

## ***Поиск в Интернете***

Интернет – это хранилище огромного количества информации: тексты, рисунки, звуковые и видео файлы, программы и архивы... Однако такое многообразие таит в себе и сложности: как найти нужное? Поиск информации в Интернете – дело очень непростое и учиться ему надо на практике.

### **Основные способы поиска в Интернете**

Самый быстрый и удобный способ поиска в сети – это когда кто-то уже нашел все до Вас. Или Вы сами обнаружили в статье ссылку на нужный Вам сайт. Поэтому, прежде чем искать информацию в сети, поинтересуйтесь у окружающих – вдруг кто-то уже все нашел? Просмотрите обзоры веб-сайтов, которые публикуют сейчас многие журналы (практически в любом из периодических изданий есть раздел, где описывают наиболее интересные сайты по тематике издания). Сейчас даже выходят целые книги, посвященные обзорам сайтов различной тематики! Конечно, данные в такой литературе очень быстро устаревают – и это тоже надо помнить.

Но если никто не знает, где искать требующуюся Вам информацию, то Вам следует использовать поисковые сайты.

### **Виды поисковых систем**

Все поисковые сайты представляют собой базы данных, где хранятся сведения о веб-страницах. Чем же одни поисковики отличаются от других?

Основное различие – в способах пополнения этой базы.

**Каталоги** – это информационно-справочные системы, в которых информация о сайтах заносится специальными людьми – редакторами – в соответствующие разделы разветвленного рубрикатора. По «охвату» каталоги бывают универсальными, охватывая весь спектр ресурсов Интернет, или тематическими – где классифицированы только ресурсы, посвященные определенной тематике.

Как правило, в каталог попадают только проверенные редакторами, достаточно серьезные ресурсы. Это очевидный плюс таких поисковых систем.

Однако есть и недостатки. Первый из них напрямую вытекает из достоинств каталога: для того, чтобы сайт проверили редакторы, составили его описание и поместили ресурс в нужной рубрике – необходимо значительное время. Многие же из веб-сайтов и вовсе не заносятся в каталоги, так как редакторы не сочли их достойными этого. Таким образом, в каталоге находится намного меньше ресурсов, чем в Интернете в целом.

Второй минус каталогов – система рубрик не всегда однозначно подсказывает вам, где искать ресурс. Если же вы не очень точно представляете, к какому разделу относится то, что вы хотите найти – то и вовсе, разветвленный рубрикатор с легкостью запутает Вас.

#### **Российские каталоги**

- <http://list.mail.ru> – Каталог на Mail.ru – универсальный каталог Рунета.
- <http://www.kinder.ru> — Каталог детских ресурсов Рунета.
- <http://www.portal.edu.ru> — Федеральный портал «Российское образование».
- <http://school.holm.ru> — Школьный мир: каталог образовательных ресурсов.

#### **Международные каталоги**

- <http://www.yahoo.com> — Yahoo! (каталог ресурсов на английском языке).
- <http://www.yahooligans.com> — Yahoooligans – веб-путеводитель для детей (на английском языке).
- <http://www.ipl.org> — The Internet Public Library — цифровая публичная библиотека (на английском языке).

Недостатков каталога лишен второй тип поисковиков – **поисковые машины**. Здесь – наоборот – весь процесс пополнения базы осуществляется автоматически. Специальные программы – поисковые роботы («пауки») – путешествуют по всемирной паутине и добавляют описание новых страниц в базу данных поисковой системы. После этого пользователь набирает в специальном окне слова, которые должны встретиться на искомой странице – и ему выдается перечень всех ресурсов, где есть требуемые слова. На первый взгляд, поисковые машины по всем параметрам превосходят каталоги: новые сайты добавляются в базу очень быстро – в течение нескольких дней или недель; количество ресурсов в базе намного больше, чем у каталогов; не нужно придумывать, в какой рубрике может оказаться необходимый ресурс – просто вводим искомые слова – и все. Однако, как всегда из достоинств вытекают недостатки.

Так, из большого количества ресурсов, которые выдаст нам поисковая система на наш запрос, намного сложнее выбрать действительно полезные, чем из небольшого числа ссылок, которые даст нам каталог.

Таким образом, оба типа поисковых систем имеют свои ограничения и достоинства. Именно поэтому большая часть современных поисковых сайтов объединяет в себе и каталог, и поисковую машину.

#### **Наиболее популярные российские поисковые системы:**

- 
- Яндекс (<http://yandex.ru>)
- Рамблер (<http://rambler.ru>)
- Апорт (<http://aport.ru>)
- Russia On The Net (<http://www.ru>)
- All Stars (<http://www.stars.ru>)
- Улитка (<http://www.ulitka.ru>)
- АУ! (<http://www.au.ru>)

Каталог русскоязычных ресурсов, содержащий более 20000 ссылок

#### **Международные поисковики:**

- <http://altavista.com>
- <http://google.com> (поиск только в русскоязычной сети - Google.ru)

### **Технология поиска**

Поставим себе задачу найти текст выступления ректора МГУ им. Ломоносова на Всероссийском съезде учителей информатики.

1. Откройте окно поисковой системы **Яндекс**.
2. В строке **Поиск** введите фразу «выступления ректора мгу на всероссийском съезде учителей информатики».
3. Нажмите кнопку **Найти** или клавишу **Enter** на клавиатуре.

Через некоторое время загрузится страница с результатами поиска. Найденные документы отсортированы по релевантности – то есть, по соответствию запросу. Каждая поисковая система определяет это соответствие по своему алгоритму, секрет которого является ноу-хау компании. Однако общие принципы известны: чем чаще встречаются искомые слова, чем ближе они к началу страницы, содержатся ли нужные слова в заголовке страницы, в ключевых словах или описании – тем ближе к началу списка находится сайт.

Ссылки на найденные ресурсы содержат следующую информацию:

**Название страницы**, которое одновременно является и гиперссылкой на найденный документ.

**Фрагмент текста** страницы, в котором встречаются искомые слова – он позволяет определить, насколько нам подходит сайт, еще до того, как мы его откроем.

**Адрес URL** найденного документа. В скобках указывается **размер файла** и далее – **дата последнего обновления документа**, если ее может определить поисковый робот.

4. Скорее всего, первая же ссылка даст Вам искомый документ – сайт съезда. Обратите внимание на адрес страницы. Если это адрес МГУ (msu), то это, видимо, то, что надо. Если в адресе нет «msu», то, скорее всего, это новостная информация.
5. Найдите на странице ссылку на пленарные доклады, а далее – на выступление В.А. Садовниченко. Сохраните страницу в **Избранное**.

Хотя мы уже нашли то, что искали – посмотрим, что еще есть на странице.

Внизу – номера страниц с найденными ресурсами, переключатель сортировки по релевантности и по дате. Кроме того, Яндекс дает возможность перейти для поиска в другие популярные системы.

В правой части окна находится небольшая картинка и текст «искать в картинках» – это еще одна функция данного поисковика.

Сравним, как будут выполнять наш запрос другие поисковые системы. Щелкните на ссылке внизу страницы **В других поисковых системах – Google**.

В этой поисковой системе порядок расположения ссылок и их количество может отличаться. Но представление найденных ресурсов очень похоже – также имеется небольшой фрагмент текста, заголовок и адрес URL.

Небольшое отличие: ссылки открываются в том же окне. Поэтому не забывайте пользоваться кнопкой **Назад** браузера или открывать ссылки в новом окне через контекстное меню.

### **Сложный поиск. Язык поисковых запросов**

Итак, с простым поиском все ясно: вводи в строку искомую фразу – и открывай нужные документы. Но что делать, когда в ответ на команду **Найти** приходит – «по Вашему запросу ничего не найдено»? Или наоборот – поисковик выдает тысячи ссылок, но подавляющее большинство их Вам не подходит? Это значит, что Вы неверно формулируете запрос. В первом случае – задаете слишком жесткие условия для поиска. Во втором – слишком широкие.

Как же правильно сформулировать запрос?

#### **Случай первый – ничего не найдено**

1. Проверьте орфографию – возможно, Вы допустили ошибку в словах, и поэтому ничего не находится.
2. Используйте синонимы – например, вместо «Лондон» – попробуйте «Столица Великобритании». Можно указать все слова-синонимы в одном запросе через знак «или» - «|» (для Яндекс) - тогда будут найдены все документы, где встречается хотя бы одно из этих слов. Например: Москва | столица России | город-герой.
3. Используйте более широкое определение – например, вместо «окислительно-восстановительные реакции с участием хлора» попробуйте просто «окислительно-восстановительные реакции» или «реакции хлора».
4. Учтите, что некоторые поисковые системы обращают внимание на наличие заглавных букв. Так, если вы пишете все слово строчными буквами, будут найдены все варианты его написания; если Вы указали хотя бы одну букву в искомом слове прописной, то система будет искать только такие варианты.

#### **Обратная ситуация – найдено множество документов**

1. Ищите больше, чем по одному слову. Слово «история» или «поэзия» дадут Вам огромный набор ссылок, где эти слова встречаются в разнообразнейшем контексте. Добавьте одно-два слова, конкретизирующего запрос – например, «история в средней школе», «поэзия Лермонтова».
2. Конкретизируйте запрос – ищите в найденном. Такая функция есть во всех поисковых системах. Укажите добавочное ключевое слово, установите флажок

**Искать в найденном** и щелкните **Найти** – поиск будет производиться только среди ранее найденных документов.

3. Если Вы ищете конкретное словосочетание – заключите его в кавычки. Тогда будут найдены только те страницы, где слова встречаются только в данной конкретной форме и сочетании.
4. Используйте знаки «+» и «-» (для Яндекса). Чтобы исключить документы, где встречается определенное слово, поставьте перед ним минус. И наоборот, чтобы определенное слово обязательно присутствовало в документе, поставьте перед ним плюс. Обратите внимание, что между знаком и словом не должно быть пробела. Например, если Вы ищете информацию про химические реакции, но вот про их скорость Вам информация неинтересна – сформулируйте запрос так «химическая реакция -скорость».

Знак «+» стоит использовать в том случае, когда нужно найти так называемые стоп-слова (наиболее частотные слова русского языка, в основном местоимения, предлоги, частицы – как правило, поисковые системы эти слова игнорируют при учете соответствия запросу). Чтобы найти цитату из «Гамлета», введите «+**быть или +не быть**».

5. Если вы не уверены, какие слова использовать в запросе, можно использовать подстановочный символ «звездочка» – «\*». Он ставится на месте одного или нескольких слов, которые система должна подставить в ваш запрос.

## Язык поисковых запросов

Для задания сложного запроса во всех поисковых системах предусмотрена страница расширенного поиска. Ссылка на нее, как правило, находится где-то рядом с поисковой формой. В полях расширенного поиска Вы можете более детально сформулировать свой запрос. Но можно обойтись и обычной строкой поиска – достаточно использовать в своем запросе специальные символы языка поисковых запросов.

Мы уже познакомились с некоторыми из них – знаками «+» и «-», логическим «или» - «|». В разных поисковых системах операторы могут обозначаться разными символами, поэтому мы лишь разберем наиболее распространенные из них.

## Логическое «и», «или» и «не»

Несколько слов в строке поиска, разделенные пробелами, означают, что все они должны входить в текст документа. То же самое обозначает знак **&** между словами.

Если мы ищем документ, куда входит только одно из указанных слов – мы используем знак «или» - для Яндекса это «|».

Если же мы ищем документ, куда входит одно слово, но не входит другое – мы используем знак «не» - для Яндекса его обозначает символ «~» или «-». Например, строка *Пушкин ~ «Руслан и Людмила»* найдет нам все документы, где упоминается Пушкин, но нет упоминания об указанной поэме (*Пушкин –«Руслан и Людмила»*). Обратите внимание, что «-» пишется слитно со словом, которое Вы хотите исключить, и в конце запроса.

## Поиск слова в конкретной форме

По умолчанию поиск учитывает все формы заданного слова согласно правилам русского языка. Так, если Вы ввели в запросе слово *учебник* – будут найдены и *учебникам*, и *учебники* и все остальные словоформы. Как правило, это удобно – ведь мы не знаем точно, в какой форме встретится искомое слово.

Однако иногда возникает необходимость найти слово в конкретной форме – тогда нам поможет знак «!». Запрос в виде **!машиной** выдаст нам ссылки только на те документы, где слово *машина* встречается именно в таком виде.

Если нас интересует не одно слово, а устойчивое словосочетание – заключите его в кавычки – тогда поисковая система будет учитывать только те документы, где

словосочетание встречается в точно таком виде, как в поисковой строке. Этот вариант очень удобен для поиска цитат.

### **Учет расстояния между словами**

В общем случае слова, введенные в строку поискового запроса, могут находиться на любом расстоянии друг от друга. Так, по запросу *новые учебники* найдутся и строка *В школу поступили новые учебники*, и *В школу завезли новые парты, сделали ремонт, а учебников по-прежнему не хватает*.

Поиск устойчивого словосочетания тоже может не помочь – так, запрос «*новые учебники*» не позволит найти строку *новые школьные учебники*, которая нам на самом деле подходит.

В таком случае можно указать максимальное удаление слов в запросе друг от друга, которое следует учитывать при поиске.

Для Яндекса расстояние между словами указывается после знака /. Например, *новые /2 учебники* – означает, что расстояние между словами может быть не более двух слов.

### **Поиск в определенных зонах веб-документа**

С использованием языка запросов можно ограничить поиск определенными зонами веб-документа. Например, проводить поиск только в заголовке страницы, в ключевых словах или в описании. Также можно ограничить поиск только именами и описаниями файлов изображений - именно так организован поиск рисунков. Синтаксис таких запросов довольно сложен и различается для поисковых систем, поэтому мы не приводим здесь конкретного описания.

Если вы заинтересовались описанием расширенного поиска при помощи самостоятельно составленных сложных запросов – познакомьтесь с тем, как создаются такие запросы в различных поисковиках:

- Синтаксис сложных запросов на Яндексе - [http://www.yandex.ru/ya\\_detail.html](http://www.yandex.ru/ya_detail.html)
- Синтаксис сложных запросов на Рамблере - <http://www.rambler.ru/doc/help.shtml#1>
- Синтаксис сложных запросов на Google - <http://www.google.ru/intl/ru/help/refinerearch.html> (на английском)