

1. Язык запросов поискового сервера.

1.1. Сужение и расширение поиска

Как отмечалось, сужение поиска — основной метод выделения нужной информации из невообразимой массы документов в Интернете. Но бывает и так, что предоставленная в результате поиска информация недостаточна, так что приходится иногда поиск и расширять. В большинстве поисковых систем сужение и расширение поиска обеспечивается применением логических операторов или включающих и выключающих условий. Но не меньшее значение имеет просто правильный выбор ключевых слов для поиска.

1.2. Использование логических операторов

Логические операторы хорошо знакомы всем, кто хоть немного изучал программирование. Есть три таких оператора (AND, OR и NOT), наименование и назначение которых применительно к поиску страниц в Интернет, указано ниже:

Оператор

Слово 1
AND Слово 2
Слово 1 OR
Слово 2
NOT Слово

Название

Логическое И
Логическое ИЛИ
Логическое НЕТ

Действие

Ищет страницы со Словом 1 и Словом 2
Ищет страницы со Словом 1 или Словом 2
Ищет страницы в которых нет Слова

Нетрудно заметить, что оператор AND сужает поиск, тогда как оператор OR его расширяет. Заметим, что если указано просто несколько слов, то поиск идет одновременно по всем словам, то есть так, как будто между ними стоит оператор OR. Оператор NOT позволяет исключить из поиска страницы, содержащие стоящее за ним слово. Все операторы могут использоваться и с большим числом слов. Доказано, что этими тремя операторами можно описать любые логические соотношения между словами.

1.3. Включающие и выключающие условия

В некоторых поисковых системах используются, включающие и выключающие заданное слово из поиска, знаки + и —. Ниже приведены примеры использования этих знаков:

- +Книги+фантастика—проза — ищутся страницы со словами «Книги» и «фантастика», но исключаются книги со словом «проза».
- +Медведь+белый— гималайский — ищутся страницы о белых медведях и отсеиваются гималайские мишки,
- +Программы-Microsoft — ищутся страницы о программах, за исключением программ фирмы Microsoft.
- +Города-Москва-Киев-Минск — ищутся города, за исключением Москвы, Киева и Минска.

В некоторых поисковых системах есть свои особые правила поиска. Поэтому перед широким применением любой поисковой системы полезно внимательно ознакомиться с принятыми в ней правилами поиска.

1.4. Манипуляции с URL-адресом

Нередко поиск конкретных сведений не приводит к успеху. Возможно, будет получено сообщение поисковой системы о том, что нужных сведений нет. Довольно часто выводится сообщение об ошибке 404 «Документ с указанным Вами адресом на сервере отсутствует». В таких случаях ситуация вовсе не выглядит безнадежной.

Прежде всего проанализируйте URL-адрес, запрошенного вами, документ. Если он многоступенчатый (с множеством разделителей /), то воспользуйтесь методом усечения адреса справа. Обычно, это ведет к расширению поиска вплоть до того, что останется адрес только начальной страницы Интернет - сайта. Нередко такой прием позволяет обнаружить нужный документ просто на другой странице сайта. Если вы вышли на начальную страницу сайта, то внимательно ознакомьтесь с ее возможностями. Часто сайты имеют свои собственные средства поиска и с их помощью можно попробовать найти иные документы.

2. Технология поиска.

2.1. Предварительный сбор информации — основа поисковых систем

Полноценный поиск в океане информации Интернета был бы просто невозможен, если бы не были созданы специальные **поисковые системы Интернета**. Многие из таких систем за секунды могут *выдать* вам список из многих тысяч страниц, относящихся к интересующей вас теме поиска. Чаще всего она задается набором ключевых слов в специальной панели (поле) поиска.

Многим такая скорость поиска кажется просто невероятной. И впрямь, трудно себе представить в столь медленной сети, как Интернет, поисковую систему так быстро отыскивающую ключевые слова во всей всемирной паутине Интернета. На самом деле этого и не происходит! Эти системы вовсе не тратят нужные сведения по всей сети в данный момент времени. Они делают это заранее!

Современные поисковые системы — это крупные серверы с мощными базами данных, взаимодействующие друг с другом и связанные между собой другими крупными узлами Интернета высокоскоростными световолоконными или спутниковыми линиями связи. Время от времени эти системы сканируют паутину Интернета в поисках новых страниц и оценки их содержания. Эти данные поступают в базу данных поисковых систем и затем используются для подготовки данных, нужных пользователю. Вы получаете информацию быстро только потому, что она была получена поисковой системой заранее и потому, что поиск идет не по всей сети Интернета заново, а в специальной базе данных с быстрым поиском, реализованной на быстрых и мощных компьютерах.

2.2. Индексные поисковые системы — пауки и черви

Вообще говоря, есть два класса поисковых систем: индексные и составители каталогов. Индексные системы по введенным пользователям ключевым словам отыскивают в своей базе данных все URL-адреса страниц Интернета и ссылки в них. Совокупность ссылок и адресов оформляется в виде отчета о поиске, который и направляется пользователю. Есть системы, которые добросовестно находят максимально возможное число страниц, а есть более разумные, которые учитывают заданные пользователем ограничения на поиск.

Индексные системы имеют специальные средства для периодического просмотра страниц в Интернете и обновления своей базы данных. Их именуют как не очень симпатичных животных, пауками (spiders), червями (worms), гусеницами (crawlers) и так далее. Впрочем, как и эти, довольно полезные твари, данные системы делают неторопливо свое важное дело: накапливают информацию о новых страницах и об изменении старых страниц, уже хранятся в базе данных поисковой системы.

2.3. Поисковые системы — составители каталогов

Другой вид поисковых систем — составители каталогов. Такие системы ищут URL-адреса по разделам Web-страниц, систематизированных по иерархическим признакам. В этом случае

анализируется не все содержание Web-страниц и даже не все страницы. Словом, круг поиска оказывается более узким.

В настоящее время грань между этими двумя классами поисковых систем становится все более и более условной. И те и другие системы, в ответ на запрос пользователя, выдают каталог обнаруженных в Интернете страниц иногда в виде списка заголовков, иногда даже с аннотациями. Они нередко являются просто несколькими первыми строками текста на обнаруженной странице.

2.4. Общие приемы эффективного поиска

Здесь мы рассмотрим некоторые общие приемы эффективного поиска информации. Некоторые из них носят общий характер, другие просто обобщают практику использования средств поиска информации. Итак, рекомендуется придерживаться следующих правил:

- как можно конкретнее и точнее сформулировать, что именно вы собираетесь искать;
- определите системы поиска наиболее близкие к выбранной тематике;
- наметьте ключевые слова по теме поиска, исключив из них слова о широко известных объектах поиска;
- установите в качестве начальной страницу, содержащую ссылки на наиболее полезные для вас данные и системы поиска информации в Интернете;
- если у вас часты отключения из-за потери активности модема, включите утилиту имитации активности модема (например, Ponger или Ponger Pro);
- запустите в работу браузер;
- при поиске хорошо известных фирм и организаций попробуйте начать поиск с указания их имени, например, **www.intel.com**, **www.amd.com**, **www.microsoft.com**, **www.Pcweek.com** и так далее;
- начните поиск с просмотра баз данных по ресурсам Интернета, например, знаменитых «желтых» страниц на CD-ROM или в Интернете, каталогов с адресами сайтов определенной направленности и т.д.;
- познакомьтесь с десятком-другим поисковых систем, и установите ссылки на наиболее полезные из них в папке «Избранные»;
- если вас интересует самая общая краткая информация об объекте поиска, можно использовать поисковые возможности браузера, вводимые кнопкой **Поиск**;
- серьезный поиск лучше производить, загрузив отдельно страницу с выбранной системой поиска, с тем, что бы она занимала всю область просмотра браузера;
- используя подготовленные ключевые слова, проведите общий поиск и оцените объем получаемой информации (например, число страниц, ссылки на которые вам предоставлены);
- если объем информации велик, проведите сужение поиска более точным указанием ключевых слов и применением логических операторов OR и NOT, а также выключающих условий;
- если объем информации мал, проведите расширение поиска, расширив набор ключевых слов, применяя логический оператор OR, включающий условие и усечение URL-адреса справа;
- при поиске книг, статей и прочих документов с известным названием укажите его в кавычках;
- при затруднениях в поиске начните его с посещения крупных электронных библиотек и просмотра электронных энциклопедий;
- учтите, что события и факты последнего времени вы скорее найдете в группах и каналах новостей, чем в поисковых системах;
- если вы ищете свободно распространяемое программное обеспечение, обратите внимание на раздел сайтов «download», обратитесь также к файловым серверам, некоторые из которых предоставляют десятки, и даже сотни тысяч подобных программ;
- используете специализированные узлы и серверы Интернета для поиска сведений профессионального характера, например, сервер **www.gov.ru** Правительства Российской Федерации для поиска Постановлений Правительства, **www.unicor.ru**; для сведений из университетов, **www.osi.ru** при поиске мероприятий, поддерживаемых фондом Сороса и так далее. Придерживаясь этих простых рекомендаций, вы довольно скоро заметите, что поиск

не только ведет вас к вашей цели, но и дает вам массу новой полезной информации, о существовании которой, вы, возможно, даже и не догадывались.

Фиксируйте немедленно все страницы, где вы обнаружили полезные для себя сведения. В первую очередь используйте для поиска специализированные узлы и серверы, а также базы данных по ресурсам Интернета. Помните, что отложив запись (обычную или электронную) заинтересовавшей вас Интернет - страницы на потом, вы рискуете уже никогда не найти эту страницу.