

*Технология  
поиска информации  
в сети Интернет*

**Автор: учитель информатики ГУО СШ № 28 г.Бобруйска Кунцевич Л.А.**



# *Введение.*

**Первые шаги по поиску необходимой информации определяются тем, в какой мере пользователь владеет изучаемой темой. Если имеется только общее представление о теме, то естественно начинать поиск с тематических каталогов. Здесь особых проблем не возникает. Пользователь обязательно найдёт ссылки на информационные базы данных или отдельные документы.**

**В тех случаях, когда пользователь владеет терминологией и ищет оригинальные, специфические документы, удобнее проводить поиск с помощью поисковых систем общего назначения.**

# *Требования к поисковым системам общего назначения.*

**С точки зрения пользователя общение с поисковыми машинами должно быть простым, не требующим подготовки и специальных навыков. Пользователь должен иметь возможность:**

- Сделать запрос на родном языке в естественной форме.**
- Ограничить область поиска документов. Это ограничение может определяться местом и временем создания документа, языком, на котором он написан, тематикой разыскиваемой информации.**
- В ответ на запрос получить перечень найденных документов с кратким их описанием. Документы должны максимально соответствовать запросу.**
- При необходимости загрузить нужные документы на свой компьютер для дальнейшего использования.**

**Действия поисковой машины, скрытые от пользователя, должны сводиться к следующему:**

- **Перевод ключевых слов на заданный язык.**
- **Подбор синонимов к ключевым словам.**
- **Учёт морфологических вариантов слов. Например, искать документы, содержащие не только слово «автомобиль», но и «автомобили», «автомобилей» и т.д.**
- **Проведение поиска на указанных пользователем языках.**
- **Смысловой отбор найденных документов по соответствию их тематике запроса.**
- **Представление найденных документов в удобном и понятном для пользователя виде.**

**Каждая из поисковых систем предлагает два способа поиска –  
**простой** и **сложный**.**



# Простой поиск

Он заключается в том, что запрос состоит из набора ключевых слов, а поисковая машина возвращает документы, содержащие хотя бы одно из этих слов. Понятно, что при таком поиске будет найдено множество лишних документов, не представляющих интерес для пользователя.

# Сложный поиск

Для повышения эффективности поиска необходимо использовать развитые средства поиска, позволяющие формулировать точные **сложные** запросы. Правила построения сложных запросов можно найти в справочных файлах каждой поисковой системы.

# Основные принципы систематического поиска

## Поиск всех форм слова.

Для поиска всех слов, имеющих общий корень, используется знак \* .  
Например, в результате поиска по запросу «**компьютер\***» будут найдены документы, содержащие слова: *компьютер, компьютеры, компьютером, компьютерные* и т.д. Ряд поисковых систем делают это автоматически, разыскивая документы, содержащие все морфологические формы ключевого слова.

## Поиск фразы.

Для поиска в документе фразы или словосочетания используются двойные кавычки. По запросу «**информационные технологии**» будут найдены документы, содержащие это словосочетание без вариантов.



# *Включение и исключение ключевых слов.*



Знак плюс, указанный перед словом или фразой без пробела, требует включения этого слова ( фразы ) в документ, а знак минус – его исключения. Это означает, что по запросу **«Поиск в Интернете – «поисковые машины»** из списка документов, найденных по запросу «Поиск в Интернете» будут исключены документы, содержащие фразу «поисковые машины».

Или на запрос **«Информатика»+«логические схемы»** в список документов, найденных по запросу «Информатика» включатся документы, содержащие фразу «логические схемы».



# Логические запросы.

В таких запросах применяются логические операторы **OR**, **AND**, **NOT** и **NEAR**, связывающие несколько слов или фраз в один запрос.

## **OR ( или )**

Когда элементы запроса связаны оператором **OR**, это означает , что в любом из найденных документов должен присутствовать хотя бы один из них. Пример: *сыщик OR детектив* . По этому запросу будут найдены документы, содержащие как оба слова, так и каждое из них в отдельности.

## **AND ( и )**

Этот логический оператор употребляется в том случае, когда необходимо присутствие всех соединённых им слов. Пример: *Информационные AND технологии*. В найденных документах будут присутствовать оба слова. Но это не означает, что они находятся в тексте близко друг от друга или связаны по смыслу.







## **NOT ( не )**

Слова или фразы, перед которыми стоит этот оператор, будут исключены из результатов поиска. Пример: *поиск в Интернете NOT поисковые машины.*

Такой запрос аналогичен запросу: *+ поиск в Интернете – поисковые машины.*

## **NEAR**

Оператор указывает на то, что два слова должны быть разделены не более чем десятью словами. В ряде поисковых систем для запроса на поиск слов, расположенных на определённом расстоянии, применяются фигурные и квадратные скобки.



# Вывод результатов поиска.

Каждая поисковая система по-разному представляет пользователю результаты поиска. Обычно это **нумерованный список** найденных по запросу документов. Для каждого документа, как правило, сообщается его название, URL, размер, время обновления документа и несколько первых фраз текста, по которым можно судить о его содержании. Ряд поисковых систем дают краткое резюме, из которого можно получить представление о том, в каком контексте употребляется ключевое слово.

Название документа и его URL оформляются **в виде гиперссылки**, по которой пользователь имеет возможность загрузить документ на свой компьютер для прочтения, сохранения и печати.

Поисковая система оценивает результаты поиска и выводит список документов в определённом порядке ( **ранжирует документы** ). Таким образом, в начале списка будут находиться документы, имеющие более высокую степень соответствия запросу.

Каждая поисковая система использует собственные критерии ранжирования документов как при анализе результатов поиска, так и при формировании индекса. Вследствие этого результат поиска по одинаковому запросу будет различным для каждого поискового инструмента.

Пользователь может сам повлиять на порядок ранжирования документов, указав в соответствующем окне наиболее значимые слова и фразы. В этом случае ранжирование будет производиться таким образом, чтобы в начале списка оказались документы, содержащие эти слова и фразы.



# ВОПРОСЫ:

1. В каких случаях целесообразно использовать тематические каталоги для поиска информации в сети Интернет?
2. Какие возможности поиска должна предоставлять пользователю поисковая система общего назначения?
3. Составьте запрос на поиск документов, содержащих слово *фрукты* и одно из слов *свежие* или *экзотические*.

Источник: «Информатика». Учебное пособие для 11-го класса с углублённым изучением информатики/А.И.Павловский, А.Е.Пупцев и др.