

Поиск информации в интернете

Информация, размещенная во Всемирной сети, исчисляется огромным количеством байт. Для поиска информации во Всемирной сети используются специальные веб-сайты – информационно-поисковые системы. Они позволяют по ключевым словам найти информационные ресурсы, связанные с ключевыми словами. Это может быть текст, содержащий ключевые слова, или графическое изображение одного из ключевых слов. Примерами информационно-поисковых систем являются системы Google и Yandex.

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

Существуют три основных способа поиска информации в Интернете:

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

Способ 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.

Не стоит забывать возможность поиска по открытой в окне браузера web-странице (Правка-Найти на этой странице...).

Способ 2: Передвижение по гиперссылкам

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

Способ 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). Адреса поисковых серверов хорошо известны всем, кто работает в Интернете. В настоящее время в русскоязычной части Интернета популярны следующие поисковые серверы: Яндекс (yandex.ru), Google (google.ru) и Rambler (rambler.ru).

Поисковая система — веб-сайт, предоставляющий возможность поиска информации в Интернете.

Большинство поисковых систем ищут информацию на сайтах Всемирной паутины, но существуют также системы, способные искать файлы на ftp-серверах, товары в интернет-магазинах, а также информацию в группах новостей Usenet.

По принципу действия поисковые системы делятся на два типа: поисковые каталоги и поисковые индексы.

Поисковые каталоги служат для тематического поиска.

Информация на этих серверах структурирована по темам и подтемам. Имея намерение осветить какую-то узкую тему, нетрудно найти список web-страниц, ей посвященных.

Каталог ресурсов в Интернете или каталог интернет-ресурсов или просто интернет-каталог — структурированный набор ссылок на сайты с кратким их описанием.

Каталог в котором ссылки на сайты внутри категорий сортируются по популярности сайтов называется рейтинг (или топ).

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.

Первой поисковой системой для Всемирной паутины был «Wandex», уже не существующий индекс, разработанный Мэтью Грэйем из Массачусетского технологического института в 1993.

Как работает поисковой индекс?

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.

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

Яндекс — российская система поиска в Сети. Сайт компании, Yandex.ru, был открыт 23 сентября 1997 года. Головной офис компании находится в Москве. У компании есть офисы в Санкт-Петербурге, Екатеринбурге, Одессе и Киеве.

Количество сотрудников превышает 700 человек.

Слово «Яндекс» (состоящее из буквы «Я» и части слова index; обыгран тот факт, что русское местоимение «Я» соответствует английскому «I») придумал Илья Сегалович, один из основателей Яндекса, в настоящий момент занимающий должность технического директора компании.

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.

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

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.

По одной из версий, Google — искаженное написание английского слова googol. «Googol (гугол)» — это математический термин, обозначающий единицу со 100 нулями. Этот термин был придуман Милтоном Сироттой, племянником американского математика Эдварда Каснера, и впервые описан в книге Каснера и Джеймса Ньютона «Математика и воображение»(Mathematics and the Imagination). Использование этого термина компанией Google отражает задачу организовать огромные объемы информации в Интернете.

Интерфейс Google содержит довольно сложный язык запросов, позволяющий ограничить область поиска отдельными доменами, языками, типами файлов и т. д.

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

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

Rambler создан в 1996 году.

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