



Введение

Основная задача Интернет – предоставление необходимой информации. **Интернет** – это информационное пространство, в котором можно отыскать ответ практически на любой интересующий пользователя вопрос. Это огромная глобальная сеть, в которую как информационные ручейки, стекаются потоки более мелких сетей. Любой пользователь, располагающий ПК и соответствующими программами, сможет подключиться к сети, используя её возможности для самых разных целей – проведения досуга, обучения, чтения научных работ, отправки электронной почты и т.д. По различным данным, в 2004 г. количество пользователей глобальной системы Интернет составило от 600 до 900 миллионов человек. Это число продолжает стремительно расти и уже в 2014 г. оно достигло 3 миллиарда человек. Сегодня глобальная сеть превращается в важный социальный и политический фактор современного информационного общества. С развитием Интернет-технологий появился новый гигантский источник информационных ресурсов, доступ к которым является не только относительно дешевым, но и очень быстрым.

Поисковые системы общего назначения

Поисковые системы общего назначения являются базами данных, содержащим тематически сгруппированную информацию об информационных ресурсах Всемирной паутины. Такие поисковые системы позволяют находить Web-страницы **по ключевым словам** в базе данных или путем поиска в **иерархической системе каталогов**.

Интерфейс таких поисковых систем общего назначения содержит список разделов каталога и поле поиска. В поле поиска пользователь может ввести ключевые слова для поиска документа, а в каталоге выбрать определенный раздел, что сужает поле поиска и таким образом ускоряет его.

Заполнение баз данных осуществляется с помощью специальных программ-роботов, которые периодически «обходят» Web-серверы Интернета. Программы-роботы читают все встречающиеся документы, выделяют в них ключевые слова и

заносят в базу данных, содержащую URL-адреса документов.

Так как информация в Интернете постоянно меняется (создаются новые Web-сайты и страницы, удаляются старые, меняются их URL-адреса и так далее), поисковые роботы не всегда успевают отследить все эти изменения. Информация, хранящаяся в базе данных поисковой системы, может отличаться от реального состояния Интернета, и тогда пользователь в результате поиска может получить адрес уже не существующего или перемещённого документа.

В целях обеспечения большего соответствия между содержанием базы данных поисковой системы и реальным состоянием Интернета большинство поисковых систем разрешают автору нового или перемещенного Web-сайта самому внести информацию в базу данных, заполнив регистрационную анкету. В процессе заполнения анкеты разработчик сайта вносит URL-адрес сайта, его название, краткое описание содержания сайта, а также ключевые слова, по которым легче всего будет найти сайт. Сайты в базе данных ранжируются по количеству их посещений в день, неделю или месяц. Посещаемость сайтов определяется с помощью специальных счетчиков, которые могут быть установлены на сайте. Счетчики фиксирует каждое посещение сайта и передают информацию о количестве посещений на сервер поисковой системы.

Поиск по ключевым словам. Поиск документа в базе данных поисковой системы осуществляется с помощью введения запросов в поле поиска. Простой запрос содержит одно или несколько ключевых слов, которые являются главными для этого документа. Можно также использовать сложные запросы, использующие логические операции, шаблоны и так далее. Через некоторое время после отправки запроса поисковая система вернет аннотированный список URL-адресов документов, в которых были найдены указанные ключевые слова. Для просмотра этого документа в браузере достаточно активизировать указывающую на документ ссылку. Если ключевые слова были выбраны неудачно, то список URL-адресов документов может быть слишком большим (может содержать десятки и даже сотни тысяч ссылок). Для того чтобы уменьшить список, можно в поле поиска ввести дополнительные ключевые слова или воспользоваться каталогом поисковой системы.

Типология методов поиска

Более или менее серьезный подход к любой задаче начинается с анализа возможных методов ее решения. Поиск информации в Интернете может быть

произведен по нескольким методам, значительно различающимся как по эффективности и качеству поиска, так и по типу извлекаемой информации. В ряде случаев приходится использовать весьма трудоемкие - результат того стоит. Можно выделить следующие основные методы поиска информации в Интернете, которые, в зависимости от целей и задач ищущего, используются по отдельности или в комбинации друг с другом:

1. Непосредственный поиск с использованием гипертекстовых ссылок.

Поскольку все сайты в пространстве WWW фактически оказываются связанными между собой, поиск информации может быть произведен путем последовательного просмотра связанных страниц с помощью браузера. Хотя этот полностью ручной метод поиска выглядит полным анахронизмом в Сети, содержащей более 60 млн. узлов, "ручной" просмотр Web-страниц часто оказывается единственно возможным на заключительных этапах информационного поиска, когда механическое "копание" уступает место более глубокому анализу. Использование каталогов, классифицированных и тематических списков и всевозможных небольших справочников также относится к этому виду поиска.

2. Использование поисковых машин. Сегодня этот метод является одним из основных и фактически единственным при проведении предварительного поиска. Результатом последнего может являться список ресурсов сети, подлежащих детальному рассмотрению.

Как правило, применение поисковых машин основано на использовании ключевых слов, которые передаются поисковым серверам в качестве аргументов поиска: что искать. Если делать все правильно, то формирование списка ключевых слов требует предварительной работы по составлению тезауруса.

3. Поиск с применением специальных средств. Этот полностью автоматизированный метод может оказаться весьма эффективным для проведения первичного поиска. Одна из технологий этого метода основана на применении специализированных программ-спайдеров, которые в автоматическом режиме просматривают Web-страницы, отыскивая на них искомую информацию. Фактически это автоматизированный вариант просмотра с помощью гипертекстовых ссылок, описанный выше (поисковые машины для построения своих индексных таблиц используют похожие методы). Нет нужды говорить, что результаты автоматического поиска обязательно требуют последующей обработки. Применение данного метода целесообразно, если использование поисковых машин не может дать необходимых результатов (например, в силу нестандартности

запроса, который не может быть адекватно задан существующими средствами поисковых машин). В ряде случаев этот метод может быть очень эффективен. Выбор между использованием спайдера или поисковых серверов является собой вариант классического выбора между применением универсальных или специализированных средств.

Специализированные поисковые систем

С появлением системы пользовательского поиска от Google(Google Custom Search Engine) Специализированные поисковые системы стали появляться как нибудь после дождя. В чем же их преимущества перед обычными поисковыми системам и? По сути оно только одно, но весьма весомое – если искать информацию по какой-то узкой теме в специализированной поисковой системе, выдача, которая получается в результате запроса, будет гораздо более чистой – не придется отсеивать десятки, а то и сотни ссылок рекламного характера и прочего мусора. Небольшой список специализированных поисковых систем:

beeMP3 – специализированный поисковые системы по музыке. Можно искать по альбому, исполнителю, песне или всему сразу. У beeMP3 достаточно интересна организована выдача — сразу получают ссылки на конкретный файл (а не на страницу с ним). Кроме того, наведя курсор мышки на ссылку, можно узнать жанр песни, альбом, год выпуска, битрейт и др.

Tagoo – русскоязычная специализированная поисковая система по музыке keeperweb.com – специализированная медиа поисковая системы по mp3 музыке, кино, клипам, мелодиям для мобильных телефонов и много другого.

FindSounds – поиск звуковых эффектов и музыкальных сэмплов. В отличие от других мультимедийных поисковых систем, которые ищут песни, радиопередачи и тому подобное, FindSounds ищет только простые звуки и короткие отрывки.

eVdb – поиск электронных книг.

poiskknig.ru – поиск электронных книг, свободно распространяемых в Интернете.

WikiPoisk – поиск по энциклопедиям.

Qwika – специализированная поисковая система, предназначенная для работы с онлайн-энциклопедией Wikipedia. Является единственной в мире поисковой системой, индексирующей информацию, полученную посредством систем автоматического перевода.

Scirus – поисковая система для ученых, ищет web-страницы с научным содержанием (ищет и русскоязычные сайты): сайты университетов, библиотек и т.д.

ILIGENT – поиск информации на бизнес-сайтах. Для поиска доступны материалы по маркетингу, менеджменту, бухгалтерскому учету, финансам, управлению кадрами предприятия, законодательству. Результаты поиска разделены на 9 типов информации: материалы (статьи), новости, словари, сообщения на форумах, право, рефераты, книги в продаже, платные материалы и мероприятия.

PureVideo – поиск видеофайлов.

Киновед – поиск всего, что связанного с фильмами: отзывов, рецензий и описаний фильмов, информации об актерах кино, биографий, кадров из фильмов, фотографий актеров, постеров, обоев для рабочего стола.

DVD-поиск – специализированная поисковая система фильмов. В базе поиска содержатся базы фильмов самых известных Интернет-магазинов. В результатах поиска выводятся: название фильма, магазин, формат и цена. У некоторых позиций также присутствует небольшая аннотация.

PicSearch – поиск изображений.

Заключение

В сети Интернет хранится очень большой объем информации по различной тематике в виде статей в электронных газетах, отчетов, справочников, графических изображений, аудио и видеофайлов и многого другого. Путешествуя по Интернет, можно найти любую информацию, не выходя из дома или офиса, иначе говоря, если какие-либо данные вводились когда-либо в компьютер, то вероятнее всего их можно найти где-то на необозримых просторах Интернет. Для этого нужен только непосредственно сам компьютер, подключенный к Интернет с установленной специальной программой – браузером, предназначенной для просмотра содержимого Web-страниц.

С появлением World Wide Web (WWW) - гипертекстовой информационной системы, опутавшей весь мир, связавшей сотни и сотни компьютеров, тысячи и тысячи страниц гипертекстовых документов - без высокоэффективных средств поиска

просто не обойтись. Благодаря разнообразию поисковых систем, специально разработанных для рядового пользователя, каждый может без труда отсеять заведомо ненужный поток информации, лишь правильно сформулировав цель поиска.

Список Литературы

Интернет глазами пользователя. Режим доступа: <http://yandex.ru/>

Средства поиска информации в INTERNET // Афанасий-биржа (Тверь). - 28 марта 1997.

Википедия - свободная энциклопедия <http://ru.wikipedia.org/>