

Поисковые системы

Послан armandr - 08.05.2010 11:44

Для поиска необходимой нам информации в интернете мы пользуемся поисковиками, и я думаю, что так как мы здесь чайники, то часто просто сложно во всём этом объёме найти именно то, что необходимо. Существует множество правил набора запросов. Хотелось бы научиться профессиональному поиску в интернете. Понять в чём различия поисковиков? Понять принцип работы и нюансы пользования каждым из них. Может подскажете ссылку на сайт, где можно прочитать о пользовании поисковиками? А может напишите статью?

=====

RE: Поисковые системы

Послан Night_hunter - 08.05.2010 17:11

Во-первых, для того, что бы разобраться, как ищет поисковая машина и какие операторы (+ прочие навороты) необходимо использовать для улучшения ее поисковых возможностей - рекомендуется потратить свое время на изучение справки (каждый сервис снабжен этими документами). Нет ничего лучше, чем работа с первоисточником! Там будут даны конкретные примеры и прочие полезные объяснения. Это что касается того, где взять информацию непосредственно по каждой поисковой системе.

Во-вторых. В зависимости от того, что вам необходимо найти, поисковые запросы можно подразделить на два типа:

- а. поиск по всему Интернету;
- б. поиск в заранее определенной части Интернета.

Поясним на конкретном примере пункт "б". Допустим, вы хотите отыскать в сети музыку. И не просто песню, а целый диск. Отыскать и скачать. Было такое?

Так вот. В данный момент времени, самый эффективный поиск альбомов тех или иных исполнителей, можно легко реализовать, если заранее знать, где такая музыка может находиться. Поясним это на вот каком примере.

Блогерская платформа гугла - blogspot - это идеальное место для того, что бы делиться с другими людьми тем, что тебе нравится, а бесплатные аплоуд хосты (rapidshare, zshare, sharebee и т.п.) - это то место где можно складировать свои файлы (рекомендуем, кстати, статью про фришные аплоуд-хосты!). Музыка (в раровском архиве) заливается на аплоуд хост, на блоге дается линк и все - пост готов. Миллионы людей давно смекнули, что простой и в то же время очень удобный blogspot - это идеальный вариант для подобных расшариваний. На сегодняшний день музыкальных блогов на блогспоте так много, что пересчитать их вряд ли кому удастся. Найти же на них можно все что душе угодно. Более того - гугл очень любит блоги находящиеся на "родных" серверах и постоянно поставяет их в серп.

Что из этого следует? Из этого следует, что вместо того, что бы извращаться и искать во всей сети тут или иную пластинку, достаточно попытаться поискать ее на гугловских блогах, которых, как мы уже сказали - не счесть.

Как это сделать? Все просто - пишете название исполнителя, название альбома и после этого добавляете слово blogspot. Гугля выдаст в первую очередь те самые блоги, где может

находиться требующийся вам диск. Этот нехитрый способ опробован на практике и почти всегда (за редкими исключениями, разумеется) дает отличные результаты. Все просто и, как говорится, - не надо придумывать велосипеды!

Тоже самое можно сказать и про поиск по трекерам. Достаточно вписать после названия альбома (или, допустим, фильма) слово torrent и мы получим весьма заманчивые результаты, которые приведут вас не в онлайн-шопы, а на конкретные трекеры с которых можно совершенно бесплатно скачать требующиеся файлы.

Ну вот. Теперь перейдем к более сложным вещам - как найти в Интернете, среди миллиардов страниц, требующуюся информацию и при этом затратить на это дело как можно меньше времени? Допустим, вам необходимо найти, расписание движения поездов или текст той или иной книги. Тут поиск по блогам или по трекерам уже отменяется, т.к. он врядли даст положительный результат. В нашем случае надо выйти на какой-то конкретный сайт, который будет содержать четкий контент. Для такого поиска существуют специальный поисковый язык, который дает машине понять, что же нам от нее надо.

Итак, приступим. Как было сказано выше, разбираться с этими вещами мы будем на примере Гугловского поисковика - он самый совершенный, самый посещаемый и, что очень важно - никогда не подводит. У гугла найдется все и сразу (а не со временем, как это обещают инженеры из Яндекса). Каждый пользователь Гугла должен знать, что все "пробелы" между словами, которые он вводит в строку поиска, это не просто пустое место, а оператор. Название же он носит простое - AND ("и" по-русски). Если мы ищем расписание поездов, то вбив в поисковик "расписание поездов" мы даем понять машине что ищем два ключевых слова в текстах проиндексированных гуглем страниц. Это очень важно понимать. Вы спросите - и что же тут важного? А все дело в том, что вместо этого самого "AND", которое по умолчанию не видно, можно поставить кое-что другое и это "кое-что" так изменит выдачу, что ого-го!

Перечислим же все эти операторы.

1. Оператор "OR" (тобишь по-русски - "или"). Запрос с таким оператором следует использовать тогда, когда вы хотите отыскать, что-то в нескольких вариантах. Допустим, вы хотите найти сразу расписания двух видов транспорта - поездов и самолетов. В этом случае, запрос должен выглядеть так - "расписание поездов OR самолетов". Поэкспериментируйте с этим запросом сами - вбейте его с оператором, а потом без! Есть разница? :-) Кстати, вместо "OR" можно использовать "|" (без кавычек, разумеется - этот символ на клавиатуре расположен там же, где и слэш в русской раскладке).

2. Оператор "+". Это очень важная штука, т.к. используя "+" можно весьма сильно преуспеть в нахождении той или иной информации. Используйте его тогда, когда хотите принудительно включить в результаты поиска какое либо слово. Например, запрос по типу "автомобили +горбатый запорожец" даст вам только Горбатые Запорожцы, но никак не BMW или Audi. Оператор "+" можно использовать несколько раз - тем самым вы все более и более усовершенствуете свой запрос и найдете все то, что вам необходимо.

3. Оператор "-" (минус). Воспользовавшись этой логической категорией вы исключите из своего запроса очень много мусора. За примером далеко ходить не надо, т.к. все и всюду хотят что-то продать, а вот бесплатно делиться что-то не очень. Минусом можно выбрасывать из серпа всевозможные магазины, которые лезут в топ по таким запросам, как скажем "книги", "фильмы", "диски" и т.п. Например, вы хотите скачать книги Харуки Мураками и постоянно попадаете на магазины, которые ими торгуют. Хотите покупать - пожалуйста, но если требуется бесплатный

файл с текстом, то возьмите и исключите из серпа все что связано с коммерцией! "Мураками -купить -заказать +скачать" - это будет вполне рабочий запрос и, скорее всего, вы добьетесь своей цели.

4. ("") - так можно искать точную фразу. Вообще-то в настройках поиска всегда имеется возможность выбрать режим поиска, но зачем туда забираться, когда можно просто набрать в кавычках то, что вам надо? Такие запросы следует делать тогда, когда вы хотите найти, скажем, кусок текста или какое-то название (например, фирмы, продукта и т.п.). По правде сказать, гугля и так понимает что мы ищем - даже без кавычек она выкидывает в топ точно соответствующие запросу страницы, а уже потом собирает по крупицам разрозненные данные. Но оператор ("") полезен вот в каком случае - если вам требуется составить сложный запрос, принудительно исключая что-то из серпа, то кавычками вы можете указывать системе на те слова и выражения, которые должны быть в тексте без каких либо вариаций. Это очень важный оператор и он существенно улучшает результаты поиска. Очень рекомендуется к использованию!

5. "*" - этот оператор поможет вам отыскать имя или словосочетание, часть которого вы забыли. Оператор называется "усечение слова" и используется он редко, но метко. Например, "Иль * Франс", "Дом который построил *" и т.д. и т.п. Просто вместо слова которое вы забыли, ставьте (*) и гугл будет подбирать более релевантное содержание, опираясь на свою индексную базу.

6. Иногда бывает так, что найдя какой-то сайт и пытаясь на него зайти вы обнаруживаете, что сайта-то уже и нет - ссылка есть, а страница не грузится. Такое случается, к слову сказать, весьма часто - сайт, например, может закрыться, переехать или же, оставаясь на месте, попросту имеет битый линк на уже несуществующую страницу. Гугля все это отслеживает и, периодически, избавляется от подобного нерабочего хлама, очищая свой индекс, но делает это не ежесекундно, а как уже было сказано - периодически. В случае, когда в серп попадают битые линки, снипет (т.е. описание) которых весьма заманчив, то следует попытаться получить требующуюся информацию из кэша поисковика. Оператор cache: - вот что нам необходимо в данном случае. Например, вбив в поисковую форму cache:deadsite.com - вы получите версию того самого нерабочего сайта, но не с сервера на котором он находился, а напрямую из индекса гугла. Короче говоря, это очень удобный оператор, т.к. с его помощью можно найти то, чего уже, как бы и нет. Согласны? Не забывайте еще и о том, что пробел между двоеточием и названием домена ставить нельзя - пишем без пробела, точно так, как показано в примере.

7. Если вам необходимо что-то отыскать по расширению файла (например, вам нужна книга в pdf или в txt), то вы можете запросто это сделать, если воспользуетесь оператором filetype: Суть проста - ищем Гоголя в pdf и для этого пишем в поисковой форме такую комбинацию - Мертвые души Гоголь filetype:pdf Можете поэкспериментировать - добавьте туда что-нибудь из вышеперечисленных логических операторов и поиск станет еще более точным.

8. Если какой-то сайт не имеет формы поиска по своим страницам, то не отчаивайтесь - гугл произведет этот поиск за доли секунды! Воспользуйтесь оператором site: Схема такая - пишем необходимое слово или словосочетание, затем ставим оператор site: и название сайта. Например, вам необходимо найти все статьи с нашего сайта в которых упоминается слово "бабло". Запрос выглядит так: бабло site:www.bigfozzy.com

9. Каждая страница в сети имеет свой заголовок (во всяком случае, подавляющее большинство документов в сети - это не Untitled Document) - поиск можно производить и по ним (т.е. по заголовкам html-документов). Для этого существует оператор allintitle: Запрос такого типа, как "allintitle:знакомства в киеве" даст возможность копаться в документах, которые несут именно это название. Напоминаем, что заголовки находятся в html-теге TITLE.

10. Если вам необходимо какое-то определение, то поисковая система очень оперативно его вам выдаст - для этого необходимо применить оператор define: Положим, вы не знаете что такое Энтропия. Вбиваем в гугл так: define:энтропия и получаем ссылки туда, где повествуется об этом термине. Кстати, гуглем можно и не пользоваться - достаточно сразу забраться в Википедию (собственно туда вас и направят).

В заключение еще несколько поисковых операторов, которые могут пригодиться:

- Оператор link:site.com - позволяет посмотреть страницы, которые ссылаются на указанный ресурс;

- Оператор allinurl: - ищет слова, которые находятся только в тексте ссылок

- Оператор inurl: - приблизительно то же самое, что и allinurl: , но с большей свободой действия.

После этого оператора можно вписать любое слово и оно будет искаться по всему документу.

Например: inurl:mp3 купить ("mp3" - ищется в урл, а "купить" - уже по тексту той или иной страницы).

- Оператор related: - ищутся похожие страницы. Например, related:yahoo.com - будут найдены другие поисковики. Оператор очень удобен для того, что бы искать различные аукционы, магазины и прочие сервисы.

:)

=====

RE: Поисковые системы

Послан armandr - 08.05.2010 19:14

Вы дали направление для действий.:woohoo:

Спасибо за время и знания!

=====