

Лекция 5

Компьютерные сети и Интернет

Поиск информации в Интернете

Языки запросов поисковых систем

В поисковых системах можно использовать *языки запросов*, позволяющие решать сложные поисковые задачи. Наиболее популярные ИПС:

Яндекс

Google

Rambler

Язык запросов ИПС Яндекс

Запрос содержит ключевые слова и операторы.

Для запросов из нескольких слов, Яндекс умеет определять, на каком максимальном расстоянии должны быть слова запроса. Требования к расстоянию могут быть указаны в явном виде.

Если запрос содержит несколько слов, Яндекс пытается определить, насколько тесно они между собой связаны:

- слова, которые сильно связаны между собой, ищутся в пределах одного предложения;
- менее тесно связанные слова ищутся на расстоянии в несколько предложений;
- несвязанные слова ищутся в одном документе.

Близость слов определяется по факту их синтаксической согласованности:

ловля рыбы спиннингом

Пример синтаксической несогласованности:

ловля рыбий спиннинговый

Требования к расстоянию могут быть заданы в явном виде:

Слова идут подряд

Такой порядок слов можно указать с помощью запроса в кавычках

“параллельные вычисления”

По этому запросу будут найдены страницы, где вначале идет слово *параллельные*, а потом *вычисления*

Слова в одном предложении

Ограничить поиск страницами, где слова запроса находятся в пределах предложения, можно, соединив слова оператором **&**:

лето & зима

Слова в одном документе

При поиске документов, где присутствуют заданные слова — неважно, на каком расстоянии и в каком порядке, используется оператор **&&**:

свет && тьма

Слова на расстоянии в несколько слов

Максимально допустимое расстояние между двумя любыми словами запроса можно указать, поставив после первого слова символ /, сразу за которым идет число, задающее расстояние

свет /2 тьма

Слова на расстоянии в несколько предложений

Расстояние между словами можно задавать в предложениях. Для этого перед оператором расстояния / указывается оператор **&&**:

свет &&/2 тьма

Порядок слов

Кроме расстояния между словами можно указывать их порядок.

Слова следуют в определенном порядке на нужном расстоянии

Если точно известно не только расстояние, но и порядок слов запроса, в котором они должны идти в тексте страниц, можно указать между символом / и числом-расстоянием символ + для прямого порядка слов или - для обратного

свет /+2 тьма

Слова расположены в заданной окрестности

В некоторых поисковых задачах слова могут идти в разном порядке, между ними может быть от n слов (минимум) до m (максимум). В таком запросе используется оператор /(n m):

свет /(-10,+15) тьма

Поиск любого из слов

При поиске любого из слов используется оператор |:

добро | зло

Исключение слов из поиска

Яндекс позволяет исключать страницы, где есть определенные слова. Для этого используется оператор ~~, слева от которого указывают, что искать, а справа - какие страницы исключать из поиска:

Чирцов ~~ декан

Исключение слов из поиска

Иногда допустимо, чтобы слово встречалось на странице, но нельзя, чтобы оно встретилось в том же предложении. С этой целью используется оператор ~:

Кузькина ~~ мать

Сложные запросы

В сложных запросах можно использовать скобки:

чирцов && (лыжи | парашют | дайвинг) ~~ (декан | ректор)

Морфология

Слова, указанные в запросе, по умолчанию ищутся с учетом морфологии. Для отключения этого учета перед словом используется оператор !:

!чирцов && (лыжи | парашют | дайвинг) ~~ (декан | ректор)

Другие возможности языка запросов

Поиск текста в заголовке страниц:

```
$title(запрос)
```

Ограничение поиска группой страниц с заданным URL:

```
url="www.gde-to.ru/chto-to/*"
```

Ограничение поиска заданным типом файла:

```
mime="HTML|PDF|DOC|PPT|..."
```

Ограничение поиска заданным языком:

```
lang="ru|en|de|fr|..."
```

Поиск только по страницам, дата которых удовлетворяет заданному условию:

```
date="ГГГГ{*|ММ{*|ДД} }"
```

Архитектура ЭВМ

Мейнфрейм
 Персональный компьютер
 Бортовой компьютер
 Наладонный компьютер

Несмотря на различный внешний облик, во внутреннем устройстве современных компьютеров разных типов много общего.

Идеи фон Неймана, высказанные им в статье, опубликованной в 1946 году совместно с Артуром Берксом и Германом Голдстайном «Предварительное обсуждение логического устройства электронного вычислителя» («Preliminary Discussion of the Logical Design of an Electronic Computing Instrument»), до сих пор лежат в основе устройства современных компьютеров

Архитектурой компьютера называют описание основных компонентов компьютера и способов их взаимодействия. Архитектура — это те свойства и характеристики вычислительной системы, которые «видны» программисту. Среди них — набор команд, разрядность машинного слова, механизмы ввода-вывода и т. д.

Фон Нейман считал, что важнейшим качеством универсального компьютера является его способность хранить программу, связанные с ней команды, данные, промежуточные результаты вычислений и другую информацию. Для этого универсальный компьютер должен иметь устройство хранения команд и данных, которые изменяются от программы к программе или даже во время выполнения одной программы. Таким устройством является *оперативная память*.

И данные, и команды могут одновременно храниться в одной памяти. Это положение считают одним из основных принципов фон-неймановской архитектуры.

