

Технология поиска в сети интернет



*Презентация подготовил учащийся гр. 037
Сидоров Александр*

Содержание

1. Поиск информации
2. Поисковые индексы
3. Яндекс
4. Google
5. Rambler

Поиск информации – одна из самых востребованных на практике задач, которую приходится решать любому пользователю Интернета.

Существуют три основных способа поиска информации в Интернет:

1. Указание адреса страницы.
2. Передвижение по гиперссылкам.
3. Обращение к поисковой системе (поисковому серверу).

1: Указание адреса страницы-Это самый быстрый способ поиска, но его можно использовать только в том случае, если точно известен адрес документа или сайта, где расположен документ.

2: Передвижение по гиперссылкам-Это наименее удобный способ, так как с его помощью можно искать документы, только близкие по смыслу текущему документу.

3: Обращение к поисковой системе-Пользуясь гипертекстовыми ссылками, можно бесконечно долго путешествовать в информационном пространстве Сети, переходя от одной web-страницы к другой, но если учесть, что в мире созданы многие миллионы web-страниц, то найти на них нужную информацию таким способом вряд ли удастся.

На помощь приходят специальные **поисковые системы** (их еще называют **поисковыми машинами**). В настоящее время в русскоязычной части Интернет популярны следующие поисковые серверы: **Яндекс** (yandex.ru), **Google** (google.ru) и **Rambler** (rambler.ru).

Поисковые индексы

Поисковые индексы работают как алфавитные указатели. Клиент задает слово или группу слов, характеризующих его область поиска, и получает список ссылок на web-страницы, содержащие указанные термины.

Как работают поисковые индексы?

Поисковые индексы автоматически, при помощи специальных программ (веб-пауков), сканируют страницы Интернета и индексируют их, то есть заносят в свою огромную базу данных.

Поисковый робот («веб-паук») — программа, являющаяся составной частью поисковой системы и предназначенная для обхода страниц Интернета с целью занесения информации о них (ключевые слова) в базу поисковика. По своей сути паук больше всего напоминает обычный браузер. Он сканирует содержимое страницы, забрасывает его на сервер поисковой машины, которой принадлежит и отправляется по ссылкам на следующие страницы. В ответ на запрос, где найти нужную информацию, поисковый сервер возвращает список гиперссылок, ведущих web-страницам, на которых нужная информация имеется или упоминается. Обширность списка может быть любой, в зависимости от содержания запроса.

Яндекс

Поиск Яндекса позволяет искать по Рунету документы на русском, украинском, белорусском, румынском, английском, немецком и французском языках с учётом морфологии русского и английского языков и близости слов в предложении.

Отличительная особенность Яндекса — возможность точной настройки поискового запроса. Это реализовано за счёт гибкого языка запросов.

По умолчанию Яндекс выводит по 10 ссылок на каждой странице выдачи результатов, в настройках результатов поиска можно увеличить размер страницы до 20, 30 или 50 найденных документов.

Время от времени алгоритмы Яндекса, отвечающие за релевантность выдачи, меняются, что приводит к изменениям в результатах поисковых запросов. В частности, эти изменения направлены против поискового спама, приводящего к нерелевантным результатам по некоторым запросам.

Яндекс



Google

Лидер поисковых машин Интернета, Google занимает более 70 % мирового рынка. Сейчас регистрирует ежедневно около 50 млн поисковых запросов и индексирует более 8 млрд веб-страниц. Google может находить информацию на 115 языках. Интерфейс Google содержит довольно сложный язык запросов, позволяющий ограничить область поиска отдельными доменами, языками, типами файлов и т.

д. Google значительно выделяется в рядах поисковых систем-конкурентов своим предельно простым и легко узнаваемым интерфейсом стартовой страницы, на которой расположены только ссылки на основные базы данных Google, страницу настроек, расширенный поиск и языковые инструменты.

Для некоторых результатов поиска Google предоставляет повторное поле для поиска, которое позволяет пользователю найти то, что он ищет внутри конкретного веб-сайта.

авное назначение страницы расширенного поиска Google — упростить работу с комплексными запросами. В результате с помощью такого раздела можно составить довольно сложный запрос, не особо вникая в тонкости применяемого языка запросов.



Rambler

Rambler Media Group — интернет-холдинг, включающий в качестве сервисов поисковую систему, рейтинг-классификатор ресурсов российского Интернета, информационный портал.

Поисковая система Рамблер понимает и различает слова русского, английского и украинского языков. По умолчанию поиск ведётся по всем формам слова.

Поисковая система Рамблер при поиске учитывает морфологию русского языка, что дает больше возможностей для эффективного поиска информации.

Реализована также система так называемых "перевязок", которая позволяет выдавать в результатах поиска не только страницы содержащие запрос, но и слова, которые являются синонимами запроса. Еще одной функцией "перевязок", думаю более значимой, является выдача контекстной рекламы не только по конкретному запросу, но и по запросам, которые тесно связаны с исходным, это позволяет перекрыть большее количество целевой аудитории.

