

Практическая работа

Тема: Поисковые системы. Осуществление поиска информации или информационного объекта в тексте, в файловых структурах, в базах данных, в сети Интернет.

Цель работы. Изучение способов поиска информации в тексте, в файловых структурах, в базах данных, в сети Интернет.

Краткие сведения

Поиск информации - задача, которую человечество решает уже многие столетия. По мере роста объема информационных ресурсов, потенциально доступных одному человеку, были выработаны все более изощренные и совершенные поисковые средства и приемы, позволяющие найти необходимый документ. Обширные возможности для работы с большими массивами информации дают поисковые сервисы Internet.

При наличии первичных сведений по теме поиска, документы можно разыскивать в поисковых системах. При этом следует различать приемы *простого, расширенного, контекстного и специального поиска*.

Под *простым поиском* понимается поиск Web-ресурсов по одному или нескольким ключевым словам. Недостаток простого поиска заключается в том, что обычно он выдает слишком много документов, среди которых трудно выбрать наиболее подходящие.

При использовании *расширенного поиска* ключевые слова связывают между собой операторами логических отношений. Расширенный поиск применяют в тех случаях, когда приемы простого поиска дают слишком много результатов. С помощью логических отношений поисковое задание формируют так, чтобы более точно детализировать задание и ограничить область отбора, например по дате публикации или типу данных.

Контекстный поиск – это поиск по точной фразе. Он удобен для реферативного поиска информации, но доступен далеко не во всех поисковых системах. Прежде всего, чтобы обеспечивать такую возможность, система должна работать не только индексированными файлами, но и с полноценными образами Web-страниц. Эта операция достаточно медленная, и ее выполняют не все поисковые системы.

Специальный поиск применяют при розыске Web-страниц, содержащих ссылки на заданные адреса URL, содержащих заданные данные в служебных полях, например в поле заголовка и т.п.

Расширенный поиск. Кроме средства простого поиска обычно поисковые службы предоставляют средства расширенного поиска. Эти средства позволяют более точно формулировать поисковое задание, но требуют определенного опыта и работают заметно медленнее. В большинстве поисковых систем команды расширенного поиска формируются с помощью логических команд. Удобство использования логических команд в частности связано с тем, что команды простого поиска у многих поисковых систем реализованы по-разному. Каждая система стремится сделать средства простого поиска наиболее удобными, а средства расширенного поиска – наиболее стандартными. Тем не менее, для обозначения логических операторов в различных поисковых системах используются разные обозначения. Поэтому желательно перед осуществлением расширенного поиска желательно изучить синтаксис поисковых запросов выбранной поисковой системы.

Рассмотрим подробнее операторы логических отношений (логические команды).

Логическая оператор **OR (ИЛИ)** служит для формирования поискового запроса, если искомый текст должен содержать хотя бы один из терминов, соединенных данным оператором. Этот оператор в различных поисковых системах может обозначаться одним из следующих способов: **|**; **OR**; **ИЛИ**.

Например, результат запроса «Чёрное OR море» - будет представлен списком ссылок на документы, в которых есть слово «Чёрное», или слово «море», или оба этих слова вместе.

В некоторых поисковых системах, как отмечалось выше, по умолчанию ключевые слова в запросе связаны именно этим логическим отношением.

С помощью логического оператора **AND (И)** осуществляется поиск документов, содержащих все термины, соединенные данным оператором. Этот оператор может обозначаться одним из следующих способов: **+**; **AND**; **&**; **И**.

Например, по запросу – «Чёрное AND море» - будут найдены документы, в которых содержатся слова «черное» и «море».

Логической оператор **NOT (НЕ)** позволяет производить поиск документов, в тексте которых отсутствуют термины, следующие за данным оператором. Этот оператор может обозначаться одним из следующих способов: **not; !; ~; НЕ**.

Например, по запросу – «Чёрное NOT море», результат - документы, в которых есть слово «Чёрное» и нет слова «море».

С помощью логических операций можно создавать достаточно сложные запросы. Запрос из нескольких слов, перемежающихся операторами, будет истолкован в соответствии с их приоритетом. Операторы AND и NOT традиционно имеют более высокий приоритет, поэтому запрос из нескольких слов при обработке сначала группируется по операторам AND и NOT, и лишь потом по операторам OR. Например, по запросу «Чёрное AND море OR Крым» будут найдены документы, либо содержащие обязательно слова: «Чёрное» и «море», либо слово «Крым», либо все три слова.

Изменить порядок группировки можно использованием скобок. Оператор, стоящий в скобках, будет выполняться в первую очередь. Использование скобок позволяет строить вложенные запросы и передавать их операторам в качестве аргументов. Так по запросу «Чёрное AND (море OR Крым)» будут найдены документы, в которых обязательно содержится слово «Чёрное» и одно из двух слов «море» или «Крым».

С помощью вложенных запросов можно значительно ограничивать область отбора, освобождая результирующий список от ненужных ссылок. Так, например, если нас интересует информация об отдыхе на юге на море, но исключительно на российском побережье, то можно попробовать использовать примерно такой запрос – «отдых AND ((Азовское OR Чёрное) AND море) NOT (Крым OR Турция OR Болгария)».

Использование круглых скобок для управления порядком исполнения задания на поиск разрешается большинством крупнейших поисковых систем.

Расширенный поиск yandex.ru

Оператор	Описание
"	Обнаруживает точные слова в кавычках или фразы
	Найти любое из слов. Достаточно поставить между словами символ , и вы получите страницы, где содержится хоть одно из слов запроса.
~	Исключает страницы, содержащие слово или фразу.
()	Вы можете строить сколь угодно сложные конструкции, подставляя в каждом из операторов вместо отдельного слова целые выражения. Чтобы Яндекс при этом правильно понимал вас, заключайте выражения в круглые скобки.
&	Ограничить поиск страницами, где слова запроса находятся в пределах предложения
&&	Если вам нужны документы, где присутствуют заданные слова — неважно, на каком расстоянии и в каком порядке — соедините их оператором
!	Слова с большой и маленькой буквы считаются разными формами одного слова, поэтому все равно, какой регистр использовать в запросе. Исключением является оператор точной формы. Это полезно, если искомое имя собственное совпадает с распространенным словосочетанием, например, группа !Черный кофе . Все слова, которые вы приводите в запросе, по умолчанию ищутся с учетом морфологии. Чтобы отключить ее, используйте оператор ! перед словом (без пробела).
/	Вы можете указать максимально допустимое расстояние между двумя любыми словами запроса, поставив после первого слова символ /, сразу за которым идет число, означающее расстояние.
*	Замена части слова. Журналист*
?	Замена любого символа. Журналистик?

Расширенный поиск rambler.ru

Оператор	Описание
+	Используется для включения общих слов
"	Слова запроса, заключенного в двойные кавычки, ищутся в документах именно в том порядке и в тех формах, в которых они встретились в запросе. Таким образом, двойные кавычки можно использовать и просто для поиска слова в заданной форме (по умолчанию слова находятся во всех формах).
NOT	Оператор NOT позволяет сформировать запрос, которому отвечают документы, удовлетворяющие левой части запроса и не удовлетворяющие правой. Так, результатом поиска по запросу собака NOT кошка будут все документы, в которых есть слово "собака" и нет

	слова "кошка". Это особенно полезно в случаях, если искомая словоформа является одновременно формой другого слова: Женя - имя собственное, а также деепричастие. Если мы ищем именно человека по имени Женя, в запросе можно написать Женя NOT женить.
&&	Два запроса, соединенные оператором &&, образуют сложный запрос, которому удовлетворяют только те документы, которые одновременно удовлетворяют обоим этим запросам. Иными словами, по запросу собака && кошка найдутся только те документы, которые содержат и слово "собака", и слово "кошка".
	Сложному запросу, состоящему из двух запросов, соединенных оператором , удовлетворяют все документы, удовлетворяющие хотя бы одному из этих двух запросов. По запросу <i>собака кошка</i> найдутся документы, в которых есть хотя бы одно из двух слов - слово "собака" или слово "кошка" (или оба эти слова вместе).
()	Использование скобок позволяет строить вложенные запросы и передавать их операторам в качестве аргументов, а также перекрывать приоритеты операторов, принятые по умолчанию.

Организация специального поиска. С помощью средств специального поиска можно осуществлять поиск документов: содержащих искомые данные в различных полях (заголовке, ключевых словах, описании); расположенных на каком-либо сайте; содержащих те или иные ссылки и т.п. Рассмотрим подробнее некоторые возможности специального поиска.

Многие поисковые системы позволяют разыскивать Web-документы по тексту, содержащемуся в заголовках. Поиск по заголовкам существенно уменьшает количество найденных ссылок, но очень точно выводит на нужные материалы. Ведь каждая Web-страница может иметь заголовок, если её автор не поленился его создать. И заголовок Web-страницы обычно точно характеризует тему материала, который содержится на ней.

Например, если вам нужна информация о дистанционном обучении, то целесообразно искать страницы, в которых это сочетание присутствует в заголовке. Таким образом, в отличие от простого запроса мы отсекаем те документы, где эти слова не являются значимыми, т.е. не определяют тему статьи.

Оператором или командой такого поиска является title, Этот оператор может обозначаться одним из следующих способов: title:, t:, title =, \$title и т.п. После оператора следуют ключевые слова. В некоторых поисковых системах ключевые слова следует заключать в скобки, в других они пишутся без скобок. Например, в Яндекс команда поиска в заголовке записывается так: \$title (очное обучение).

Поиск по сайту. С помощью поисковых систем можно осуществлять поиск информации не во всем Web пространстве, а на каком то конкретном сайте (если конечно последний проиндексирован поисковой системой). Соответствующий оператор может обозначаться следующим одним из следующих способов: url=, url:, u:, #url=". Далее следует адрес Web-узла, некоторые системы требуют заключения адреса в кавычки.

Если в запросе просто записать данный оператор с адресом какого-либо Web-узла, то будет получен список документов, проиндексированных поисковой системой на данном сайте. Но этот оператор можно комбинировать с другими, тем самым, осуществляя поиск информации по всем правилам построения запросов на данном сайте. Например, если мы хотим найти информацию об очном обучении на сайте БелГУ, то запрос в поисковой системе Апорт будет выглядеть следующим способом: URL=http://www.bsu.edu.ru + "очное обучение".

Кроме того, поисковые системы могут предлагать другие возможности специального поиска: поиск по тексту ссылок, поиск в описании документа, поиск в списке ключевых слов Web-страниц, поиск по подписям к рисункам и т.п. Следует знать, что синтаксис поисковых запросов, да и состав доступных операторов, отличается в различных поисковых указателях. Поэтому перед осуществлением поиска в той или иной поисковой системе, следует изучить страницу справки по поиску в данной системе.

Специальный поиск aport.ru

Оператор	Описание
title=	Указанное после равенства слово или конструкция в круглых скобках должны искаться в заголовках документов (перед круглыми скобками знак равенства можно опускать). По запросу title=(папа или мама) будут найдены документы, содержащие в заголовке слово папа или слово мама, или оба слова одновременно.
anchor=	Указанное после равенства слово или конструкция в круглых скобках должны искаться в тексте ссылок. По запросу anchor=(ненавижу Интернет) будут найдены документы, в тексте ссылок на которые, встречаются оба слова: ненавидеть и Интернет.

text=	Указанное после равенства слово или конструкция в круглых скобках должны искаться только в обычном тексте. По умолчанию слова запроса ищутся как в тексте, так и во всех указанных выше полях. Чтобы искать только по тексту, выдаваемому в основное окно браузера, следует использовать данный оператор. По запросу text=(оглавление или содержание) будут найдены документы, в которых любое из указанных слов встречается в пределах основного текста документа.
-------	---

Задание к работе

Задание 1. Выполнить поиск информационного объекта в тексте документа в текстовом процессоре Microsoft Word.

Порядок работы

1. Запустить на выполнение текстовый процессор Microsoft Word и открыть документ, в котором будет выполняться поиск фразы.
2. Выполнить команду Правка-Найти и указать искомый текст

Задание 2. Выполнить поиск информационного объекта с расширением PDF в файловых структурах диска D:

Порядок работы

1. В операционной системе Windows выполнить команду Пуск – Поиск
2. Указать искомый объект *.PDF
3. Указать место поиска – локальный диск D:

Задание 3. Выполнить поиск информационного объекта в сети Интернет

Порядок работы

1. Произведите простой поиск информации по ключевым словам «информационные технологии» в поисковых указателях Яндекс (<http://www.yandex.ru>), rambler (<http://www.rambler.ru>), АПОРТ (<http://www.aport.ru>). Результаты поиска по каждой поисковой системе запишите в отчет.
2. Осуществите контекстный поиск (поиск по точной фразе) информации в поисковой машине rambler по ключевым словам «информационные технологии». В файл отчета запишите запрос, удовлетворяющий данному условию и результат выполнения запроса в rambler.
3. Выполните поиск информации в поисковой машине Aport по ключевым словам «информационные технологии», но с помощью операторов расстояния задайте поиск так, чтобы слова «информационные» «технологии» стояли рядом.
4. Используя любую поисковую систему, найдите материал, который помог бы Вам ответить на следующие вопросы:
 1. Сколько медалей было завоевано российскими спортсменами на летней олимпиаде 2012 года в Лондоне?
 2. Годы царствования Петра I.
5. Выполните поиск документов, в которых ключевые слова «информационные технологии» находятся в заголовке. В файл отчета запишите запрос, удовлетворяющий данному условию и результат выполнения запроса в Яндекс.
6. Перейдите на сайт по адресу <http://www.edu.ru/>. В разделе учреждения выполните поиск информации о вашем техникуме. Запишите последовательность поиска в файл отчета.

Контрольные вопросы

Обязательная часть

1. Как выполнить поиск информационного объекта в текстовом процессоре MS Word?
2. Как выполнить поиск информационного объекта в файловых структурах Windows?
3. Перечислите известные вам поисковые машины.
4. Для чего в некоторых поисковых системах используется расширенный поиск?
5. Каким логическим оператором связаны ключевые слова в простом запросе в рассмотренных поисковых машинах?
6. Как в поисковой машине Яндекс осуществить поиск точной формы слов?
7. Какие области поиска можно определить в Яндекс?
8. Как в Яндекс указать расстояние между ключевыми словами?

Дополнительная часть

9. Что обозначает двойное & (&&) в запросе Яндекс?
10. Как правильно в запросе Яндекс поставить знаки «+» и «-»?
11. Как в поисковых машинах осуществить поиск по точной фразе?
12. Для чего используются скобки в запросах?
13. Что обозначает двойное | (||) в запросе Rambler?
14. Как выполнить поиск фразы к заголовке в Aport?