

Как не заблудиться во Всемирной паутине.

Современный студент, вооруженный персональным компьютером, неплохо осведомлён о том, что и где находится в Интернет. Он виртуозно добывает в Сети то, что нужно для написания реферата, курсового проекта, дипломной работы и т.п. И после небольшой доработки, которая часто заключается лишь в указании своей фамилии и номера группы, распечатав на принтере, сдаёт «свою работу» преподавателю. Такой подход вряд ли способствует будущей карьере. Чтобы достичь успеха в конкурентной борьбе, следует научиться обрабатывать колоссальные объёмы информации, уметь просматривать образцы письменных работ, подмечая в них слабые и сильные стороны, пробовать «препарировать» чужой текст, чтобы выделить из него наиболее существенную часть, и уже на основе полученных компонентов текста создавать действительно нужную работу.

Сегодня в Интернет около двух миллиардов Web-страниц. В результате каталогизация имеющихся в сети ресурсов становится серьёзной проблемой. Пространство Web быстрее наполняется, чем систематизируется. С помощью Интернет пользователь получает доступ к громадному массиву фактической информации. Однако за каждый результат приходится расплачиваться. Сегодня даже школьник начальных классов может опубликовать в Интернет собственную страничку. Услуга бесплатна; создание страницы практически не требует специальных знаний. Потому наряду с правдивой информацией Интернет наводнен информационным мусором, слухами и откровенной фальсификацией. Поэтому необходимо предельно критично относиться к информации, получаемой со страниц Интернет, пользоваться сведениями, взятыми только из надежных источников.

Поисковые системы – самые популярные среди ресурсов Интернет, ежедневно их посещают миллионы пользователей. Есть два типа поисковых систем. Первые относят к разряду каталогов – сведения о страничках Internet рассортированы в них по тематике и, как правило, добавляются в поисковик самими создателями страничек. Однако далеко не все создатели страничек могут (и хотят) вносить информацию о собственном детище в каталог. И тогда на помощь приходят поисковые системы второго типа, называемые «пауками». Эти периодически обследуют пространство Сети и заносят информацию о найденных страницах в специальный индекс – базу данных, по которой и ведется поиск. Существуют специализированные поисковики – например по научной информации, по ресурсам Сети для детей и т.д.. Собственные поисковые машины существуют и для серверов FTP. Кроме этого, Интернет обеспечивает удалённый доступ к базам данных. Базы данных и прочие источники информации доступны через многие Web-интерфейсы. Некоторые из них бесплатны, а для других необходима регистрация и оплата услуг. База данных – это набор записей, состоящих из нескольких полей, который является идеальной структурой для каталога неструктурированных источников информации. Для пользования определенной справочной базой данных необходимо знать ее организацию. Сама сеть Интернет в целом не может считаться базой данных по ряду признаков, например, ее записи (Web-страницы) отличаются по структуре, а Web-узлы – по качеству и возрасту информации.

В самых общих чертах поиск можно определить как упорядоченную последовательность действий, которые необходимо выполнить, чтобы найти интересующую информацию. Системный подход к этой задаче делает поиск и квалифицированным, и успешным.

Алгоритм поиска:

Определение вопроса

Вопрос – это то, на что вы хотите получить ответ; запрос – это то, что вы вводите в компьютер. Это совершенно разные вещи.

Вопрос формулируется на естественном языке, и только вы способны это сделать, так как только вам известен контекст. Библиотекари могут помочь, но уже на этапе превращения вопроса в запрос.

Выбор источника информации

Уяснив проблему, надо решить, где искать. Если ограничены технические или финансовые возможности, тем не менее, следует иметь представление, чего вы лишаетесь. Следует помнить, что ни одна из баз данных не является всеобъемлющей. Обдумайте характерные особенности каждого возможного источника информации и определите, где поиск будет наиболее успешным – в базе данных или менее упорядоченных источниках (например, Web).

При оценке источника информации приходится выбирать между его современностью (свежестью) и качеством. Качество обеспечивается при помощи рецензирования, редактирования и общественного мнения, хотя электронная почта и компьютерные средства доставляют информацию быстрее, а конференции и другие форумы доносят новости до читателей оперативнее, чем статьи.

Формулировка запроса

Определив вопрос и выбрав ресурсы для запроса поиска, надо сформулировать запрос, т.е. выразить вопрос на языке, понятном машине.

Для этого следует:

- разбить вопрос на понятия, которые будут искаться отдельно, а затем объединить их соответствующими логическими операторами (AND, OR, NOT). Подробнее о назначении логических операторов можно узнать в разделе «Помощь в поиске» любого поисковика.

- решить, насколько широко проводить поиск: сделать ли акцент на полноту отклика, т.е. найти по данной теме, часто получая при этом много ненужной информации, или на *точность*, т.е. быстро найти ограниченную конкретную информацию по теме, теряя некоторые относящиеся к ней важные сведения; заметьте, это будут совершенно разные запросы;

- учесть все варианты написания слов и синонимы, если вы ищете при помощи произвольного текста;

- выбрать соответствующую тематическую рубрику, если вы пользуетесь управляемым словарем; это эффективнее ключевых слов;

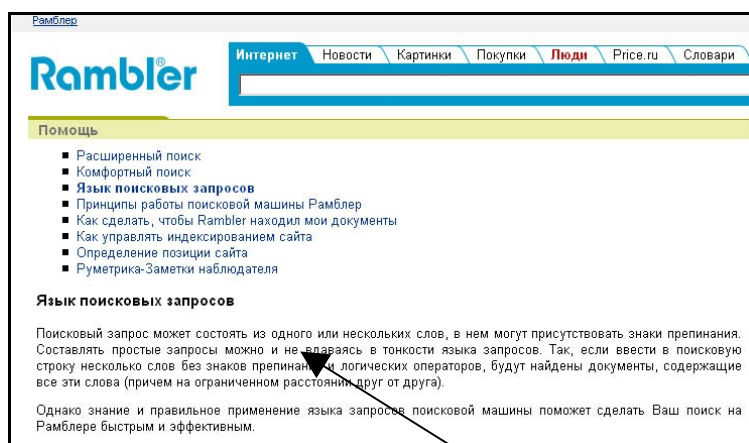
- решить, как вы бы ограничили поиск определенными полями базы данных или выполнили уточняющие и другие специальные функции.

При поиске произвольного текста вы ищете слова, которые были употреблены, например, в заголовках статей и рефератах, перебирая варианты написания и синонимы. Очевидно, уверенности, что найдено все возможное, нет, так как автор мог использовать редкий и специфический термин. В некоторых системах поиска алгоритм позволяет вводить набор слов и учитывает некоторые формы слова, например множественное число.

Оценка результатов поиска – это необходимый критический этап поиска, которым часто пренебрегают.

Критерии для оценки достаточности поиска:

- просмотрите результаты своего поиска, чтобы понять, почему выбраны эти



элементы, и какую часть составляют нужные, т.е. был ли поиск «прицельным»;

• не забывайте, что ошибки могли быть допущены на любой стадии поиска.

Например:

• слишком широкая или узкая формулировка запроса в начале;

• выбор не лучшей базы данных для поиска;

• неправильная формулировка запроса (выбор не той рубрики, ошибочное представление о выбранной рубрике, выбор многозначного ключевого слова или сокращения);

• если известны некоторые из найденных ссылок, но результаты поиска отрицательные, следует предположить наличие ошибки в стратегии. Если вы нашли слишком мало или слишком много, то считайте это первым сигналом о неэффективности поиска.

Не делайте скоропалительных выводов, если ничего (или почти ничего) не найдено. Часто это не является свидетельством отсутствия информации. Обычно доказать отрицательный результат (т.е. отсутствие литературы по этой теме) непросто. Прежде чем прийти к такому заключению, надо попробовать несколько других подходов.

Чрезмерно обширный результат поиска – явление довольно распространенное, часто из-за ошибочного определения проблемы или формулировки запроса. В таких случаях можно:

- ограничить поиск дополнительными критериями, предусмотренными системой
- прочитать названия и аннотации и выбрать самостоятельно наиболее подходящие ссылки, хотя это поверхностный подход, не учитывающий полного содержания статьи.

Если вы удовлетворены результатом поиска, запишите стратегию – она может быть неоднократно использована в дальнейшем для поиска более современной литературы.

И.А.Долгова, заведующая медиатекой.

