

Как правильно искать в Интернете

В этой инструкции по поиску в Интернете Вы узнаете:

- как эффективней и быстрее искать в Интернете;
- как правильно сделать запрос информации;
- как сужать и расширять запросы;
- что такое индексация сайта и ее особенности;
- чем каталоги лучше и хуже поисковых систем;
- адреса поисковых систем и их синтаксис.

Объем информации даже в Русском Интернете занимает сотни гигабайт на тысячах серверах, расположенных не только на территории России, но и в Европе, США, Украине, и т.д. На первый взгляд здесь можно найти что угодно, однако когда дело доходит до поиска, то сразу же появляются проблемы.

Для начала рассмотрим поисковые системы и каталоги.

[RAMBLER.ru](http://rambler.ru)

Хорошая поисковая система, но игнорирует знаки препинания и цифры в теле запроса. Ищет даже во FLASH (по крайней мере так говорят)

[APORT.ru](http://aport.ru)

В отличие от RAMBLER воспринимает любые символы в теле запроса, как и RAMBLER игнорирует предлоги в запросе.

[YANDEX.ru](http://yandex.ru)

Самая большая поисковая система в России по количеству ресурсов, но не самая популярная. Отлично ищет по сайтам, разработанным не только на обычном HTML, но и на программном обеспечении, обеспечивая тем самым самый глубокий поиск.

[YAHOO.com](http://yahoocom.com)

Самый большой в мире каталог ресурсов Интернет, но имеет не совсем удачную реализацию поиска по каталогу.

[LIST.ru](http://list.ru)

Популярный Российский каталог ресурсов.

- С чего начать поиск?

Поиск надо начать с определения ключевых слов. Например, нас интересует адрес фирмы, которая бы нам продала и смонтировала отопительный котел средней мощности, например фирмы Мора. Поскольку мы находимся в Екатеринбурге, то нас интересует именно фирмы в Екатеринбурге.

Получаем следующие ключевые слова:

Отопительный котел Mora, Екатеринбург.

'Отопительный котел' - именно так это должно быть написано в прайс-листе на сайте. Если, например, там написано 'Продажа отопительных котлов', то поисковая система нам покажет и этот сайт, так как все современные поисковые системы имеют средства лексического разбора запросов.

'Мора' - фирма-производитель. Тоже обычно содержится в прайс-листах на сайте. Заметьте, 'средней мощности' мы не пишем, если точно не знаем. Если мы это укажем, то при строго заданном запросе поисковая система скорее всего не найдет ни одного сайта с таким содержанием, потому, что редко в прайс-листах используют такие заголовки. Ведь обычно прайс-листы содержат в .ZIP файлах, которые недоступны для поиска. А вот краткий перечень продукции и фирм-производителей приводят все.

'Екатеринбург' - местоположение. Обычно адрес всегда указывается на сайте, поэтому это тоже необходимо использовать, чтобы поисковик не выдал в списке всех Российских продавцов данного товара.

- Осталось просто ввести ключевые слова в строку поиска и нажать на Enter?

- Можно, но результат будет не самым эффективным.

- Почему?

- Потому что по умолчанию поисковик настроен на максимальные возможности поиска и без использования синтаксиса поисковая система выдаст все сайты, где находится хотя бы одно из указанных слов.

- Но ведь это обычно работает!

- Да, но... Во-первых следует отметить, что алгоритм построения списка очень сложный, то есть сортировка происходит по заданному сценарию. Обычно в начало выносятся сайты с максимальным количеством совпадающих слов. То есть, при определенных условиях вперед может попасть страница с 25 словосочетаниями 'Отопительный котел' и 10 словосочетаниями 'Мора', а страница, содержащая все слова запроса может оказаться в середине списка из 25000 страниц и у Вас не хватит терпения до нее дойти. Во-вторых, в таком гигантском списке Вам все равно не выбрать все подходящие сайты - результат поиска придется сужать.

- Понятно. Как тогда мне правильно создать запрос?

- Использовать возможности синтаксиса.

Синтаксис у разных поисковых систем немного отличается, но ненамного.

1. Поисковая система не чувствительная к регистру текста запроса, до тех пор, пока Вы не поставите первую букву заглавной. Например, на запрос 'урал' Вы получите список всех страниц, которые содержат слова с указанным сочетанием букв в любом регистре (например: 'Урал', 'Уралмаш', 'уральские'), на запрос 'Урал' получите страницы, слова на которых начинаются с 'Урал' (например: 'Уралмаш'), но первые в списке все же будут сайты со словом 'Урал'.

2. Если запрос поставить в двойные кавычки, то поисковая система будет искать указанное словосочетание (точное совпадение) или только те страницы, где присутствуют все указанные в кавычках слова, а не хотя бы одно из указанных, как по умолчанию. Причем обычно, даже ищется только точное совпадение (без описанного выше лексического разбора).

3. В запросах можно использовать круглые скобки и булеву логику (OR, AND, NOT или соответственно символы '|', '&', '!'), например: ((Урал OR Сибирь) NOT Екатеринбург) - найдет все страницы с содержанием слов 'Урал' или 'Сибирь', но не содержащих слово 'Екатеринбург'.

4. Можно использовать специальные символы '*' и '?' для замены части слова или буквы в слове соответственно (как в MS-DOS).

Более сложные конструкции поиска уже зависят от системы. Более подробную инструкцию по каждой системе Вы найдете в конце статьи.

У нас получился следующий запрос:

Отопительный AND котел AND мора AND Екатеринбург

или

Отопительный && котел && мора && Екатеринбург

Обратите внимание, что мора - с маленькой буквы. Это потому, что возможно написание 'MORA' - многие пишут такие имена полностью заглавными буквами.

- Что делать, если даже на такой запрос много подходящих страниц?

Если на запрос поисковая система Вам все равно выдает список и сотен и тысяч страниц, - сужайте поиск дальше. Для данного примера можно попробовать использовать в качестве дополнительных условий тип котла по виду топлива и числу контуров. То есть можно попробовать ввести следующий запрос:

Отопительный AND котел AND газовый AND мора AND Екатеринбург AND монтаж

или

Отопительный && котел && газовый && мора && Екатеринбург && монтаж

- А если наоборот - ни одной страницы не найдено?

Если ни одной страницы не найдено, тогда можно попробовать перефразировать запрос, подобрать другие ключевые слова, например не "строительство зданий" а "сооружение зданий" или "строим здания" проверить орфографию в запросе и расширить условия поиска, то есть уменьшить количество слов в запросе с логическим **И** и добавить больше ключевых слов с логическим **ИЛИ**, например "Екатеринбург && (строительство || строим || строит)" - принцип, думаю, Вы поняли. Главное - не забывайте, что когда Вы делаете запрос, например, ядл поиска строительной компании, Вам надо писать фразу не ту, которая Вам подходит, как, "кто построит дом", а слова, которые могут находиться на сайтах тех, кто это делает, например "строительно-монтажные работы" (вид деятельности) или просто "строительство".

- Мы пробовали - не помогает. Как быть?

Если все же поиском Вы ничего не нашли, то можете попробовать аналогичный запрос в другой поисковой системе. Возможно, что не все страницы нужного Вам сайта проиндексированы, и поиск на них невозможен.

- Что значит 'проиндексирован', и почему может помочь поиск в другой системе?

Во первых, для поиска доступны только страницы, которые были специально добавлены или же найдены по ссылкам с уже известных поисковой системе страниц. Все найденные страницы обрабатываются специальной программой, которую обычно называют роботом. Этот робот индексирует страницу, то есть создает ее некий уникальный образ, который обычно содержит такие параметры страницы как адрес страницы, слова, название страницы, дата последней обработки страницы роботом, тэги и верхнюю часть страницы (ее фрагмент).

Отсюда следует, что не все страницы в Интернете могут быть найдены и тем более проиндексированы. Так как алгоритмы поиска новых страниц и индексации у разных поисковиков отличается, то поиск в разных системах даст разный результат. Страницы, добавленные недавно, а также изменения внесенные в них могут быть некоторое время недоступны. А если вебмастер не постарался, то новые страницы, возможно, вообще никто и никогда не найдет. Индексация страниц проходит с различными интервалами - от одного дня до месяца и более. Интервал индексации зависит от количества страниц на сайте, настроек страниц, местоположения на сайте и от самой поисковой системы.

- Ладно, что насчет каталогов?

Обычно все каталоги доступны для поиска из поисковых систем, но здесь информации будет явно меньше чем на конечных сайтах, так как содержат только их краткое описание. Обычно каталоги содержат собственные поисковые системы, но поиск осуществляется только внутри самого каталога. Каталоги удобны тем, что они содержат уже упорядоченную систему. И например, для нашего запроса можно войти в каталог и пройти по следующему пути:

Предприятия ->

Строительство ->

Отопительное оборудование ->

Свердловская область.

И там Вы найдете, хоть и не все, адреса сайтов с кратким описанием. По этой причине лучше пользоваться большими или местными каталогами. Еще одна проблема каталогов - у них нет общих стандартов и найти местоположение нужного раздела может оказаться более сложным, чем поиск в поисковой системе.

- Это все, что нужно знать?

Самое главное, о чем стоит помнить всегда, это терпение - на поиск надо время. И не забывайте, что в Интернете есть далеко не все - иногда ту же книгу дешевле купить в обычном магазине, чем занимать выделенную линию для поиска и перекачивания. Для успешного поиска нужен опыт - у меня это стало хорошо получаться только после первых 100 часов работы в Интернете.

Бурусов Д.В.

www.burusov.ru