

Тема — Деятельность в сети Интернет

Чтобы найти необходимую информацию во Всемирной паутине, мы можем указать адрес нужного документа. Но чаще всего обращаются к услугам поисковых систем. **Поисковая система** — это программно-аппаратный комплекс, предназначенный для поиска информации во Всемирной паутине.

Поисковая машина — программная часть поисковой системы; комплекс программ, предназначенный для поиска информации.

Поисковые системы располагаются на специально выделенных компьютерах с мощными каналами связи. Ежеминутно они обслуживают огромное количество поисковых запросов клиентов.

По принципу действия различают несколько типов поисковых систем, а именно:

- поисковые каталоги, управляемые человеком;
- системы, использующие поисковых роботов;
- гибридные поисковые системы.

Поисковые каталоги (веб-каталоги или тематические каталоги) одержат базу данных ссылок на веб-сайты, распределённых по отдельным тематическим рубрикам. Такие каталоги заполняются специалистами вручную. Поиск в них осуществляется спуском по дереву каталога. На главной странице поискового каталога располагается список основных тематических разделов (рубрик). Например: финансы и экономика, государство и право, наука и образование, компьютеры, политика и т. д. Каждое имя в этом списке является внутренней гиперссылкой, т. е. ссылкой на внутренний документ сайта, поискового каталога. Щелчок мышью на гиперссылке вызывает на экран список заголовков следующего подраздела и т. д. Перемещаясь по дереву внутренних гиперссылок каталога, пользователь в конечном итоге получает список внешних ссылок на искомые Web-документы.

Поисковые каталоги заполняются вручную специалистами, поддерживающими данную службу. При ручном способе отбора невозможно получить в каталоге исчерпывающую информацию о ресурсах Сети, и в этом состоит основной недостаток каталогов. Однако такой способ классификации исключает ссылки на случайные документы, не имеющие отношения к указанной теме. Таким образом, поисковые каталоги при относительно небольшом охвате ресурсов Сети обеспечивают хорошее качество подборки документов.

Действие поисковых систем, использующих поисковых роботов (другое название этого вида поисковой службы — поисковые указатели), основано на постоянном, последовательном изучении всех страниц всех сайтов Всемирной паутины. Для каждого документа составляется его поисковый образ — набор ключевых слов, отражающих содержание этого документа — индексные списки. Поисковые машины строят индексные списки автоматически.

Основной составляющей поисковых машин являются программы просмотра Web-документов, которые называют по-разному: роботами, червяками, пауками и пр. Непрерывно, днем и ночью, они сканируют все информационное пространство WWW, просматривая все документы, определяя в них ключевые слова и записывая в свою базу индексов данное слово с указателем на документ, в котором оно присутствует. В

связи с постоянным обновлением информации, поисковые системы периодически возвращаются к ранее изученным страницам, чтобы обнаружить и зарегистрировать изменения. Информация о ключевых словах исследованных таким образом страниц сохраняется в поисковой системе.

Гибридные поисковые системы сочетают в себе функции систем, управляемых человеком, и систем, использующих поисковых роботов.

Число поисковых указателей превышает число поисковых каталогов. На данный момент самой популярной в мире поисковой системой является Google, а крупнейшей отечественной поисковой системой — Яндекс.

Работа ПС основывается на формировании запроса, по которому происходит отбор нужных документов из базы данных, хранящейся на сервере. Запрос формируется с помощью ключевых слов (одного или нескольких). Результаты поиска выдаются пользователю в виде списка адресов (гиперссылок) и краткой аннотации к ним.

Ключевое слово — это лексическая единица, являющаяся существительным, прилагательным, глаголом, числительным, наречием или местоимением естественного языка, которая в наибольшей степени отражает содержание всего искомого документа. При формировании запроса могут использоваться не только отдельные ключевые слова, но и словосочетания, состоящие из нескольких ключевых слов.

Запрос — это набор соединенных операторами ключевых слов, с помощью которых поисковая система автоматически ведет поиск и отбор необходимых документов. Другими словами, запрос — это инструкция (команда) для ПС на поиск нужных документов.

Поиск — это процедура отбора нужных документов, хранящихся в сети.

В большинстве поисковых систем реализовано три основных типа поиска:

1. *Поиск по любому из слов* — результатом является огромный список всех страниц, содержащих хотя бы одно из ключевых слов. Применяется, когда пользователь не уверен в ключевых словах.

2. *Поиск по всем словам* — в этом режиме формируется список всех страниц, содержащих все ключевые слова в любом порядке.

3. *Поиск точно по фразе* — в результате составляется список всех страниц, содержащих фразу, точно совпадающую с ключевой (знаки препинания игнорируются).

Чтобы поиск стал более продуктивным, во всех поисковых системах предусмотрены специальные языки формирования запросов со своим синтаксисом. Эти языки во многом похожи. Выяснить особенности определённого языка можно с помощью справочной системы, входящей в состав поисковой машины.

Рассмотрим задачу.

В языке запросов некой поисковой машины для обозначения логической операции ИЛИ используется символ «|», а для логической операции И — символ «&».

В таблице приведены запросы к поисковому серверу. Расположите запросы в порядке возрастания количества страниц, которые найдет поисковый сервер по каждому запросу.

1	принтеры продажа
2	принтеры & продажа
3	принтеры & сканеры & продажа
4	принтеры продажа сканеры

С подобными задачами вы встречались в курсе информатики основной школы и знаете, что условие «логическое И» приводит к уменьшению количества найденных страниц. Соответственно, применение нескольких «логических И» еще более уменьшает число найденных страниц. В отличие от «И», «логическое ИЛИ» приводит к увеличению числа найденных страниц. Соответственно, применение нескольких «логических ИЛИ» еще более увеличивает количество найденных страниц.

Наименьшее количество страниц будет найдено по запросу **принтеры & сканеры & продажа**, (принтеры и сканеры и продажа) так как в документах должны встречаться все три слова.

Наибольшее количество страниц будет найдено по запросу: **принтеры | продажа | сканеры** (принтеры или продажа или сканеры). В этом случае достаточно хотя бы одного слова в документе. По запросу **принтеры & продажа** будет найдено меньше документов, чем по запросу **принтеры | продажа**, так как «логическое И» требует присутствия каждого слова из запроса.

Ответ: принтеры & сканеры & продажа

принтеры & продажа

принтеры | продажа

принтеры | продажа | сканеры

Рассмотрим ещё одну задачу:

В языке запросов некой поисковой машины для обозначения логической операции ИЛИ используется символ «|», а для логической операции И — символ «&». При попытке осуществить запросы к поисковому серверу было найдено следующее количество страниц по каждому запросу.

Запрос	Количество найденных страниц (тысяч страниц)
Тыква	3500
Арбуз	2800
Тыква & Арбуз	1400
Тыква & Дыня	900
Арбуз & Дыня	1700
Тыква & Арбуз & Дыня	500
Тыква Арбуз Дыня	7200

Сколько страниц будет найдено по запросу **дыня**?

Для решения задачи воспользуемся формулой включений и исключений из комбинаторики.

Заметим, что для трех переменных (X, Y, Z) формула включений и исключений выглядит так. $X|Y|Z = X+Y+Z - (X\&Y + X\&Z + Y\&Z) + X\&Y\&Z$

Обозначим количество страниц, найденных по запросу **дыня**, через X. И подставим данные в формулу:

$$7200 = 3500 + 2800 + X - (1400 + 900 + 1700) + 500$$

$$7200 = 2800 + X$$

$$X = 4400.$$

Как бы точно вы ни пытались составить запрос, почти наверняка вам будут представлены документы, не удовлетворяющие вашу информационную потребность. Информация, соответствующая информационной потребности, называется **релевантной**. Эффективность поиска обычно оценивают по двум параметрам — полноте поиска и точности поиска.

Полнота поиска — это отношение числа выданных релевантных документов к общему числу релевантных документов, имеющихся в распоряжении поисковой системы. В идеале это число должно было бы равняться 1. На практике такое почти никогда не достигается. В реальных поисковых системах коэффициент полноты поиска может достигать значений 0,7—0,9.

Точность поиска — это отношение числа выданных релевантных документов к общему числу документов, выданных системой по данному запросу. Значение этого параметра колеблется от 0,1 до 1. Как вы понимаете, здесь многое зависит от того, насколько хорошо был составлен запрос пользователем

Интернет предлагает колоссальное количество возможностей. Но есть и большая доля информации, которую никак нельзя назвать ни полезной, ни надёжной. Пользователи Сети должны мыслить критически, чтобы оценить точность материалов, поскольку абсолютно любой может опубликовать информацию в Сети.

На что необходимо обращать внимание при оценке достоверности информации, полученной в результате поиска в сети Интернет?

Выясните репутацию сайта, на котором размещена информация, представляющая для вас интерес. Проверенные данные публикуют официальные сайты государственных, коммерческих, научных и других структур, являющиеся первоисточниками информации. Ответственность за любую опубликованную ими информацию несут ресурсы, имеющие свидетельство о регистрации средства массовой информации. Избегают недостоверной информации известные ресурсы, занимающие высокие места в соответствующих рейтингах. Представление о репутации сайта можно получить и по имеющимся в сети отзывам об этом ресурсе. Если веб-сайт не обладает широкой известностью, то следует обратить внимание на следующие моменты:

- указано ли, для кого предназначен ресурс и какова цель его создания;
- насколько регулярно обновляются данные на веб-сайте;

- не устарела ли информация (узнать дату размещения материалов);
- не требуют ли разработчики веб-страницы ввода ваших личных данных.

Поинтересуйтесь информацией об авторе представляющего интерес материала. Следует убедиться, что на веб-странице приведены данные об авторе, в том числе описание его квалификации и контактная информация. Можно попытаться найти и познакомиться с другими работами этого автора, комментариями отзывами читателей на его работы.

Попытайтесь проверить фактический материал. Любые фактические и статистические данные имеют источник. Хорошо, если ссылки на авторитетные источники имеются на страницах заинтересовавшего вас сайта. Если таких ссылок нет, то данные можно выборочно сверить с официальными источниками самостоятельно. Если обнаружится, что какие-то данные не согласуются с данными официальных источников, то и остальному материалу также не стоит доверять. Хорошо, если данные подаются с разных точек зрения, если они согласуются с тем, что вы изучали в школе или узнали из других источников.

Мы уже довольно подробно обсудили преимущества Интернета как огромного хранилища информации. Но в Сети можно воспользоваться и другими сервисами. Например, **геолокационные сервисы** — это специфический вид услуг по определению местонахождения мобильного аппарата пользователя. Они также позволяют определить точный маршрут при перемещении на персональном и общественном транспорте или пешком.

Геолокационные приложения применяются не только отдельными абонентами, но и туристическими компаниями, маркетинговыми организациями, транспортными и диспетчерскими предприятиями и предприятиями в других видах промышленности. Условно задачу определения местоположения можно разделить на два вида, в зависимости от поставленных задач, — это географическая и социальная геолокация.

Первая основывается на предоставлении абоненту интерактивных карт, позволяющих определить, где человек находится в данный момент, найти на карте пункт назначения и определить оптимальный путь к нему. А задача второй состоит в том, чтобы предоставить информацию об общественных заведениях, социальных объектах, достопримечательностях и тому подобном, — это скорее средство коммуникации. Здесь же можно проследить за передвижением своих знакомых и родственников.

Следующее быстро развивающееся направление — это **интернет-торговля** — коммерческая деятельность в сфере рекламы и распространения товаров и услуг посредством использования сети Интернет. Свою деятельность электронная торговля осуществляет через интернет-магазины. Интернет-магазин является витриной магазина в интернет пространстве, здесь представлен весь каталог товаров с ценой, также все товары распределены в соответствии со своей категорией. Удобство для покупателя состоит в том, что он может осуществлять поиск нужного товара по сайту. Оплата в сфере интернет магазинов чаще всего производится с помощью банковских карт, а доставка осуществляется курьером или почтой. Рассмотрим преимущества и недостатки интернет – магазинов. Преимущества заключаются, во-первых, в том, что покупатель может выбрать товар, не выходя из дома, в любое удобное для него время,

даже ночью, сокращая транспортные расходы. Во-вторых, добавить товар к сравнению по техническим характеристикам и цене и выбрать наиболее подходящий. В-третьих, в интернет-магазине представлен более широкий ассортимент, что привлекает покупателей. Интернет торговля расширяет географию присутствия, так как товар может быть доставлен в любую точку мира. Минусами электронной торговли является, во-первых, то, что покупатель не взаимодействует с продавцом напрямую, и здесь возможно мошенничество. Во-вторых, товар из интернет-магазина покупатель не может увидеть, поэтому есть вероятность, что при его получении покупателя ждёт разочарование. Так как интернет торговля начала развиваться сравнительно недавно, к минусам можно отнести несовершенства и недочёты системы, например, задержка доставки или оплаты.

Онлайн-бронирование — бронирование через Интернет, в интерактивном режиме. Термин «онлайн-бронирование» применяется по отношению к бронированию номеров в гостиницах, билетов (авиа, железнодорожных, автобусных и т. п.), мест в ресторанах и театрах, прокату автомобилей и т. д.

Общие принципы онлайн-бронирования:

1. Поиск доступных предложений. Пользователь на сайте системы бронирования выбирает маршрут, даты, стоимость, количество человек, расположение, и другие параметры, в зависимости от предмета онлайн-бронирования. Например, для бронирования отелей характерны следующие основные критерии отбора: страна, город, отель, дата заезда, дата выезда, количество человек, количество и возраст детей.

2. Заполнение формы с контактной и платежной информацией. Осуществление платежа. Обычно при прохождении платежа деньги сначала блокируются на карте пользователя, а затем, при условии подтверждения правильности всех данных, списываются.

3. Получение документа, подтверждающего бронирование. Клиент получает документ установленного образца об успешном завершении бронирования, который гарантирует ему оказание услуги. Это могут быть электронный авиабилет, ваучер на заселение и другие.