

Стратегия и методика профессионального информационного поиска

Приступая к информационному поиску в Интернет, следует всегда помнить несколько основных моментов. Прежде всего, никакие средства навигации - справочники или поисковые машины - не охватывают всего текущего информационного массива Интернет. По некоторым оценкам, даже такие признанные лидеры сетевого поиска как Google, отражают не более трети совокупного содержания Сети. Причина этого - постоянный колоссальный прирост объемов информации в Интернет, который, несмотря на все усилия навигационных служб, содержит огромное число белых пятен.

Помимо быстрого роста и изменения местоположения документов, большинство поисковых систем имеют внутренние ограничения на отражение материалов одного сайта и на объем индексируемой части страницы. Программы-роботы зачастую не идут в глубь сервера дальше определенной директории, что также сокращает число отраженных материалов.

В тоже время многие крупные сайты имеют собственную систему поиска, которая отражает весь их информационный массив. Выявив такие сервера с помощью справочников, можно провести более детальное их обследование, используя локальный поисковый механизм. Например, при поиске сведений о конкретном виде креветки, искусственно разводимой человеком, весьма рациональным будет найти и просмотреть сервера, посвященные в целом аквакультуре, отрасли, занимающейся выращиванием морепродуктов в искусственных теплых водоемах, а при выявлении данных о конкретном заболевании - сервера учреждений, ведущих исследования данной области. Таким образом, для достижения наиболее полных результатов следует применять справочники и поисковые системы в сочетании друг с другом.

Существует также ряд общих требований к поисковой деятельности, соблюдение которых повышает эффективность и экономит время, затрачиваемое на разыскание данных.

Для поиска материалов по крайне узкой специфической тематике стоит начинать с метамашин, дабы сразу получить представление о том, насколько богата данная проблематика представлена в Интернет.

Для получения более полных результатов по сложному запросу (например, там, где есть ограничения не только по содержанию документа, но и по дате обновления или местоположению документов) поиск рекомендуется проводить отдельно в каждой поисковой машине. Поисковые системы имеют сильный разнос в отражении документов и их последовательное использование в значительной степени расширяет охват материала.

При разыскании документов об отдельной стране или на конкретном языке следует отдать предпочтение национальным/региональным поисковым средствам. Так, например, при поиске материалов на испанском языке стоит обращаться не к глобальным, а к испанским поисковым системам, например, Trovator (<http://www.trovator.com>).

Формировать запрос надо максимально точно, используя все возможности механизма составления запроса. Затраты времени на детальное составление поискового предписания окупаются при анализе результатов поиска. При точном формировании запроса процент информационного шума будет намного ниже.

При составлении запроса на текстовый материал всегда стоит мысленно представить, каким может быть предполагаемое содержание документа. При поиске источников об А.С.Пушкине недостаточно просто ввести его фамилию (список результатов будет изобиловать адресами всевозможных учреждений, расположенных на улицах Пушкина в разных городах). Гораздо больший эффект принесет поиск по фамилии в сочетании с названиями произведений поэта. Для разысканий же полных текстов произведений, наоборот, стоит вводить в поисковую строки из произведения, которые однако не являются расхожими фразами (например, для поиска полного текста поэмы "Евгений Онегин" не стоит использовать "Мой дядя самых честных правил", гораздо лучший результат принесет "Когда на бледном небосклоне звезд исчезает хоровод " или другая характерная, но не столь распространенная точная фраза).

При систематическом обращении к поисковым средствам, необходимо постоянно следить за новостями, относящимися к поисковому сервису. В окружающем нас мире вообще, а в цифровом мире тем более, нет ничего вечного. Каждая поисковая система переживает периоды зарождения, расцвета и упадка. Не бойтесь отказаться от использования любимой, но устаревающей поисковой системы и перейти на использование новых поисковых инструментов, обладающих большей эффективностью.

Принципы поиска в Интернете:

Точность и простота запроса

Не пытайтесь полностью описать поисковой системе тему вашего реферата или диплома. Пытайтесь быть лаконичными и точными. Попытайтесь побольше узнать об искомом тексте, обдумать поисковый запрос, при необходимости, подобрать синонимы к ключевым словам!

Важнейшим фактором и залогом нашего успеха является правильность поискового запроса. К примеру, если вы хотите найти информацию о чемпионате по футболу Евро 2012 в Украине, вам следует вводить в строке поиска Евро 2012 Украина, а не просто футбол Украине. Помните, машина ищет именно слова и словосочетания, она не знает, что вы имели в виду, запросив какую-либо информацию, она не будет думать за вас! Также полезно знать, что любой современный поисковик сортирует найденную информацию по релевантности (степени совпадения вашего запроса с информацией на сайте), т. е. самый подходящий вам сайт будет находиться на первом месте и т. д.

Также, обратите внимание, что при поиске был опущен предлог «в». А все это от того, что поисковая система не индексирует и не ищет по предлогам и таким частым словам, как «где», «когда», «кто»; также она опускает знаки пунктуации (исключения составляют лишь очень распространенные комбинации, как \$10 или C++).

Поиск цитаты или точного названия

Как быть с поиском цитат или точных названий фирм, музыкальных групп и пр., если поисковик не берет во внимание сильно распространенные слова? Очень просто. Стоит лишь заключить в кавычки весь запрос, и система тотчас же воспримет его как единое целое, не смея менять слова местами либо опускать какие-либо слова или символы. К примеру, если вы ищете текст песни «Белые розы», просто наберите фразу «что с вами сделали лед и морозы». Также вы можете заставить поисковую машину искать любое слово или символ, поставив перед ним «+». Результатом запроса «Обитель зла +3», будет именно ссылки на сайты о третьей части этого фильма.

Поиск по сайту или домену

Предположим, вы читали интереснейшую статью на каком-либо сайте, однако забыли сохранить ссылку в Избранном и теперь не можете ее найти. И с этой задачей способен справиться наш поисковик. К примеру, Гуглу стоит лишь указать, на каком сайте искать, например: Как купить компьютер `site:mirsovetov.ru`. Данный запрос направит поисковую систему на указанный сайт, и найдет там нужную информацию. В Яндексе такой запрос выполняется немного сложнее: Как собрать компьютер `<< url="mirsovetov.ru"`. Также эта функция позволяет искать информацию в определенном домене, например, для поиска с помощью Гугла информации, расположенной на сайтах с адресом *.ru, можно ввести: Как купить компьютер `site:ru` – что может быть полезно для поиска данных определенного региона.

Исключение из поиска

Поисковую машину можно не только «заставлять» искать нужный текст в Интернете, но и исключать некоторые слова из запроса. Если вы не хотите, чтобы при поиске реферата на тему «История Древнего Египта» вам были предложены документы с его Культурой, просто введите в поисковое поле Google: история древнего египта -культура. В Яндексе вместо знака «-» используется сочетание знаков «~». При таком запросе, слово «культура» было полностью исключено из критериев поиска.

Поиск с учетом регистра

Поисковые системы не учитывают регистр, все заглавные символы воспринимаются машинами как строчные, за исключением использования специального операнда. Таковым является знак восклицания «!», стоящий перед словом. Эта функция очень полезна, чтобы искать села или города с распространенными названиями, к примеру «село !Кошки». В данном случае, поисковик не будет искать сайты, где речь ведется о селе, где живут кошки, а будет вести поиск с учетом заглавной буквы.

Для более структурированного запроса обратитесь в раздел «Расширенный поиск», которым обладают все современные поисковики.