

Занятие №3. Научные базы данных: правила составления поискового запроса, поиск по ключевым словам

Цель лекции: получить практические навыки поиска и анализа информации в Интернете

Ключевые слова: сеть, интернет, информация, достоверность

Основные вопросы:

1. Практические методы поиска и анализа информации в Интернете.
2. Полнота, достоверность и скорость поиска
3. Планирование процесса поиска
4. Основные методы поиска в сети и их использование

Практические методы поиска и анализа информации в Интернете.

Прежде всего, необходимо обозначить проблему и задать интересующий вас вопрос. В зависимости от вопроса, можно в той или иной степени получить на него ответ через Интернет. А если учесть, что за каждым ресурсом стоят конкретные специалисты, то еще и существует возможность обратиться к ним с вопросом (например, используя электронную почту).

Это все, что касается «вопросов широкого профиля». А если нужна специализированная информация? Ну что ж, в этом случае Интернет может оказаться вообще единственным источником сведений. Дело в том, что специалисты узкого профиля разбросаны по всему миру достаточно «разреженно». Кроме того, у них не всегда есть возможность, регулярно публиковать свои материалы в научной или специальной литературе. Поэтому иногда бывает достаточно трудно найти интересующего специалиста, используя традиционные печатные источники информации.

Таким образом, в Интернете можно найти, либо уже готовый ответ на любой интересующий вопрос, либо, по крайней мере, специалистов, которые смогли бы ответить на него (а чаще всего и то и другое).

Итак, что же такое поиск информации в сети и чем он отличается от поиска информации в других источниках?

Поиск информации в сети – это последовательность действий, от определения предмета поиска, до получения ответа на имеющиеся вопросы с использованием всех поисковых сервисов, которые предоставляет Интернет сегодня.

Основные преимущества использования сети Интернет при поиске информации:

- Использование максимально возможного «пространства поиска» информации. Ни один из существующих на сегодня сетевых ресурсов не обладает тем объемом информации, который представлен в Интернете;
- Ни один другой источник не обладает такой оперативностью и доступностью. Интернет предоставляет доступ круглосуточно вне зависимости от места нахождения;

- Информацию, полученную через Интернет можно легко переслать для обсуждения или, например, распечатать в нужном числе экземпляров.

В целом поиск информации в сети – это достаточно специфическая и кропотливая работа, требующая определенных знаний и навыков. Для проведения поисковых работ пригодятся: знания основных информационных ресурсов и умение хорошо в них ориентироваться, практические навыки работы – это приходит со временем, хорошая зрительная память и умение быстро читать, а так же некоторые навыки аналитической работы.

При проведении поиска необходимой информации, необходимо просмотреть и обработать достаточно большой объем документов. Конечно, бывает и так, что ответ на вопрос находится в первом же документе и заключен в рамки одного абзаца, состоящего из трех строк. Но чаще всего – ответ формируется на основе «тридцати трех» разноплановых документов, хотя состоит все из тех же трех строк.

В основном сказанное касается поиска различных сведений достаточно общего характера или при необходимости составления аналитического отчета по интересующему вопросу. Поиск более конкретных сведений – потребует значительно меньше времени и усилий.

Основной проблемой при поиске можно назвать неумение пользователя эффективно искать информацию в сети.

Как правило, у начинающих поисковиков или у пользователей, которые только начинают «жить» в сети, существует несколько ошибочных мнений:

- Поисковые машины ищут информацию по всему Интернету;
- Не составляет большой проблемы еще раз найти заинтересовавший материал;
- Если начальный поиск закончился неудачей, то данной информации в сети нет.

Остановимся несколько подробнее на каждом из этих моментов.

Поисковые машины ищут информацию по всей сети Интернет.

На самом деле это не совсем верно. Если бы при реализации алгоритма работы поисковых машин был использован такой подход, то для обработки только одного запроса и выдачи результатов потребовалось бы несколько дней.

Поэтому, практически реализована иная схема работы поисковой машины. Каждая поисковая машина имеет и постоянно пополняет свою (локальную) базу данных. База данных поисковой машины содержит основные параметры (индексы) каждого известного данной машине (проиндексированного) документа. Каждая поисковая машина использует свои методы индексации. Кроме того, различные поисковые машины имеют разные объемы базы данных.

В результате, механизм обработки запроса пользователя поисковой машиной выглядит следующим образом:

- в соответствии с заданным в запросе ключевым словом или словосочетанием, машина проводит поиск в своей локальной базе данных,

сверяя ключевое слово с наборами ключевых слов, соответствующих каждому документу из её базы данных;

- затем, используя соответствующие алгоритмы, поисковая машина сортирует результаты поиска и выдает их пользователю;

- в результате сортировки результатов, в начало списка помещаются наиболее соответствующие (с точки зрения поисковой машины) ключевым словам документы.

В связи с огромным количеством информации, размещенной в сети, ни одна из поисковых машин не в состоянии просмотреть все документы. Каждая поисковая машина индексирует только часть их. Все остальные документы, а к сожалению это большая часть ресурсов, найти с ее помощью не удастся.

Не составляет большой проблемы еще раз найти заинтересовавший Вас материал.

Это второе очень большое заблуждение начинающих пользователей, приносящее немало вреда и значительно осложняющее жизнь. Достаточно часто бывает так, что интересные Вам материалы встречаются совершенно случайно, в процессе поиска по другой теме или при просмотре указанных в материалах сервера ссылок. А Интернет, "увлекая и заманивая" иногда лишает возможности вернуться. Отвлекаясь на более интересную информацию, забывая нужные адреса, можно потратить немало времени на их повторный поиск. И надо отметить, что не всегда вновь можно найти именно ту, потерянную информацию.

Для того чтобы избежать подобных ошибок, стоит сразу записывать адреса заинтересовавших ресурсов в раздел "Избранное" браузера, либо в текстовый файл. Если всё же не удалось сохранить адрес, то необходимо вспомнить какое-либо "кодовое" слово или словосочетание, которое точно было на "потерянном" сайте.

Главное – чтобы эта информация была оригинальной, так как использование стандартных сочетаний вида "общая психология" или "общая педагогика" не сильно продвинет на пути повторного поиска. Если же, например, использовать запомнившееся название заголовка статьи "Развитие личностной рефлексии", то есть шанс найти сайт с помощью поисковой машины.

Задав соответствующий запрос и нужно помнить – фразу надо вводить в кавычках.

В том случае, если ничего очень оригинального и своеобразного не запомнилось, стоит попробовать следующий вариант: в окне браузера найдите кнопку "журнал" и нажмите её. В левой части окна возникнет колонка, в которой будут надписи "сегодня, вчера, ...5 дней назад". Выберите мышкой надпись "сегодня" и нажмите на неё. В результате появится список серверов, которые Вы посетили за сегодняшний день. Если их число не слишком велико, то можно заново просмотреть все страницы.

В. Если начальный поиск закончился неудачей, то данной информации в сети нет.

Еще одно заблуждение, обусловленное в основном малым опытом поиска. В Интернете найти можно практически любую информацию, главное знать, что, где и как необходимо искать. Кроме того, очень важно научиться грамотно формулировать запрос поисковой машине. И если на начальном этапе не удалось найти то, что необходимо, это вовсе не означает, что данная информация отсутствует. Это значит, что Вы либо не достаточно четко формулируете запрос поисковой машине, либо ищите необходимую информацию не там, где её стоит искать.

Полнота, достоверность и скорость поиска

Для того чтобы найденная информация действительно отражала реальное положение дел, поиск должен удовлетворять следующим критериям:

- полнота охвата ресурсов;
- достоверность информации;
- высокая скорость проведения поиска.

Почему важна полнота охвата и достоверность информации, видимо, объяснять не нужно. По поводу скорости проведения поиска можно сказать лишь то, что чем меньше времени затрачено на проведение поисковых мероприятий, тем меньшими будут издержки (на оплату услуг сети, на оплату времени работы специалистов и т.д.).

Гораздо сложнее вопрос проверки: насколько удастся достигнуть полноты охвата и достоверности информации? Ответить на этот вопрос заранее и со стопроцентной уверенностью вряд ли возможно. Можно лишь дать рекомендации, основанные на опыте тех специалистов, которые занимаются поиском профессионально – что надо делать, чтобы вероятность достижения заданных критериев была максимальной.

Контроль полноты охвата ресурсов – достаточно серьезная проблема, в том случае если необходимо не просто найти любую информацию об интересующем предмете, а Вы хотите иметь полное представление о предмете поиска и состоянии дел по данному вопросу. В этом случае лучше не ограничиваться использованием только одной поисковой машины или просмотром одного, пусть даже самого любимого каталога. Для того чтобы провести полномасштабный сбор информации необходимо работать со всеми известными каталогами, поисковыми машинами, базами данных, региональными телеконференциями, электронными досками объявлений и листами рассылок. Только в этом случае Вы будете уверены, что нашли если не всю, то хотя бы большую часть существующей информации. При этом не стоит забывать и об электронных СМИ.

Контроль достоверности информации. Еще одна важная и достаточно сложная часть процесса поиска. В силу своей специфики, Интернет содержит достаточное количество устаревшей либо не достоверной информации. Во многом это объясняется возможностью анонимного размещения материалов, особенно на бесплатных сервисах. В основном, контроль достоверности информации – это аналитическая работа. При этом важно произвести сверку найденного фактического материала, выяснить статус документов, получить информацию о компетентности автора материала и т.д.

Но есть и более очевидные вещи. Например, вряд ли стоит доверять организации, на сайте которой нет других координат связи, кроме адреса электронной почты. Либо сайту, который не встречается в рейтингах популярности в первой сотне, но авторы которого утверждают, что он имеет грандиозный успех.

Результаты поиска.

Как уже отмечалось выше, при поиске в Интернете важны:

- полнота охвата – ничего не потеряно из имеющейся информации
- точность – не найдено лишней или недостоверной информации.

Получив нужную информацию по сети, постарайтесь для начала ее проверить. Проверяется любая адресная информация: какими бы возможностями ни обладала современная техника, но заносит информацию в сеть люди, а им свойственно ошибаться, проявлять неаккуратность, не оперативность и т.д.

Аналитическую информацию проверяют, сравнивая данные, полученные из нескольких источников. Таким же способом проверяют разнообразные статистические данные.

Если Вы предполагаете использовать полученную информацию для публичного обозрения, не сочтите за труд точно записать источник (адрес в сети), имя автора (или название организации) и дату публикации используемых сведений. Обратите внимание на возможное наличие особых указаний относительно соблюдения авторских прав.

Помните о том, что если некоторую информацию смогли получить Вы, то таким же образом ее может получить и любой другой.

Наконец, самое главное. Задайтесь вопросом: принесла ли найденная информация ту пользу, на которую Вы рассчитывали? Удалось ли Вам получить информацию в необходимом объеме за отведенное время?

Описание процедуры поиска информации в сети выглядит довольно объемным. На практике, после приобретения некоторого навыка работы в сети процесс поиска вряд ли будет представлять какие-либо трудности.

Скорость проведения поиска в Сети, если не принимать во внимание технические характеристики подключения пользователя, зависит в основном от двух факторов:

- грамотного планирования поисковой процедуры;
- опыта работы с ресурсом выбранного типа.

Особое значение скорость проведения поиска имеет в том случае, когда Вы имеете дело с быстро обновляющейся информацией.

3. Планирование процесса поиска

Эффективность любого вида деятельности определяется четким представлением того, что, как и в какой последовательности мы собираемся делать, т.е. четким планированием работ. Сказанное в полной мере относится и к процедуре поиска информации в сети Интернет.

Рассмотрим, из каких этапов состоит процесс поиска информации:

№	Этап	Содержание работ этапа
1.	Определение предмета поиска	На этом этапе определяем, что конкретно нас интересует.
2.	Составление списка ключевых слов	На этом этапе выявляем, как может называться то, что нас интересует.
3.	Выбор информационного пространства	На этом этапе определяем, где может находиться то, что нас интересует.
4.	Определение инструмента для поиска	На этом этапе принимаем решение о том, как проще и быстрее найти то, что нас интересует.
5.	Предварительный поиск	Пробуем найти.
6.	Анализ полученной информации	Смотрим на полученные результаты. Если это необходимо (в том случае, когда полученные результаты нас не устраивают), проводим корректировку всех предыдущих действий.
7.	Дополнительный поиск	Ищем дальше, пока не получаем ответ на свой вопрос.

Неплохо, также, определить время, в течение которого информация должна быть найдена, оценить альтернативные способы получения и степень важности этой информации для Вас.

Общие советы при поиске.

- Потратьте несколько лишних минут, чтобы максимально "сузить" описание предмета поиска – это поможет сэкономить Вам много времени и денег;
- Искать что-то конкретное лучше всего с помощью поисковых машин, так как если вы знаете "хорошие" ключевые слова, четко определяющее то, что вы хотите найти, то и поиск не представляет никакого труда;
- Используйте для поиска нескольких поисковых машин;
- Если на просматриваемой Вами странице существует несколько заинтересовавших Вас ссылок – открывайте несколько окон, пока Вы читаете информацию на одной странице, остальные успешно (или не очень) успевают загрузиться;
- Старайтесь использовать подходящую поисковую систему: файлы ищите с помощью FTP-поисковика, MP3 (файлы музыкальных произведений, записанные и обработанные в цифровом формате) – с помощью MP3-

поисковика, благо сейчас практически любая поисковая машина предоставляет такую возможность. В случае поиска файлов, в запросе указывайте наиболее вероятное имя файла, например, «name.jpg»;

- Если вы хотите найти популярные, часто посещаемые ресурсы – ищите с помощью рейтинга;
- Телеконференции и страницы ссылок на тематических сайтах помогут Вам быстрее найти специализированные или редкие вещи;
- Если нужна очень редкая информация – попробуйте найти ресурсы, посвященные более общей теме. Возможно, там будут размещены ссылки на необходимые Вам ресурсы или будут опубликованы требуемые материалы;
- Старайтесь найти ответ, а не задать вопрос;
- Создавайте свою коллекцию интересующих Вас ссылок;
- Если Вы нашли что-то интересное, сразу запишите адрес ресурса в "Избранное" либо в текстовый файл;
- Записывая адреса ресурсов в "Избранное", старайтесь дать им более четкое название, максимально соответствующее содержащейся в них информации, постарайтесь не использовать слишком длинные названия;
- Разработайте свой собственный классификатор и для каждой темы создавайте отдельную папку в "Избранном". Поверьте, поиск в своих собственных не разобранных архивах занимает не меньше времени, чем поиск в Интернете.

Основные методы поиска в сети и их использование

Существует два основных метода поиска информации в Интернет – с использованием поисковой машины или с использованием каталога. При этом сам механизм поиска в обоих случаях практически одинаков. Различия возникают на этапе 2 (составление списка ключевых слов). Для поисковой машины – это составление списка ключевых слов, а для каталога на этом этапе производится определение тематики разделов, в которых может находиться необходимая информация.

Например, Вам необходимо найти информацию о психических процессах

Если Вы решили использовать поисковую машину, то необходимо подобрать набор ключевых слов, необходимых для поиска:

- *психические процессы;*
- *познавательные процессы;*
- *психические процессы в структуре личности;*
- *внимание, память, мышление;*
- *и т.д.*

А, если Вы используете для поиска каталог, то требуется понять, в каких разделах может находиться интересующая Вас информация:

- *Психология ---> Психология познавательных процессов;*
- *Социальная психология ---> Общение ---> Коммуникативный процесс;*
- *Педагогика ---> обучение ---> дидактика---> методы обучения и т.д.*

Теперь давайте подробнее рассмотрим этапы поиска информации, описанные в предыдущем разделе.

Определитесь с предметом поиска. Основное правило поиска – определитесь, что именно Вас интересует, и что Вы хотите найти? Ведь если Вы хорошо знаете, предмет поиска, то искать намного легче.

Постарайтесь, как можно более подробно конкретизировать, что же Вы будете искать.

Например, Вы хотите найти информацию по истории отечественной психологии. Какие именно психологические направления Вас интересуют? Или психологические школы? Или конкретные фамилии? Можно конечно собрать информацию обо всём, но если Вы будете искать что-то конкретное, то потратите значительно меньше времени и получите более качественный результат.

Итак, как выглядят правильное и неправильное описание предмета поиска?

Неправильное: Найти историю отечественной психологии.

Здесь задача поставлена слишком широко, при поиске непременно возникнут вопросы по дальнейшей конкретизации необходимой Вам информации. Что нужно найти: информацию о направлениях психологии, либо об авторах научных статей, либо выдающихся учёных или практиках или же книги по истории психологии. Какая информация требуется: только фамилии, информация о работах, школы и т.д.

Правильное: Найти отечественных учёных-психологов живших в период с 1960 по 2000 годы. Работающих над проблемами социальной психологии.. Регион: Россия. Требуемая информация: Научные работы. Статьи. Открытия.

Если же Вы ничего не знаете о предмете поиска кроме общей темы, то начать придется с поиска информации о самом предмете поиска. Задача переходит в другую плоскость – "найти хоть что-нибудь". Можете попробовать найти интересующую Вас тему в каталоге ссылок, кроме того, источниками информации могут быть специальные справочники или энциклопедии общего профиля, где можно почерпнуть начальные сведения. Если Вас постигла неудача – воспользуйтесь поисковой машиной. Правда в этом случае Вас ждет "море информации" и Вы потратите значительное время, чтобы его "переплыть".

Подбор ключевых слов. Найдете ли Вы нужную Вам информацию, во многом зависит от правильного подбора ключевых слов. Поэтому очень важно научиться подбирать ключевые слова и их комбинации.

На основе предварительно полученной информации о предмете поиска составьте максимально широкий набор ключевых слов в виде отдельных терминов, словосочетаний, названий и фамилий, тесно связанных с проблемой, профессиональной лексики.

Естественный язык запросов привлекает своей простотой, однако, как показывает практика, он менее эффективен при серьезном поиске, когда ставится задача найти все или большинство документов на заданную тему.

Рассмотрим основные моменты, которые помогут вам правильно сформулировать запрос к поисковой машине:

- Не задавайте только одно слово. Используйте целые фразы, или, по крайней мере, несколько слов;
- Указывайте слова, которые не должны встречаться в искомым документах. Обычно для этого используют либо знак "-", либо ключевое слово NOT;
- Начинайте имена собственные: названия книг, городов, фамилии людей, и др. с заглавных букв, а все остальные слова пишите только маленькими буквами;
- Если вы хотите найти фразу целиком – заключите ее в кавычки;
- Если вы вводите запрос к поисковой машине, состоящий из нескольких слов, то в результате получаете список документов, в которых встречается хотя бы одно слово.

Стоит отметить, что поисковая машина обычно производит сортировку документов по принципу релевантности.

Итак, что же такое релевантность. При индексации документов поисковые машины высчитывают так называемый "вес" слова на странице – соотношение количества повторов на странице заданного Вами слова к общему количеству слов на странице документа. Если Вы задаете запрос, состоящий из нескольких слов, то более релевантными будут документы, в которых совокупный вес слов будет максимальный. Однако, при подсчете веса не учитывается, рядом или отдельно стоят данные слова, и поэтому нет гарантий, что в первых документах содержится максимальное количество повторений словосочетания. Вполне возможно, что такого словосочетания там вообще не будет.

Поэтому, если Вы хотите найти заданное словосочетание – задавайте запрос в окне поисковой машины в кавычках. В этом случае будет высчитываться вес словосочетания целиком. Соответственно, гарантируется наличие именно данного словосочетания в найденных документах.

Критерием правильности выбора ключевых слов для Вас будет служить количество документов, найденных поисковой машиной при предварительном поиске. Если оно исчисляется сотнями – то Вам необходимо уточнить запрос, введя дополнительные ключевые слова или воспользовавшись специальными средствами расширенного поиска. Если не найдено ни одного документа – такое редко но случается, составьте список синонимов данного термина или проверьте орфографию.

Выбор информационного пространства. Четкое представление, где может находиться запрашиваемая Вами информация, в одних случаях избавит Вас от необходимости просматривать большое количество лишней информации, в других позволит найти те сведения, которые при заданных по умолчанию в поисковой машине условиях поиска, вообще не отыскиваются.

Например: конкретное название сайта по психологии и педагогики, статьи или программы.

Выбор поисковых инструментов. Этот этап определяет всю дальнейшую технологию поиска – переходить к нему стоит только после анализа предварительно полученной на предыдущих этапах информации.

Поскольку на сегодняшний день наиболее популярным информационным пространством является *World Wide Web*, то остановимся более подробно на вопросе выбора инструментов поиска именно для WWW. Как правило, у пользователя регулярно занимающегося поиском существует свой отработанный механизм, однако некоторые рекомендации применительно к работе с основными видами инструментов поиска в сети дать можно:

Поисковые машины – с помощью поисковых машин всегда легче найти что-то конкретное, и труднее – что-то общее и неопределенное. Естественно, размер базы данных поисковой машины – решает не все. Для поисковой машины самое главное – релевантность (соответствие запросу) выдаваемых ссылок, а на этот показатель, кроме размера, влияет структура базы данных и синтаксис языка запросов.

Каталоги. Именно с помощью каталогов удобно искать информацию на общие темы (природа, искусство, наука, компьютеры, медицина и т.д.). Чем удобен поиск в каталоге? Любой каталог имеет рубрикатор. Таким образом, информация в определенной степени отсортирована. Для этого и существуют подкаталоги. Каждый адрес снабжен аннотацией представленных материалов. При решении достаточно стандартной поисковой задачи именно каталог, а не поисковая машина оказываются более приемлемыми для начала поиска.

Во многом тип требуемого Вам информационного ресурса определяется характером искомой информации:

- Адресная информация – адресные справочники;
- Новости – новостные порталы и информационные агентства, списки рассылок;
- Конкретная информация – поисковые машины;
- Общие сведения – каталоги;
- Статистика – сервера статистических служб, рейтинги;
- Хотите что-то купить – Интернет-магазины и доски объявлений;
- Хотите проконсультироваться – сайты, специализирующиеся на данной тематике и списки рассылок.

Для выбора наиболее подходящих Вам в данный момент информационных ресурсов используйте тестовые запросы из 1-2 ключевых слов. Проанализируйте количественный отклик и выберите наиболее представительные источники информации.

Предварительный поиск и анализ полученных результатов. Конечно, предварительный поиск может оказаться сразу и окончательным, если все предыдущие этапы были выполнены полно и аккуратно, а результат Вас вполне устроил.

Такая ситуация возникает чаще всего в случае, когда Вам необходима какая-либо конкретная информация по хорошо известным Вам вопросам. В случае, когда параметры или предмет поиска Вам самому не вполне точно

известны, тогда по анализу результатов предварительного поиска можно существенно уточнить, что, где и как нужно искать.

Вопросы

1. Какие практические методы поиска и анализа информации в Интернете вы знаете?
2. Что такое полнота, достоверность и скорость поиска?
3. Как осуществить планирование процесса поиска?
4. Каковы основные методы поиска в сети и их использование?

Основная литература:

Интернет-ресурсы:

1. <https://knastu.ru> › files ›
2. <https://www.scopus.com/home.uri>
3. thomsonreuters.com
4. <https://elsevierscience.ru>