

Поиск информации в Интернете

Система основных понятий

Поисковая служба Интернета			
Поисковые каталоги		Поисковые указатели	
<i>Накопление и систематизация информации о ресурсах</i>	<i>Способы поиска ресурсов пользователем</i>	<i>Накопление и систематизация информации о ресурсах</i>	<i>Способы поиска ресурсов пользователем</i>
Иерархический предметный каталог. Формируется вручную силами экспертов	Спуск поддерева каталога. Запросы по ключевым словам	Индексные списки (предметные указатели). Формируются автоматически	Запросы по ключевым словам

Справочный материал

Поисковые системы – это автоматические системы, опрашивающие серверы, подключенные к глобальной сети, и сохраняющие в своей базе информацию об имеющихся на серверах данных.

Как правило, поисковые системы состоят из трех частей: *робота*, *индекса*, и *программы обработки запроса*.

Робот (Spider, Robot или Bot) – это программа, которая посещает веб-страницы и считывает их содержимое, определяет ключевые слова и записывает в базу индексов данное слово с указателем на документ.

Индекс – это хранилище поисковых образов, посещенных роботами страниц. Поисковый образ – это описание содержания документа на специальном информационно-поисковом языке.

Программа обработки запроса – это программа, которая в соответствии с запросом пользователя «просматривает» индекс на предмет наличия нужной информации и возвращает ссылки на найденные документы. Множество ссылок на выходе из системы распределяется системой в порядке убывания релевантности, то есть от наибольшей степени соответствия ссылки запросу к наименьшей.

Название	Поисковый каталог	Поисковый указатель
Апорт	http://search.aport.ru/catalog/	http://search.aport.ru/
Рамблер	http://top100.rambler.ru/?src=main	http://www.rambler.ru/
Яндекс	http://yaca.yandex.ru/	http://www.yandex.ru/
Ау!	http://www.au.ru	

Основные характеристики поисковых систем:

- ✓ объем документов в индексе;
- ✓ частота обновления информации;
- ✓ информационное пространство, которое охватывает робот поисковой системы, и разнообразие типов документов, о которых собирается информация;
- ✓ скорость обработки запроса;
- ✓ критерии определения релевантности;
- ✓ возможность детализации и уточнения запроса.

При профессиональном поиске в интернете должны соблюдаться следующие требования:

- ✓ высокая скорость проведения поиска;
- ✓ достоверность получения информации;
- ✓ полнота охвата ресурсов при поиске.

Десять простых правил формирования запроса в поисковой системе Яндекс

№	Пример	Значение
1	Если в запросе было указано слово «знаю», то условию поиска будут удовлетворять и слова «знаем», «знаете» и т. п.	При поиске учитываются все формы слова по правилам русского языка не зависимо от формы слова в запросе
2	«фарфоровая посуда»	Для поиска устойчивого словосочетания следует заключить слова в кавычки
3	!тихий !Дон (Поиск документа, в котором есть слово тихий, начинающееся со строчной или заглавной буквы, или слово Дон, начинающееся с заглавной.)	Для поиска по точной словоформе перед словом надо поставить восклицательный знак
4	Приключенческий & роман	Для поиска в пределах одного предложения слова в запросе разделяются пробелом или знаком &
5	Волга +автомобиль	Для поиска документов, в которых обязательно встречается определённое слово, перед словом ставится знак «+»
6	Волга -автомобиль	Для исключения слова из результатов поиска перед словом ставится знак «-»
7	ребенок малыш младенец	Для поиска любого из перечисленных слов между словами ставится знак « »
8	книги ~ магазин	Знак «~» (тильда) позволяет найти документы с предложением, содержащим первое слово, но не содержащим второе
9	лев ~~ астрология (Будут найдены документы со словом «лев», не относящиеся к астрологии)	Если оператор повторяется один раз (например, & или ~), поиск производится в пределах предложения. Двойной оператор (&&, ~~) задает поиск в пределах документа
10	(компьютер ПК ЭВМ) ++(устройство архитектура)	Вместо одного слова в запросе можно подставить целое выражение. Для этого его нужно взять в скобки