

Содержание:

Введение

Поисковые системы уже давно стали неотъемлемой частью российского Интернета. Поисковые системы представляют собой огромные и сложные механизмы, которые являются не только инструментом поиска информации, но и заманчивые сферы для бизнеса.

Большинство пользователей поисковых систем никогда не задумывались о принципе их работы, о схеме обработки запросов, выдачи результатов и о том, из чего эти системы состоят и как функционируют.

Актуальность работы состоит в том, что поисковые системы постоянно используются в современном мире. Если пользователь умеет правильно составлять запрос и понимает структуру работы поисковых систем, улучшается качество и скорость работы.

Цель работы: теоретический анализ работы поисковых систем

Задачи:

- раскрыть понятие поисковых систем;
- изучить характеристики и принципы работы поисковых систем.

Основой теоретических исследований были работы А.Б. Экслера, А.Б. Егорова, В.С. Гусева.

1.

1 Понятие и функции поисковой системы

1.1 Определение поисковых систем

Поисковая система - это программно-аппаратный комплекс, предназначенный для поиска информации в сети Интернет, отвечающий на запрос пользователя.

Пользователь задает запрос в виде текстовой фразы (поискового запроса), затем осуществляется выдача списка ссылок на источники информации, в порядке релевантности (в соответствии запросу).

Наиболее крупные международные поисковые системы: «Google», «Yahoo», «MSN». В русском Интернете это – «Яндекс», «Рамблер», «Апорт».

Для поиска информации с помощью поисковой машины пользователь формулирует поисковый запрос. Работа поисковой машины заключается в том, чтобы по запросу пользователя найти информацию, содержащую либо указанные ключевые слова, либо слова, как-либо связанные с ключевыми фразами. При этом поисковая машина генерирует страницу результатов поиска. Такая поисковая подборка может содержать различные типы результатов, например: веб-страницы, изображения, аудиофайлы. Некоторые поисковые системы также извлекают информацию из подходящих баз данных и каталогов ресурсов в Интернете.

Для получения наиболее качественного результата по запросу, пользователь должен корректно сформулировать запрос, в соответствии с тем, что он хочет найти. Запрос должен быть максимально кратким и простым для понятия. Допустим, необходимо найти информацию в «Яндексе» о том, как выбрать автомобиль. Для этого, нужно открыть главную страницу «Яндекса», и ввести текст поискового запроса «как выбрать машину». После введения запроса пользователю предоставляется возможность выбрать из найденных результатов ссылки с интересующей его информацией.

Однако, если запрос составлен некорректно и найденная информация не подходит пользователю, то либо нужно перефразировать свой запрос, либо в базе поисковой системы действительно нет никакой актуальной информации по заданному запросу (такое может быть при задании очень «узких» запросов, как, например, «как выбрать машину в городе Кирсанов»).

Основная задача любой поисковой системы – предоставление информации, необходимой пользователям. Обучить всех пользователей составлять корректные запросы, соответствующие принципам работы поисковых систем, практически невозможно. В связи с этим, разработчики создают такие алгоритмы и принципы работы поисковых систем, которые бы позволяли находить пользователям информацию, в которой они нуждаются.

Это означает, что поисковая система должна «думать» так же, как думает пользователь при поиске информации. Когда пользователь обращается с запросом

к поисковой машине, он хочет найти то, что ему нужно, максимально быстро и просто. Получая результат, он оценивает работу системы, руководствуясь несколькими основными параметрами: «Нашел ли он то, что искал? Если не нашел, то сколько раз ему пришлось перефразировать запрос, чтобы найти искомое? Насколько актуальную информацию он смог найти? Насколько быстро обрабатывала запрос поисковая машина? Насколько удобно были представлены результаты поиска? Был ли искомый результат первым или же сотым? Как много ненужного мусора было найдено наравне с полезной информацией? Найдется ли нужная информация, при обращении к поисковой системе, скажем, через неделю, или через месяц?».

Так же важно отметить что поисковая система тем лучше, чем больше документов, релевантных запросу пользователя, она будет возвращать. Результаты поиска могут становиться менее релевантными из-за особенностей или вследствие человеческого фактора.

Пример работы запроса представлен на рисунке 1.

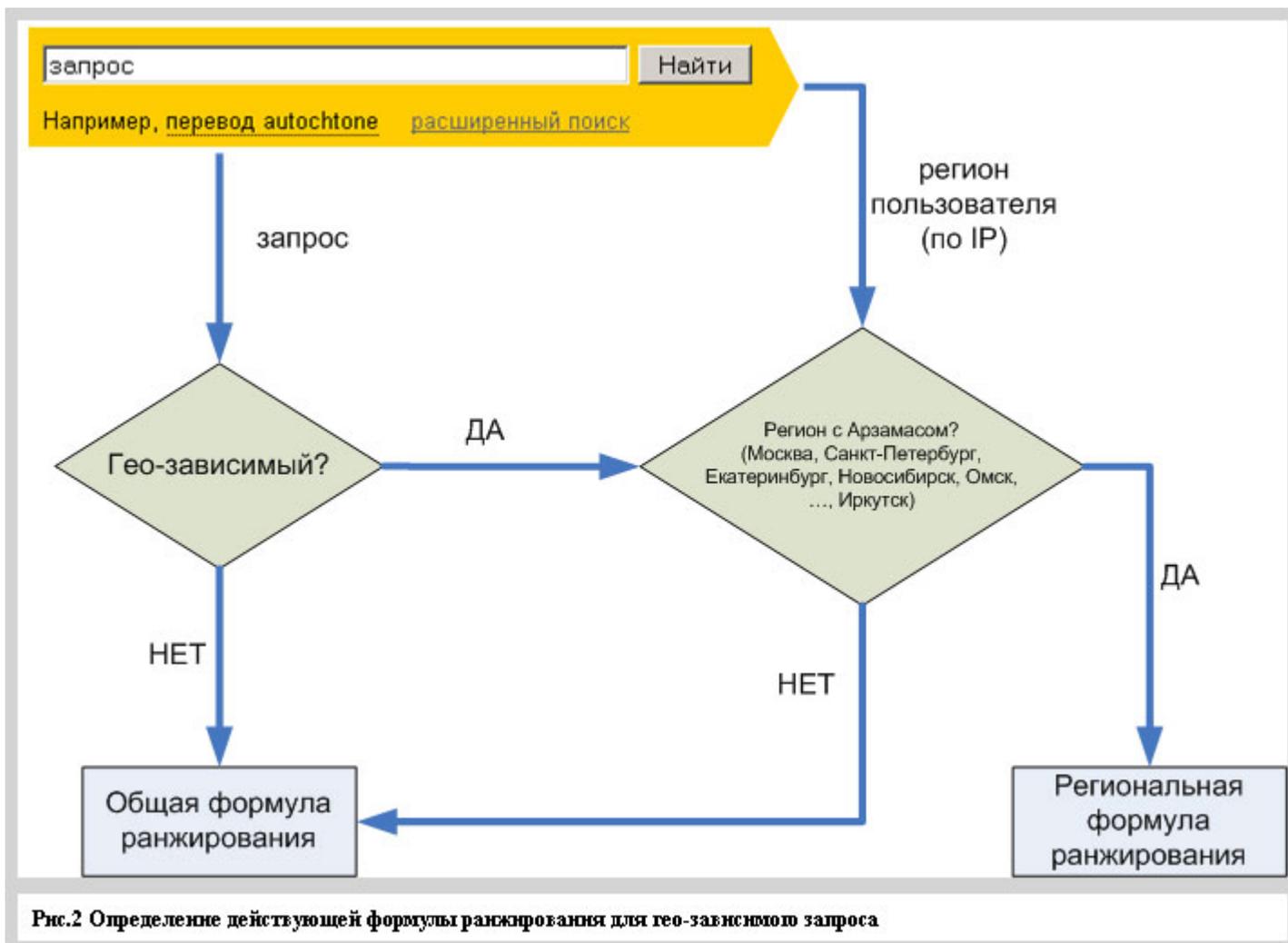


Рисунок 1. Ранжирование результатов запроса по геологическому признаку.

По состоянию на 2015 год самой популярной поисковой системой в мире является Google, однако есть страны, где пользователи отдали предпочтение другим поисковикам. Так, например, в России «Яндекс» обгоняет Google больше, чем на 10 %.

Таблица 1.

Сведения популярности рынка международных поисковых систем

Поисковая система	Процент рынка в июле 2014	Процент рынка в октябре 2014	Процент рынка в сентябре 2015
Google	68,69 %	58.01 %	69.24%

Поисковая система	Процент рынка в июле 2014	Процент рынка в октябре 2014	Процент рынка в сентябре 2015
Baidu	17,17 %	29.06 %	6,48%
Bing	6.22 %	8.01 %	12,26%
Yahoo!	6.74 %	4.01 %	9,19%
AOL	0,13 %	0.21 %	1,11%
Excite	0.22 %	0,00 %	0.00%
Ask	0,13 %	0,10 %	0,24%

Россия и русскоязычные поисковые системы

В России популярность поисковых систем среди пользователей выглядит иначе. Большая часть рынка принадлежит Яндексу – его доля составляет 56,2 %, а Google, в свою очередь, занимает всего 37,6 %.

Данные распределения популярности международных поисковых систем в России [\[1\]](#):

- Google (37,6 %);
- Bing (0,3 %);
- Yahoo! (0,1 %) и принадлежащие этой компании поисковые машины: Inktomi, AltaVista, Alltheweb.

1.2 Введение в историю информационных систем

На раннем этапе развития сети Интернет Тим Бернерс-Ли поддерживал список веб-серверов, размещённый на сайте ЦЕРН. Со временем, список сайтов расширялся, и

поддерживать вручную такой список становилось всё сложнее. На сайте NCSA был специальный раздел «Что нового!» (англ. What's New!), где публиковали ссылки на новые сайты.

Первой компьютерной программой для поиска в Интернете, была программа Арчи (англ. archie — архив без буквы «в»). Программа была создана в 1990 году студентами, изучающими информатику в университете Макгилла в Монреале - Аланом Эмтэджем (Alan Emtage), Биллом Хиланом (Bill Heelan) и Дж. Питером Дойчем (J. Peter Deutsch).

Программа скачивала списки всех файлов со всех доступных анонимных FTP-серверов и строила базу данных, в которой можно было выполнять поиск по именам файлов. Однако, программа Арчи не индексировала содержание этих файлов, так как объём данных был настолько мал, что всё можно было легко найти вручную.

Распространение и развитие сетевого протокола Gopher, придуманного в 1991 году Марком Маккэхилом (Mark McCahill) в университете Миннесоты, привело к появлению двух новых поисковых программ, Veronica и Jughead. Как и Арчи, они искали имена файлов и заголовки, сохранённые в индексных системах Gopher. Veronica (англ. Very Easy Rodent-Oriented Net-wide Index to Computerized Archives) позволяла выполнять поиск по ключевым фразам большинства заголовков меню Gopher во всех списках Gopher. Программа Jughead (англ. Jonzy's Universal Gopher Hierarchy Excavation And Display) извлекала информацию о меню от определённых Gopher-серверов.

Хотя, вручную поддерживались многочисленные специализированные каталоги, к лету 1993 года ещё не было ни одной машины для поиска в интернете. Оскар Нирштрасс (Oscar Nierstrasz) в Женевском университете написал ряд сценариев на Perl, которые периодически копировали эти страницы и переписывали их в стандартизированный формат. Это стало основой для W3Catalog, первой примитивной поисковой системы сети, запущенной 2 сентября 1993 года.

Считается, что первым поисковым роботом, написанным на языке Perl, был «World Wide Web Wanderer» — бот Мэтью Грэя (Matthew Gray) из Массачусетского технологического института в июне 1993 года. Этот робот создавал поисковый индекс «Wandex».

Цель робота Wanderer состояла в том, чтобы определить размер всемирной паутины и найти все веб-страницы, содержащие слова из запроса. В 1993 году

появилась вторая поисковая система «Aliweb».

Aliweb ожидала уведомлений от администраторов веб-сайтов о наличии на их сайтах индексного файла в определённом формате, при этом не используя поисковых роботов.

В декабре 1993 года Дж. Флетчер создал JumpStation. Принцип работы исходил из поиска веб-страниц и построения их индексов с помощью поискового робота, и использования веб-форм в качестве интерфейса для формулирования поисковых запросов.

Это был первый инструмент поиска в Интернете, который сочетал три важнейших функции поисковой системы:

- проверка информации;
- индексация;
- поиск.

Из-за ограниченности ресурсов компьютеров того времени индексация и, соответственно, поиск были ограничены только названиями и заголовками веб-страниц, найденных поисковым роботом.

Первой полнотекстовой поисковой машиной, проводящей индексацию ресурсов при помощи робота («crawler-based»), стала система «WebCrawler», запущенная в 1994 году.

В отличие от своих предшественниц, она давала возможность пользователям искать информацию, по любым словам, расположенным на любой веб-странице — с тех пор это стало являться стандартом для большинства поисковых систем. А также, это был первый поисковик, получивший широкое распространение. В 1994 году была запущена система «Lycos», разработанная в Университете Карнеги-Меллон и ставшая достаточно серьёзным коммерческим предприятием.

Вскоре появилось множество других конкурирующих поисковых машин, таких как: «Magellan», «Excite», «Infoseek», «Inktomi», «Northern Light» и «AltaVista». В определенном смысле, они конкурировали с популярными интернет-каталогами, такими как «Yahoo!». Но, поисковые возможности каталогов могли ограничиваться поиском по самим каталогам, а не по текстам веб-страниц. Позже каталоги объединялись или снабжались поисковыми роботами с целью улучшения поиска.

В 1996 году компания Netscape хотела заключить уникальную сделку с одной из поисковых систем, установив её поисковой системой по умолчанию на веб-браузере Netscape. Это привлекло настолько большой интерес, что Netscape заключила контракт сразу с пятью крупнейшими поисковыми системами (Yahoo!, Magellan, Lycos, Infoseek и Excite). За пять миллионов долларов США в год они предлагались по очереди на поисковой странице Netscape.

Поисковые машины участвовали в «Пузыре доткомов» конца 1990-х. Несколько компаний эффектно вышли на рынок, получив рекордную прибыль во время их первичного публичного предложения. Некоторые отказались от рынка общедоступных поисковых движков и стали работать только с корпоративным сектором, например, Northern Light.

В 1998 году Google взял на вооружение идею продажи ключевых слов. Тогда это была маленькая компания, обеспечивающая работу поисковой системы по адресу goto.com. Этот шаг ознаменовал для поисковых систем переход от соревнований друг с другом к одному из самых выгодных коммерческих предприятий в сети Интернет. Поисковые системы стали продавать первые места в результатах поиска отдельным компаниям.

В настоящее время существуют три основные международные поисковые системы – Google, Yahoo и MSN, имеющие собственные базы данных и алгоритмы поиска. Большинство остальных поисковых систем используют, в том или ином виде, результаты трех вышеперечисленных систем. Например, поиск AOL (search.aol.com) использует базу Google, а AltaVista, Lycos и AllTheWeb – базу Yahoo.

Стоит отметить, что поисковая машина Google занимает первые позиции с начала 2000-х. Компания добилась высокого положения благодаря хорошим результатам поиска с помощью алгоритма PageRank. Алгоритм был представлен общественности в статье «The Anatomy of Search Engine», написанной Сергеем Брином и Ларри Пейджем, основателями Google. Этот итеративный алгоритм ранжирует веб-страницы, основываясь на оценке количества гиперссылок на веб-страницу в предположении, что на «хорошие» и «важные» страницы ссылаются больше, чем на другие.

Интерфейс Google выдержан в сдержанном стиле, где нет ничего лишнего, в отличие от многих своих конкурентов, которые встраивали поисковую систему в веб-портал. Поисковая машина Google стала настолько популярной, что появились системы, подражающие ей, например, Mystery Seeker (тайный поисковик).

К 2000 году Yahoo! осуществлял поиск на основе системы Inktomi. Yahoo! в 2002 году купил Inktomi, а в 2003 году купил Overture, которому принадлежали AlltheWeb и AltaVista. Затем Yahoo! работал на основе поисковой системы Google вплоть до 2004 года, пока не запустил, свой собственный поисковик на основе всех купленных ранее технологий.

Осенью 1998 года фирма Microsoft впервые запустила поисковую систему Microsoft Network Search (MSN Search), используя результаты поиска от Inktomi. В начале 1999 года сайт начал отображать выдачу Looksmart, смешанную с результатами Inktomi. Также MSN search использовал результаты поиска от AltaVista.

В 2004 году фирма Microsoft начала переход к собственной поисковой технологии с использованием собственного поискового робота — msnbot. После проведения ребрендинга 1 июня 2009 года компанией Microsoft была запущена поисковая машина Bing. 29 июля 2009 года Yahoo! и Microsoft подписали соглашение, согласно которому Yahoo! Search работал на основе технологии Microsoft Bing. На момент 2015 года союз Bing и Yahoo! дал первые настоящие плоды. Теперь Bing занимает 20,1 % рынка, а Yahoo! 12,7 %, что в общем занимает 32,60 % от общего рынка поисковых систем в США.

1.3 Поиск информации на русском языке

В 1996 году был реализован поиск с учётом русской морфологии на поисковой машине Altavista и запущены оригинальные российские поисковые машины Рамблер и Апорт.

Поисковая машина Яндекс была открыта 23 сентября 1997 года. 22 мая 2014 года компанией Ростелеком была открыта национальная поисковая машина Спутник, которая на момент 2015 года находится в стадии бета-тестирования. 22 апреля 2015 года был открыт новый сервис «Спутник. Дети» специально для детей, с повышенной безопасностью.

Примеры работ поисковых систем Спутник и Яндекс можно увидеть на рисунках 2 и 3.

Как выбрать ноутбук

[http://www.eldorado.ru/...](http://www.eldorado.ru/)

Интернет-магазин электроники и бытовой техники: телевизоры, холодильники, компьютеры, **ноутбуки**, стиральные машины и телефоны. Список сервис-центров. Что **выбрать ноутбук** или нетбук.

Notebook.ru - ноутбуки цены, выбрать и купить ноутбук...

<http://www.notebook.ru/>

Интернет-сервис поиска и подбора **ноутбуков**, планшетов, нетбуков. Поисковая система, информация о производителях, отзывы. **Ноутбук выбрать** по параметрам. Как **выбрать ноутбук** для дома.

🕒 Сегодня: 10:00-19:00, без обеда (открыто)

☎️ +7 (495) 5454939

📍 Москва, Верейская, 29 ст154

ROZETKA — Новости, статьи и обзоры / Советы интернет...

[http://rozetka.com.ua/...](http://rozetka.com.ua/)

Фото-, видео- и аудиотехника, **ноутбуки**, компьютерные комплектующие и др. Условия оплаты и доставки. ...время и силы на изучение данной операционной системы, то логичнее будет переустановить систему самостоятельно или **выбрать ноутбук**...

Выбор ноутбука. Как выбрать ноутбук. Основные...

http://www.aimax.ru/choice_noteb

Продажа нетбуков, **ноутбуков**, коммуникаторов. Планшетные компьютеры Acer, Asus, Dell. Электронные книги. Во-вторых, при выборе **ноутбука** нужно правильно **выбрать** процессор.

Ноутбуки - лучшие предложения и цены. Где купить...

<http://torg.mail.ru/notebooks/>

Только лучшие предложения где купить **ноутбуки**. Обсуждения и отзывы, широкий выбор моделей и удобный подбор по параметрам на Товары@Mail.ru.

Как выбрать ноутбук для дома, работы и игр

[http://masterservis24.ru/...](http://masterservis24.ru/)

Следовательно, выбирайте тот **ноутбук**, у которого будет меньше всего оперативной памяти, но при этом будет большой потенциал. Какой **выбрать ноутбук**.

Как выбрать ноутбук

[http://www.computerhom.ru/...](http://www.computerhom.ru/)

Какой **ноутбук выбрать**. Если вы плохо разбираетесь в **ноутбуках**, то лучше выбирайте **ноутбук** с типом видеопамять – дискретное. ...внимание, так как на экране **ноутбука** зависит качество выводимой картинки на экран **ноутбука**.

Как выбрать ноутбук для игр - лучшие советы

[http://entercomputers.ru/...](http://entercomputers.ru/)

1 **Как выбрать ноутбук** — формат **ноутбука**. А теперь мы рассмотрим каким должен быть **ноутбук** в зависимости от его назначения:

Как выбрать ноутбук

http://www.sun-hands.ru/23kak_v

Выбор хорошего **ноутбука**, приспособленного для ваших нужд, - это очень важно. Это не очень легкая задача, которая требует от вас анализа собственных потребностей и выбора подходящих решений.

Как выбрать ноутбук

[http://antonkozlov.ru/...](http://antonkozlov.ru/)

Прежде чем **выбрать ноутбук**, необходимо определиться с собственными финансовыми возможностями и основными требованиями к устройству, потому что такое устройство. Например, **как выбрать ноутбук** для дома или для работы?

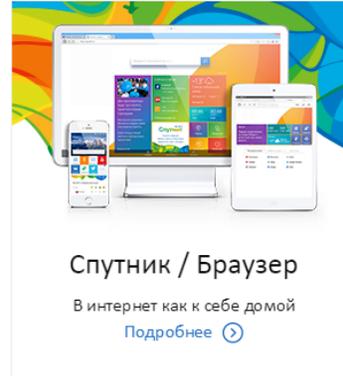


Рисунок 2. Поисковая система «Спутник».

Поиск

 **Выбираете ноутбук? / intel.ru**Реклама [intel.ru/Процессоры-Core](#)

Переходите на устройства 2 в 1 на базе лучшего процессора в истории Intel!

Устройства 2в1 Процессор Intel Intel & Windows 10 Сайт Intel

Картинки

Видео

До 10ч без подзарядки · HD графика · Быстрое включение · 3Дигры на 40%лучше

Карты

Маркет

Ещё

 **НОТИК: Выбираете ноутбук? / notik.ru**Реклама [notik.ru](#)

Поможем подобрать ноутбук для дома/работы/учебы. Ведущие бренды. Звоните!

Новинки Акции Доставка за рубль Услуги

Контактная информация · +7 (495) 777-01-99 · пн-пт 10:00-20:00, сб-вс 11:00-19:00

★★★★★ Магазин на Маркете · м. Рижская

 **Ноутбуки ASUS – Выбирайте с Яндекс.Маркетом**Реклама [market.yandex.ru](#)

Рекомендации. Отзывы. Актуальные цены.

[Ноутбуки Lenovo](#) [Ноутбуки Acer](#) [Ноутбуки ASUS](#) [Ноутбуки HP](#) **Как выбрать ноутбук для дома, работы и игр**[masterservis24.ru](#) > [13-kak-vybrat-noutbuk.html](#)

Как выбрать ноутбук по параметрам. Выбрать идеальный ноутбук для дома, работы или игр, Вам помогут технические характеристики (параметры) ноутбука...

 **Как выбрать ноутбук 2015-2016, а не тупо купить ноутбук...**[yoursputnik.ru](#) > [Как выбрать ноутбук](#)

Тупо купить ноутбук Asus, HP или разобраться всё-таки, как выбрать ноутбук правильно, лучше, хороший, по параметрам и т.д...

 **Как правильно выбрать самый лучший ноутбук**[dxdigital.info](#) > [2013/04/vibor-noutbuka-2013.html](#)

Как выбрать хороший надежный ноутбук? Кто производит самые лучшие ноутбуки - Asus, Acer, Lenovo, HP, Dell, Toshiba, Samsung, Sony, Apple?

 **Как выбрать ноутбук, лучшие ноутбуки 2015-2016 года**[world-mans.ru](#) > [kak-pravilno-vybrat-noutbuk...](#)

Задаваясь вопросом, как правильно выбрать ноутбук для работы, важно также обратить внимание на количество ОЗУ.

 **КАК ВЫБРАТЬ НОУТБУК, советы по выбору**[Ros-Potreb.ru](#) > [41-kak-vybrat-noutbuk.html](#)

Как выбрать ноутбук? Давайте узнаем, на что нужно обратить внимание при выборе ноутбука

 **Как выбрать ноутбук для дома и не пожалеть потом**[it-like.ru](#) > [kak-vyibrat-noutbuk-dlya-doma/](#)

Выбирайте ноутбук с углом наклона крышки от 130 градусов. ... Итак, я довольно подробно рассказал как выбрать ноутбук для дома.

 **Как выбрать ноутбук правильно. По каким параметрам...**[OtvetKak.ru](#) > [home-comfort/kak-vybrat-noutbuk-po...](#)

Перед тем как выбрать ноутбук, обратите на это внимание. ... Видеокарта в ноутбуке является встроенной, заменить ее или выбрать отдельно нельзя.

 **Как выбрать ноутбук. Какой лучше выбрать ноутбук?**[2notebook.net](#) > [content/articles/kak-vybrat-noutbuk](#)

Что выбрать: ноутбук или стационарный компьютер? Ассортимент компьютерной техники сегодня велик как никогда.

 **Как выбрать ноутбук**[eldorado.ru](#) > [research_center/article/detail.php...](#)

Как выбрать ноутбук. Отправляясь в магазин за новым ноутбуком, можно растеряться, глядя на множество представленных моделей.

 **Как выбрать ноутбук: какой лучше купить для игр, дома...**[where-money.com](#) > [gaznoe/kak-vybrat-noutbuk-kakoj...](#)

Как выбрать хороший ноутбук для дома, игр, работы и учебы? Какой ноутбук лучше купить бюджетный недорогой или самый мощный игровой?

 **Продайте старый ноутбук дорого! – Деньги сразу!**Реклама [nb100.ru](#)

Выкупаем ноутбуки в любом состоянии! Оценка по телефону - 5мин

Выезд оценщика 0 руб · Перенос данных 0 руб · Деньги сразу · Дорого

Контактная информация · +7 (495) 150-35-12 · пн-пт 10:00-18:00

Выбрать ноутбук / kupim-noutbuk.ruРеклама [kupim-noutbuk.ru](#)

Петли для ноутбука asus, петли для ноутбука HP, купить петли для ноутбука

 **Выберите ноутбук по низкой цене! – 220 магазинов техники**Реклама [budenovskiy.ru](#)

Помощь в выборе ноутбуков. Низкие цены! Приходите!

Нашлось 118 млн результатов

13 тыс. показов в месяц

[Дать объявление](#) [Показать все](#)

Рисунок 3. Пример работы поисковой машины «Яндекс».

Большую популярность получили методы кластерного анализа и поиска по метаданным. Из международных систем такого плана наибольшую известность получила «Clusty» компании Vivisimo. В 2005 году в России при поддержке МГУ запущен поисковик «Нигма», поддерживающий автоматическую кластеризацию и экспериментирующий с визуальной кластеризацией. В 2006 году открылась российская метамашин Quintura, предлагающая визуальную кластеризацию в виде облака тегов.

Местонахождение поисковых систем среди ключевых уровней интернета можно увидеть на рисунке 3.



Рисунок 3. Ключевые уровни интернета.

1.

2 Основные свойства поисковых систем

2.1 Характеристики поисковых систем

Поисковые системы обладают основными характеристиками:

1. Полнота;

Одна из основных характеристик поисковой системы – это полнота. Представляет собой отношение количества найденных по запросу документов к общему числу документов в сети Интернет, удовлетворяющих данному запросу. Например, если в сети Интернет имеется 100 страниц, содержащих словосочетание «как выбрать машину», а по соответствующему запросу было найдено всего 60 из них, то полнота поиска будет равна 0,6.

Из этого выходит, что чем полнее поиск по запросу, тем больше вероятность того, что пользователь найдет нужный ему документ, при условии, что он вообще существует в Интернете.

1. Точность;

Точность - еще одна важная характеристика поисковой системы, которая определяется степенью соответствия найденных документов по запросу пользователя. Для примера, если по запросу «как выбрать машину» находится сто документов, в пятидесяти из них содержится словосочетание «как выбрать машину», а в остальных просто присутствуют эти слова («как правильно выбрать радио и установить в машину»), то точность поиска считается равной $50/100 (=0,5)$.

Чем точнее поиск, тем быстрее пользователь может найти нужные ему документы, тем меньше лишней информации среди них будет встречаться, тем чаще найденные документы будут соответствовать запросу.

1. Актуальность;

Не менее важная составляющая поиска – это актуальность, которая характеризуется временем, которое проходит с момента публикации документов в сети Интернет, до занесения их в индексную базу поисковой машины. Например, на

следующий день после появления интересной новости, большое количество пользователей могут обратиться к поисковым машинам с соответствующими запросами. Объективно, с момента публикации новостной информации на эту тему прошло меньше суток, а основные документы уже были проиндексированы и доступны для поиска. Это осуществляется, благодаря существованию у крупных поисковых машин так называемой «быстрой базы», обновляющейся несколько раз в день.

1. Скорость поиска;

Скорость поиска зачастую связана с его устойчивостью к нагрузкам. Для примера, на сегодняшний день в рабочие часы к поисковой машине Рамблер приходит около 60 запросов в секунду[2]. Данная загруженность требует сокращения времени обработки каждого запроса.

Здесь интересы пользователя и поисковой системы совпадают: пользователь желает получить результаты как можно быстрее, а поисковая машина должна обрабатывать запросы в максимально короткие сроки, чтобы не затормаживать вычисление следующих запросов.

1. Наглядность;

Наглядность представления результатов - важный компонент удобного поиска. По многим запросам поисковая машина находит сотни, а иногда и тысячи результатов. Вследствие некорректно составленных запросов или неточности поиска, даже первые страницы выдачи не всегда могут содержать только нужную информацию. Это означает, что пользователь зачастую должен производить свой собственный поиск внутри найденного списка.

Различные элементы, содержащиеся на странице выдачи поисковой системы, помогают ориентироваться в результатах поиска.

Пример результата работы поисковой системы, обладающей всеми основными вышеперечисленными характеристиками, можно увидеть на рисунке 4.

Результатов: примерно 206 000 (0,41 сек.)

От него зависит нужный диаметр колёс и длина рамы **велосипеда**. Чтобы выяснить, какие колёса подходят вашему **ребёнку**, разделите его рост на 2,5 и ещё на 2,54 (длина 1 дюйма в см). Например, рост 110 см: $110 / 2,5 / 2,54 = 17,3$ дюйма. Следовательно, вам нужно **выбирать** из **велосипедов** с колёсами 16 или 18 дюймов. 29 мая 2015 г.

Как выбрать велосипед для ребёнка? | Памятки и инструкции ...
www.aif.ru/dontknows/instruction/kak_vybrat_velosiped_dlya_rebenka

[Подробнее...](#) • [Оставить отзыв](#)

Как выбрать велосипед для ребенка - Велосайт.ру - Велосипеды

www.velosite.ru/vibor_kid/ ▾

Как выбрать велосипед для ребенка? Детство без велосипеда — это как пирожок без начинки или соевый шоколад. Чего-то не хватает! Велосипед в ...

Как выбрать детский велосипед? | Полная информация

<https://dreambikes.ru/page/kak-vybrat-detskij-velosiped> ▾

В этой статье подробно рассмотрены все актуальные вопросы по выбору детского велосипеда - от классификации по размерам колес и росту ребенка, ...

Как выбрать велосипед для ребёнка? | Памятки и инструкции ...

www.aif.ru/dontknows/instruction/kak_vybrat_velosiped_dlya_rebenka ▾

29 мая 2015 г. - Чтобы у ребёнка не возникло проблем с управлением велосипедом и было меньше падений и травм, следует уделить особое ...

5 советов как выбрать детский велосипед - Домашний

www.domashniy.ru/semya_i_dety/5_sovetov_kak_vybrat_detskiy_velosiped/ ▾

Выбирайте велосипед, чья конструкция подходит ребёнку по возрасту и ... Следовательно, вам нужно выбирать из велосипедов с колесами 16 или 18 ...

Подбираем велосипед для ребёнка - Велосипеды

www.velostrana.ru > Подбор велосипеда ▾

Как выбрать велосипед для ребёнка | Купить велосипед в интернет-магазине VeloСтрана.

Как выбрать велосипед ребенку | Сайт учителя физической культуры

www.f-trener.ru/kak-vybrat-velosiped-rebenku ▾

Ребенок растет - велосипед должен к тому же становиться и более, красивее и быстрее. Как выбрать велосипед ребенку. Классификация детских и ...

Как правильно выбрать двухколесный велосипед для ребенка

www.detskiysad.ru/articles/073.html ▾

Как выбрать велосипед ребенку от 3-х лет? Самое главное правило: ни в коем случае не приобретайте велосипед «на вырост». Это может обернуться ...

Как правильно выбрать велосипед для себя и ребёнка? Советы ...

https://www.1tv.ru/.../94120-kak_pravilno_vybrat_velosiped_dlya_sebya_i_rebyonk... ▾

6 июн. 2012 г. - Крайнее оживление с началом лета наступило в магазинах, торгующих велосипедами. Что может быть лучше для прогулок по паркам ...

Выбор велосипеда 12, 14, 16, 18 или 20 дюймов? | Ма+Па игрушки ...

https://vk.com/topic-46273338_27990147 ▾

12 февр. 2013 г. - Если вы покупаете ребенку велосипед на годы, то не советуем экономить ... Эльвира, посоветуйте, какой велосипед выбрать 12 или 14 ...

Как правильно выбрать велосипед взрослый и детский по росту ...

remontvelosipeda.ru/kak-vybrat-velosiped-po-rostu/ ▾

Многие родители совсем не думают над вопросом как выбрать велосипед ребенку, хотя это немаловажный факт в развитие ребенка. Если выбрать ...

Вместе с как выбрать велосипед для ребенка часто ищут

- как выбрать велосипед для ребенка 3 лет
- как выбрать велосипед для ребенка 6 лет
- спортмастер
- как выбрать велосипед для ребенка 8 лет
- как выбрать велосипед для ребенка 1 год
- как выбрать велосипед для ребенка 10 лет
- как выбрать велосипед для ребенка 4 лет
- как выбрать велосипед для ребенка 2 года

Рисунок 4. Пример качественного результата работы поисковой системы Google.

2.2 Состав и принцип работы поисковых систем

Основной поисковой системой в России является «Яндекс», далее - Rambler.ru, Google.ru, Aport.ru, Mail.ru. На данный момент, Mail.ru использует базу поиска и механизм работы «Яндекса».

Практически каждая крупная поисковая машина имеют свою собственную структуру, отличную от других. Однако можно выделить общие основные компоненты, присущие для всех поисковых машин. Различия в структуре могут быть лишь в виде реализации механизмов взаимодействия этих компонентов.

1. Модуль индексирования

Модуль индексирования состоит из трех вспомогательных программ (роботов):

- **Spider** (паук) – программа, предназначенная для скачивания веб-страниц.

«Паук» обеспечивает скачивание страницы и занимается извлечением всех внутренних ссылок с этой страницы. Скачивается html-код каждой страницы. Для скачивания страниц используются протоколы HTTP.

Робот «паук» работает следующим образом – на сервер передается запрос “get/path/document” и некоторые другие команды HTTP-запроса. В ответ робот получает текстовый поток, который содержит служебную информацию и, непосредственно, сам документ.

Ссылки извлекаются из тэгов a, area, base, frame, frameset, и др. Наряду со ссылками, многими роботами обрабатываются и редиректы (перенаправления).

Каждая скачанная страница сохраняется в следующем формате:

- URL страницы;
- дата, когда страница была скачана;
- http-заголовок ответа сервера;
- тело страницы (html-код).
- **Crawler** («путешествующий» паук) – это программа, автоматически проходящая по всем ссылкам, найденным на странице.

Выделяет все ссылки, присутствующие на странице. Задача Crawler - определить, основываясь на ссылках или исходя из заранее заданного списка адресов, куда дальше должен идти паук. Следуя по найденным ссылкам, он осуществляет поиск новых документов, которые еще неизвестны поисковой системе.

- **Indexer** (робот - индексатор) – это программа, анализирующая веб-страницы, скачанные пауками.

Индексатор «разбирает» страницу на составные части и, применяя собственные лексические и морфологические алгоритмы, анализирует их. Анализу подвергаются различные элементы веб-страницы, такие как текст, заголовки, ссылки структурные и стилевые особенности, специальные служебные html-теги и т.д.

Таким способом, модуль индексирования позволяет обходить по ссылкам заданное множество ресурсов, скачивать встречающиеся страницы, извлекать ссылки на новые страницы из получаемых документов и проводить полный анализ данных документов.

1. База данных

Индекс поисковой системы или база данных - это система хранения данных - информационный массив, в котором специальным образом хранятся преобразованные параметры всех скачанных и обработанных модулем индексирования документов.

1. Поисковый сервер

Поисковый сервер - это важнейший элемент всей системы, так как от алгоритмов, лежащих в основе ее функционирования, напрямую зависит качество и быстрота поиска.

Поисковый сервер работает следующим образом:

Запрос, полученный от пользователя, подвергается морфологическому анализу. После генерируется информационное окружение каждого документа, который содержится в базе.

Полученные данные передаются в качестве входных параметров специальному модулю ранжирования. Происходит обработка данных по всем документам, в результате чего, для каждого документа рассчитывается персональный рейтинг,

характеризующий релевантность запроса, введенного пользователем, и различных составляющих этого документа, которые хранятся в индексе поисковой машины.

В зависимости от выбора пользователя, этот рейтинг может быть скорректирован дополнительными условиями, примером является «расширенный поиск».

Далее генерируется сниппет, это означает, что для каждого найденного документа из таблицы документов извлекаются заголовок, краткая аннотация, наиболее соответствующая запросу, и ссылка на сам документ, причем найденные слова подсвечиваются.

После, пользователю передаются полученные результаты поиска в виде SERP (Search Engine Result Page) – страницы выдачи поисковых результатов.

Очевидно, все эти компоненты тесно взаимосвязаны друг с другом и работают во взаимодействии, образуя четкий, достаточно сложный механизм работы поисковой машины, который требует огромных затрат ресурсов.

Обработка поискового запроса в системе «Рамблер» происходит, так, как это изображено на рисунке 5[3].

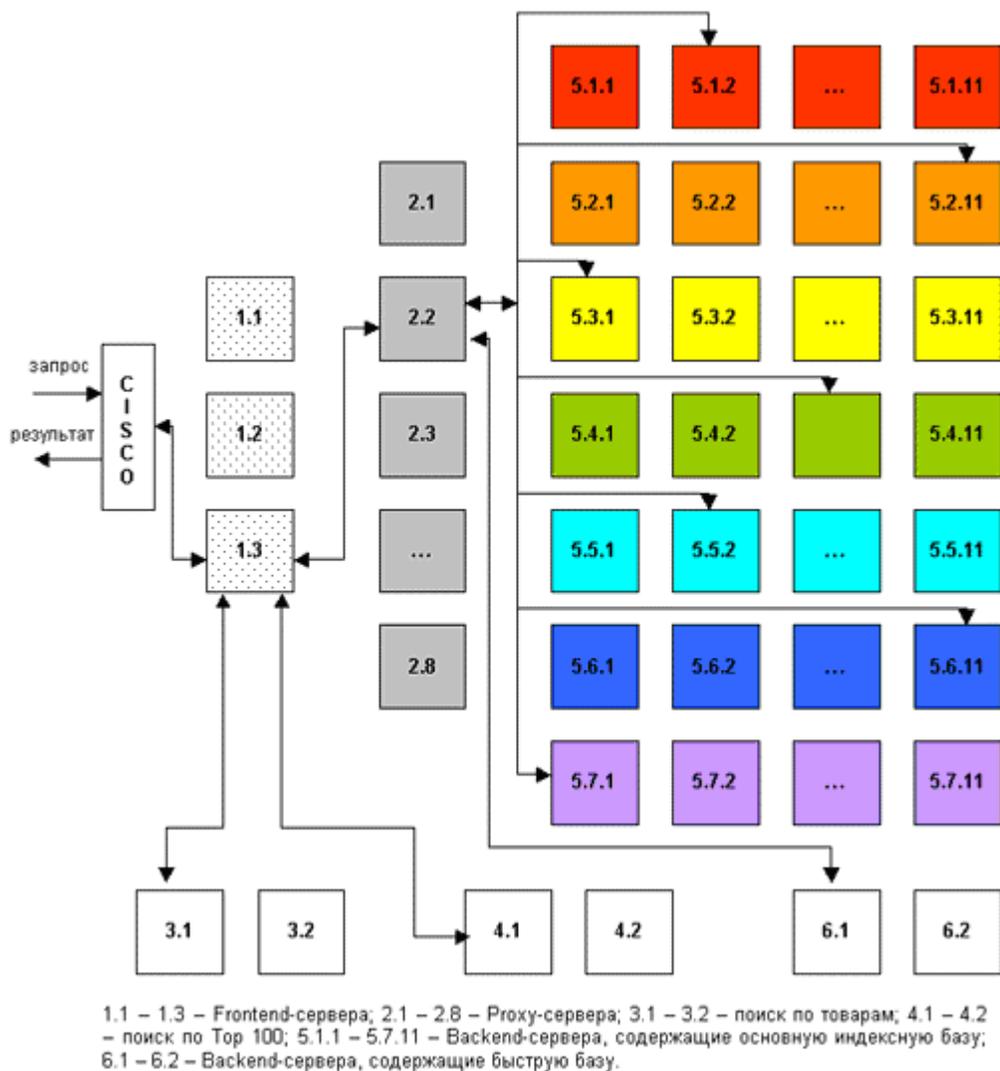


Рисунок 5. Обработка поискового запроса.

Запрос поступает в поисковую систему через маршрутизатор Cisco 6000 series. Затем Cisco передает его наименее загруженной машине первого уровня - frontend (1.1 - 1.3, на рис. машине 1.3). Машина Frontend, в свою очередь, отправляет запрос дальше, на один из восьми проху-серверов, также выбирая наиболее свободный сервер (2.1 - 2.8, на рис. машине 2.2). Одновременно frontend отправляет запрос на машины, осуществляющие поиск по товарам (3.1 - 3.2, на рис. машине 3.1) и по базе Top 100 (4.1 - 4.2, на рис. машине 4.1). На проху проводится поиск по ссылочному индексу, и его результаты вместе с поисковым запросом передаются на индексные машины - backends (5.1.x - 5.7.x, на рис. машинам 5.1.2, 5.2.11, 5.3.1 и т.д.). Та же информация отправляется на машины с "быстрой базой" (6.1 - 6.2).

На текущий момент в поиск включено 77 backend'ов. Они сгруппированы по 11 машин, и каждая группа содержит копию одной из частей поискового индекса.

Таким образом, информация о сайтах, условно входящих в красный сектор Интернета, находится на backend'ах первой группы (5.1.1 - 5.1.11 на рис), оранжевый сектор - на backend'ах второй группы (5.2.1 - 5.2.11) и т.д. Проху-сервер выбирает наименее загруженный backend в каждой группе машин и отправляет на него поисковый запрос с результатами ссылочного поиска. На backend'ах осуществляется поиск по частям индексной базы и ранжирование с учетом результатов поиска по ссылочному индексу. При ранжировании для всех найденных документов высчитывается вес по конкретному запросу.

После того, как запрос обработан на backend'ах, информация о результатах и ранжировании отдается обратно на проху-сервер. Туда же поступают отсортированные результаты с машин "быстрой базы". Проху интегрирует данные, полученные с восьми машин: склеивает дубли, объединяет зеркала сайтов, переранжирует документы в общий список по весу, рассчитанному на backend'ах. Итак, первым в списке найденного может быть документ с машины 5.3.1, вторым и третьим - с 6.1, четвертым - с 5.5.2 и т.д. На проху-сервере также реализуется построение цитат к документам и подсветка слов запроса в тексте. Полученные результаты отдаются на frontend.

Кроме информации с проху-сервера, frontend получает результаты из поиска по товарам и из базы Top 100, которые сортируются, с цитатами и подсветкой слов запроса. Frontend осуществляет окончательное объединение результатов, генерирует html со списком найденной информации, вставляет баннеры и перевязки и отдает html Cisco, который маршрутизирует информацию пользователю.

Принцип работы запросов можно увидеть на рисунке 6.



Рисунок 6. Принцип работы запроса.

Заключение

Изучив информационные системы, их состав и принцип работы, можно сделать некоторые выводы.

Поисковая система является программно-аппаратным комплексом, который служит для осуществления поиска информации в интернете и, реагирующий на запросы пользователей, которые задаются в виде текстовой фразы, и выдаются списком ссылок на источники информации, в порядке релевантности.

История появления поисковых систем начинается с первой полноценной поисковой машины, которой стал проект WebCrawler, вышедший в свет в 1994 году.

Первоочередная и наиболее важная задача любой поисковой системы – доставлять людям именно ту информацию, которую они ищут.

Основные и наиболее важные характеристики поисковых систем:

1. полнота;
2. точность;
3. актуальность;
4. скорость поиска;
5. наглядность.

В состав поисковой системы входят компоненты:

- модуль индексирования;
- база данных;
- поисковый сервер.

Список использованной литературы

1. Ашманов И. С., Иванов А. А. Продвижение сайта в поисковых системах. — М.: Вильямс, 2011. — 304 с.
2. Байков В.Д. Интернет. Поиск информации. Продвижение сайтов. — СПб.: БХВ-Петербург, 2013. — 288 с.
3. Гусев, В.С., Google. Эффективный поиск - Москва, Санкт - Петербург, Киев.: Диалектика, 2012. — 240 с.
4. Гусев, В.С., Поиск, Internet - Москва, Санкт - Петербург, Киев.: Диалектика, 2011. — 300 с.
5. Гусев, В.С., Яндекс. Эффективный поиск - Москва, Санкт - Петербург, Киев.: Диалектика, 2014. — 138 с.
6. Егоров, А.Б., Поиск в Интернете - Санкт - Петербург.: НиТ, 2014г. — 220 с.
7. Колисниченко Д. Н. Поисковые системы и продвижение сайтов в Интернете — М.: Диалектика, 2011. — 272 с.
8. Кузьмин А.В. Золотарева Н.Н. Поиск в Интернете - Санкт - Петербург.: Издательство НиТ, 2011. — 169 с.
9. Куприянова, Г.И., Информационные ресурсы Internet - М., 2012. — 145 с.
10. Ландэ Д. В. Поиск знаний в Internet. — М.: Диалектика, 2005. — 272 с.
11. Ландэ Д. В., Снарский А. А., Безсуднов И. В. Интернетика: Навигация в сложных сетях: модели и алгоритмы. — М.: Либроком (Editorial URSS), 2013. — 264 с.
12. Экслер, А.Б., Самоучитель работы в Интернете - Москва.: NT Press, 2014. — 302 с.

1. Согласно данным LiveInternet, июнь 2015 года. «Об охвате русскоязычных поисковых запросов». [↑](#)
2. По данным ООО «Рамблер Интернет Холдинг». [↑](#)
3. По информации ООО «Рамблер Интернет Холдинг». [↑](#)