

Поисковая система

Содержание

1. Введение
2. Поисковые системы
 - 2.1 Виды поисковых систем
- Первая поисковая система
3. Задачи и стратегии поиска
4. Инструменты поиска в Интернете
5. Средства поиска информации
6. Заключение

ВВЕДЕНИЕ

Сегодня, почти каждый человек хотя бы раз пользовался поисковой системой. Неважно какой именно. Цифровые технологии и IT – технологии настолько плотно вошли в нашу жизнь, что некоторые уже не могут жить без них и интернета. Каждый день среднестатистический человек, у которого есть доступ в сеть Интернет, пользуется поисковой системой в среднем около 25-30 раз. Поиск в компьютерных сетях становится искусством и требует вполне определённых знаний, которыми современному человеку необходимо овладеть. Я выбрала данную тему реферата, так как считаю, что одной из составляющей грамотной работы в Интернет является умение работать в глобальной сети. И для того, чтобы не утонуть во всём изобилии информации необходимо научиться не только её искать, но и находить сейчас подробно рассмотрим тему поисковых систем.

Актуальность

Цель

задачи

ГЛАВА 1 ПОИСКОВАЯ СИСТЕМА

1.1 Виды поисковых систем

Поисковая система (поисковик) — это сайт, который осуществляет поиск ответа на запрос пользователя по всем известным данной системе сайтам. Существует множество поисковых систем, каждая из которых использует собственные алгоритмы построения списка сайтов, содержащих по мнению поисковика ответ на запрос пользователя. Помимо алгоритмов поисковая система использует в работе роботов, которые индексируют сайты, изображения, проверяют доступность сайтов. Поиск может осуществляться не только по текстовому запросу, вводимому в строку поиска, но, например, и по картинке или голосовому сообщению.[1]

Большинство поисковых систем учитывают региональность сайта, пользователя и его запроса, выдавая посетителю наиболее корректный по мнению поисковика ответ в виде списка сайтов.

Как не удивительно это не звучало, но даже поисковые системы делятся на несколько видов. Поисковые системы делятся на следующие виды:

Поисковые системы разрабатываемые изначально для поиска сайтов внутри конкретной страны, т.е. для внутреннего рынка. Большинство из них постепенно вышли за рамки своего государства, но при этом не перешли в разряд транснациональных.

Транснациональные поисковые системы. Поисковые системы, осуществляющие поиск ответа на запрос пользователя по сайтам всех стран, независимо от их доменной зоны и страны нахождения.

Различают поисковые системы: глобальные, предназначенные для поиска информации в интернете; локальные, осуществляющие поиск в рамках локальных сетей или определенных ресурсов. Глобальные поисковые системы могут быть: Универсальными. Такие ПС предоставляют своим

пользователям возможность поиска любого вида контента: текстового, графического, аудио, видео. Сам поиск осуществляется по всем ресурсам интернета.

Первой системой, придуманной для нахождения информации во Всемирной паутине был «Wandex», уже не существующий индекс, который создавал «World Wide Web Wanderer» — бот на языке Perl, разработанный Мэтью Грэм из Массачусетского технологического института в 1993 году. Создатель «Вандекса» также принимал участие в разработке Apache, создании сайта MIT и других проектах. Сейчас, как бы удивительно это не было, Мэтью Грэй является сотрудником компании Google.

Нужно отметить, что в 1993 году во всём Интернете насчитывалось всего лишь 623 веб-сайта, так что работы у «Вандекса» было поменьше, чем у современных поисковых систем. К тому же, благодаря таким скромным масштабам, для составления индексов и выдачи результатов не требовались ни мощные сервера с многоуровневой кластеризацией, ни сложная функциональность алгоритмов поиска и ухищрения по их оптимизации.

Wandex и другие первые сервисы имели весьма несовершенную систему поиска. При запросе слова они выдавали не наиболее релевантные ссылки, а все страницы, его содержащие, без анализа и ранжирования. Поэтому некоторое время реальными конкурентами поисковиков были каталоги страниц — сайты с вручную отсортированными коллекциями ссылок на веб-ресурсы.

1.2 Задачи и стратегии поиска

Прежде чем начать поиск следует четко сформулировать задачу. До тех пор, пока задача не будет сформулирована об успешности, а тем более оптимизации поиска, не может быть и речи. Объем информации в Интернете слишком велик, и как нигде здесь верна поговорка "Время - деньги". Поэтому, здесь нельзя искать по принципу "Пойди туда, не знаю куда. Возьми то, не знаю, что". Иначе результат будет плачевный. Дать общий рецепт эффективной стратегии поиска информации в Интернет, пожалуй, невозможно. Есть лишь некоторые принципы, позволяющие тратить меньше времени. Вот некоторые из них. Например, вам необходимо узнать, где обитает трёхлапый ленивец, то вряд ли вы пойдете в алфавитный каталог библиотеки. Может быть, вы найдете нужную литературу с помощью систематического каталога. С несколько большей вероятностью - с помощью предметного. Но, скорее всего, ни один из библиотечных каталогов вам не поможет. Но, зайдя в информационно-библиографический отдел крупной библиотеки, дежурный библиограф достанет библиографический указатель по животным или какую-то похожую книжицу, из которой вы и найдете ответ на свой вопрос. Подобную стратегию можно с успехом применять и в Интернет. В ИПС (информационно-поисковых сетях) общего назначения можно утонуть в тысячах ссылок, выданных вам на простой запрос. Целью использования универсальной ИПС общего назначения может быть поиск специализированной ИПС, посвященной тематике вашего поиска. Такая ИПС может быть распознана по наличию слов "информация", "документ" и т.п. Но часто специализированная ИПС может скрываться на сервере общественной, профессиональной или специализированной организации, издательства.

Этот пример иллюстрирует еще один элемент стратегии: читайте найденные документы в поисках наиболее точных терминов и связей между терминами. Возможно, вы мыслите в совершенно не тех терминах, которые

используют авторы искомых документов (ведь большую роль играют культурные различия!).

Третий элемент стратегии: используйте несколько ИПС. Если вы регулярно занимаетесь поиском информации по какой-то тематике, отметьте те ИПС, которые для вас наиболее эффективны.

Инструменты поиска в Интернете. По принципу организации и использования поискового инструмента можно разделить на каталоги (каталоги, каталоги) и поисковые системы.

Каталоги — это каталоги, содержащие списки интернет-адресов, сгруппированных по определенным атрибутам. Обычно они сгруппированы по тематическим областям (наука, искусство, новости и т.д.), причем каждая тематическая область разветвляется на несколько подуровней. Особенностью этих поисковых инструментов является то, что создание структуры, базы данных и ее постоянное обновление осуществляется «вручную» командой редакторов и программистов, а сам процесс поиска требует непосредственного участия пользователя, который переходит от ссылки к ссылке самостоятельно. Основным преимуществом поисковых каталогов ресурсов является то, что хранящаяся в них информация тщательно отбирается, сортируется и разделяется на отдельные категории.

Все каталоги структурированы по одному принципу. Страницы каталога содержат тематические классификаторы. На стартовой странице находится тематический каталог верхнего уровня. Подкатегории могут быть расширены под каждой категорией каталога. При нажатии на одну из тематических ссылок открывается страница с подчиненными тематическими категориями, которые, в свою очередь, также подразделяются на подкатегории. Все категории и подкатегории выделены соответствующим образом.

Каталоги организованы иерархически и расположены в виде тематического дерева, корнем которого является домашняя страница каталога и которое содержит определенное количество тематических ветвей.

Вы можете перемещаться вверх или вниз по тематическому дереву, нажимая на названия страниц. Страницы большинства каталогов содержат специальные элементы — индексы, содержащие ссылки на узлы, связанные с данной темой. При нажатии на одну из ссылок появится окно, показывающее организацию данной ветви дерева тем. Ускоряет отображение тематических категорий и подкатегорий.

Одним из самых известных каталогов является Yahoo, который содержит множество иерархически упорядоченных тематических категорий. Категории на всех уровнях разветвлены на подкатегории. Корень — домашняя страница каталога, содержащая 14 тематических ветвей. Вы можете перемещаться вверх или вниз по тематическому дереву, нажимая на названия страниц. Индексы всех веб-ресурсов, связанных с темой, включены в Yahoo! Все каталоги и подкатегории выделены жирным шрифтом, а гиперссылки на существующие сайты выделяются жирным шрифтом. На каждой странице Yahoo! встроены специальные иконки.

Поисковая система Yahoo может быть активирована на каждой странице каталога. Для поиска во всем каталоге или на одной из страниц используются специальные кнопки и поле поиска. Поиск Yahoo! основывается на типе адресов, выбранных через опции меню (под поисковым полем). Помимо системы адресации, используемой в тематическом каталоге, в дополнение к электронной почте можно обратиться в информационную службу Usenet. При выборе опций меню Yahoo! осуществляется поиск среди сайтов, содержащихся в каталоге тематических новостных сервисов.

В Интернете размещены узлы и другие каталоги, которые оснащены поисковыми системами, аналогичными поисковым системам Yahoo!

ГЛАВА 2 СРЕДСТВА ПОИСКА ИНФОРМАЦИИ

1. Средства поиска информации
2. Тематические каталоги ресурсов;
3. Поисковые системы;
4. Метапоисковые системы.
5. Тематические каталоги.

В каталогах информация о сайтах упорядочена в соответствии с рубрикаторм. В отличие от простых подборок ссылок имеется механизм поиска. Работа поисковых механизмов ограничивается поиском в кратких аннотациях сайтов.

Кроме основных разделов многие каталоги имеют дополнительные разделы, в которых сайты классифицированы по другому основанию:

1. региону, стране;
2. алфавиту;
3. популярности.

Информация о сайтах в каталоги может вноситься вручную как авторами сайта через механизм регистрации, так и модераторами - ведущими рубрик каталога. Преимущество каталога - создание списка "отборных" сайтов. Недостатки: неоднозначность структуры, требуется время для поиска нужной категории, количество сайтов в каталоге недостаточно большое.

Особенность каталогов в том, что они более эффективны при поиске подборок информации на определенную общую тему, например, "развитие образования", "состояние промышленности", но не при поиске ответа на конкретный вопрос!

Каталоги могут быть:

1. специализированными,
2. универсальными.

Специализированные каталоги включают только ссылки на сайты определенной, узкой тематики. Подобные каталоги удобны для поиска информации по конкретной тематике. Пример специализированного каталога: RadioLink - каталог ресурсов о телекоммуникациях.

Универсальные каталоги позволяют производить поиск по различным темам. Информация сгруппирована по разделам. Каждый раздел имеет несколько подразделов. Проблема классификации каталогов. В настоящее время в интернете функционируют миллионы каталогов. Они неравнозначны по качеству исполнения, информационному содержанию и функциональности. Значительная часть каталогов не производит индексацию сайтов и, следовательно, не производит поиск по их содержимому. Необходимо различать обычные каталоги от "продвинутых".

"Продвинутые" каталоги отличаются расширенной, а иногда и полной индексацией содержимого сайтов. К имеющимся возможностям каталогов (рубрикатор и поиск по аннотациям) добавляются мощные механизмы поиска в материалах сайтов.

ЗАКЛЮЧЕНИЕ

По итогам сделанной мной работы я могу заключить что; поисковые системы уже давно стали неотъемлемой частью Интернета. Поисковые системы сейчас – это огромные и сложные механизмы, представляющие собой не только инструмент поиска информации, но и заманчивые сферы для бизнеса.

По моему мнению, самой лучшей иностранной поисковой системой является Google , так как для меня основное значение имеет точность и полнота предоставляемых данных. Но можно заключить также что, каждая поисковая система будь то Российская или зарубежная предоставляет различные возможности поиска, из различных баз данных, поэтому сказать точно какой именно лучше пользоваться было бы неправильно. Поэтому для удобства поиска и полноты информации следует пользоваться несколькими поисковиками вводя в них нужные запросы. По моему мнению, из многих Российских поисковиков выделяются Яндекс и Рамблер, для них характерно постоянное обновление баз данных что, обеспечивает именно актуальность и точность предоставляемой информации.

СПИСОК ИСТОЧНИКОВ ЛИТРАТУРЫ

Примечание

1.