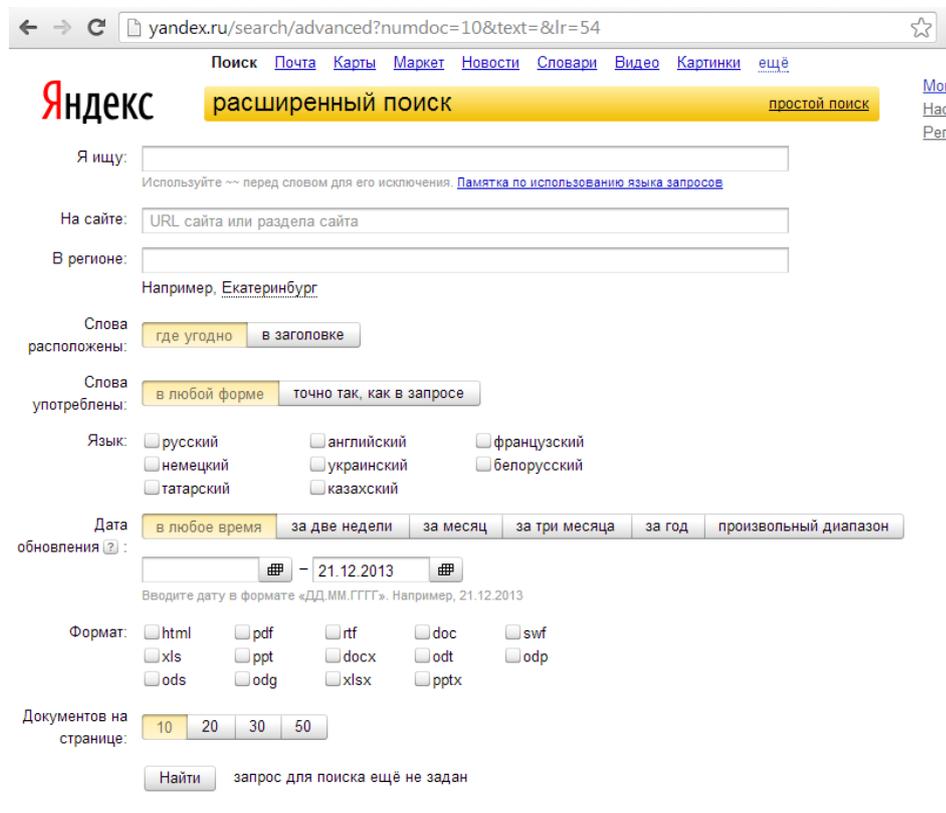


Рисунок 8 – Окно расширенного поиска

Уточнив условие в окне расширенного поиска, мы сократим количество ссылок более чем в 20 раз. Однако, результат поиска можно еще улучшить.

Для того чтобы сделать поиск более продуктивным, во всех поисковых системах существует специальный язык формирования запросов со своим синтаксисом. Эти языки во многом похожи. Изучить их все достаточно сложно, но любая поисковая система имеет справочную систему, которая позволит вам освоить нужный язык.

Для того, чтобы вывести на экран окно расширенного поиска достаточно ничего не указывая в строке поиска нажать на кнопку Найти, а затем, после загрузки страницы, под строкой поиска выбрать появившуюся ссылку расширенный поиск (рис. 9).



Рисунок 9 – Ссылка на расширенный поиск

Задание 2

Изучив окно расширенного поиска сформируйте запрос по точному названию или цитате.

Вам известно точное название документа, например «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы». Сформулируйте запрос для поиска в Интернете полного текста документа (вариант уточните у преподавателя).

1. Гражданский кодекс Российской Федерации (часть первая).
2. Гражданский кодекс Российской Федерации (часть вторая).
3. Окинавская хартия глобального информационного общества.
4. О стратегии сотрудничества государств-участников СНГ в построении и развитии информационного общества и Плане действий по ее реализации на период до 2015года.
5. О Концепции правовой информатизации России.
6. О президентских программах правовой информатизации.
7. О Государственной автоматизированной системе Российской Федерации «Выборы».
8. Об информации, информационных технологиях и о защите информации.
9. Доктрина информационной безопасности Российской Федерации.

Результат поиска сохраните в виде копии экрана (клавиша PrintScrin). Добавьте данное изображение в отчет, который вы начали готовить выполняя задание 1. Над картинкой укажите текст задания. Сохраните отчет в вашу папку.

Правила формирования поисковых запросов в системе Yandex

Самым важным фактором успешного поиска информации является то, какой критерий поиска пользователь ввел в строку запроса, как составил запрос.

Если вы в строке поиска в качестве запроса указали фразу типа:

Скажи-ка, дядя, ведь не даром Москва, спаленная пожаром, Французу отдана?

то любая поисковая система сразу выдаст вам перечень источников, где содержится либо стихотворение М.Ю. Лермонтова «Бородино», либо информация о бородинском сражении. Если вы зададите дополнительные параметры поиска (имеются не во всех поисковых системах!), то вы сможете отобрать более конкретную информацию по данному запросу. Такой быстрый результат поиска, в данном случае, мы получили в связи с тем, что эта фраза достаточно распространенная и однозначная. Если же мы в строке поиска укажем фразу:

О, подскажи мне всемогущий Интернет, где я могу купить для ужина котлет?

то вряд ли поисковая система выдаст нам что-то вразумительное, приближенное к тому, что мы хотели найти.

Для более точного поиска информации необходимо убрать из этого запроса лишние слова:

О, подскажи мне всемогущий Интернет, где я могу купить для ужина котлет?

Из всего множества слов в запросе ключевыми являются только два: **купить** и **котлеты**.

Еще один пример, приближенный непосредственно к вашему направлению подготовки. Если мы в строке поиска укажем фразу:

В чем заключается юридическая сила нормативного правового акта?

то в качестве результата поиска будет выдано огромное количество ссылок на сайты, на которых встречаются либо все слова и словосочетания, имеющиеся в данном запросе, но в разной последовательности, либо какие-то от-

дельные его части. Для того чтобы поисковая система включала в результаты поиска как можно меньше ненужной информации, необходимо убрать из поискового запроса лишние слова. Очень важно, для правильного поиска информации, указывать в поисковой строке только те ключевые слова и словосочетания, которые важны для результатов поиска. Например, **юридическая сила нормативного правового акта**.

Помимо ввода в строку поиска ключевых словосочетаний, можно повлиять на результаты отбора информации с помощью специальных символов.

Если вы указали в строке поиска фразу **нормативный правовой акт**, то в результаты поиска будут отображены страницы, на которых встречаются по отдельности, а так же в любой последовательности слова и словосочетания:

- нормативный;
- правовой;
- акт;
- правовой акт;
- нормативный акт.

Кроме того, при поиске информации будут учтены все формы слова по правилам русского языка, независимо от формы слова в запросе. Таким образом указав в запросе слово «акт», вам будут найдены результаты со словом «акты», «актом», «актов» и т.п.

Для поиска по точной словоформе перед словом необходимо поставить восклицательный знак – например, **!акт**.

Для поиска устойчивого словосочетания необходимо заключить данную фразу в кавычки – **«нормативный правовой акт»**. Тем самым объем результатов поиска значительно сократится.

Если вы хотите, чтобы были отображены только те документы, в которых встретилось каждое слово, указанное в запросе, поставьте перед каждым из них знак плюс «+». Если вы, наоборот, хотите исключить какие-либо слова из результата поиска, поставьте перед этим словом минус «-».

Знаки « + » и «-» надо писать через пробел от предыдущего и слитно со следующим словом

Например, если вы хотите найти словосочетание «правовой акт», то необходимо написать запрос следующим образом – **правовой +акт**. Если наоборот, вы хотите найти все документы, а которых встречается слово «правовой», но не «акт», то запрос следует составить так – **правовой -акт**.

При поиске синонимов или близких по значению слов между словами можно поставить вертикальную черту «|». Например, **акт | документ**.

Для поиска внутри одного предложения слова в запросе разделяют пробелом или знаком &.

Например, **правовой акт** или **правовой&акт**. Несколько набранных в запросе слов, разделенных пробелами, означают, что все они должны входить в одно предложение искомого документа.

Знак «~» (тильда) позволяет найти документы с предложением, содержащим первое слово, но не содержащим второе.

Например, по запросу **нормативный ~ правовой** будут найдены все документы, содержащие слово «нормативный», рядом с которым (в пределах предложения) нет слова «правовой».

Если оператор повторяется один раз (например, & или ~), поиск производится в пределах предложения. Двойной оператор (&&,~~) задает поиск в пределах документа.

Например, по запросу **правовой ~ акт** будут найдены документы со словом «правовой», не относящиеся к актам.

Задание 3

Путем формирования сложных запросов найдите следующую информацию:

1. составьте запрос для поиска информации о русской бане. Исключите предложения об услугах, рекламу банных принадлежностей и прочую рекламу. Сосредоточьте поиск на влиянии русской бани на организм.
2. Составьте сложный запрос на поиск информации по уходу за домашними кошками. Исключите из поиска крупных кошек (например, львов), а также предложения о покупке, продаже, фотографии для обоев и т. п.

Текст запроса и результат поиска оформите в вашем отчете, предварительно указав текст задания.

Задание 4

С помощью поисковой системы найдите Федеральный закон РФ «Об информации, информационных технологиях и о защите информации» от 27 июля 2006 г. № 149-ФЗ и выпишите в свой отчет следующие определения:

- Информация;
- Документированная информация;
- Общедоступная информация.

Задание 5*

В сети Интернет найдите информацию, представленную в вашем варианте, и вынесите ее в ваш отчет.

При выполнении задания обязательно указывать источник, согласно следующему правилу (в соответствии с ГОСТ 7.0.5-2008): Название интернет-ресурса, полный URL-адрес, дата обращения.

Например: Справочники по физике // [Персональная страница В.Р. Козака] / URL: <http://www.inp.nsk.su/%7EKosak/start.htm> (дата обращения: 13.03.12).

1. Перечислить формы представления информации;
2. Дать определение информационного процесса и приведите 2 примера информационных процессов из области биологии;
3. Привести примеры 3-х различных информационных процессов из различных областей знаний (или деятельности людей);
4. Привести 3 примера, иллюстрирующих отличие информации от данных;
5. Приведите 3 примера, при которых данные могут являться информацией или не могут быть признаны в качестве информации;
6. Приведите 3 примера актуальной и достоверной информации из области математики.
7. Приведите примеры, когда к информации применимо понятие репрезентативности;
8. Дайте определение такому свойству информации, как адекватность и приведите по 3 примера адекватной информации из физики.
9. Приведите примеры из социальной (или культурной, или общественно-политической) сферы, когда неполное раскрытие информации делало ее недостоверной.
10. Дайте определение свойству достоверность информации. Что может служить критерием достоверности информации в точных науках, в гуманитарных науках?
11. Определите, в чем появляется сходство и различие в толковании свойств достоверность и адекватность информации. Приведите не менее 2-х примеров.
12. Приведите 3 примера неактуальной, но достоверной информации из области математики.

13. Дайте определение такому свойству информации, как адекватность и приведите по 2 примера адекватной и неадекватной информации из физики.

Задание 6

Сохраните ваш отчет и отправьте его на электронный адрес вашего преподавателя.

Пригласите преподавателя и побеседуйте с ним о том, чему вы научились, выполняя данную лабораторную работу.

Контрольные вопросы:

1. Что такое поисковая система?
2. Какие типы поисковых систем вы знаете?
3. Что такое поисковый каталог? Как осуществляется поиск информации в таком каталоге?
4. Дайте определение поисковому роботу.
5. Назовите способы улучшения результатов поиска.
6. Что представляет собой рубрикатор поисковой системы?
7. Какова технология поиска по ключевым словам?
8. Какие правила формирования поисковых запросов в системе Yandex вы знаете?
9. Когда в критерии поиска надо задавать + или -?
10. Какие критерии поиска в Yandex заданы следующей фразой:
(няня|воспитатель|гувернантка)++(уход|воспитание|присмотр)?
11. Что означает удвоение знака (~~ или ++) при формировании сложного запроса?