

# Поиск информации в ИНТЕРНЕТЕ

Для слушателей курсов.

Халкечева Л.В.

A stylized, layered mountain range graphic in shades of teal and blue, located in the bottom right corner of the slide.

**Цель занятия:** ознакомиться с различными типами поисковых систем, научиться находить необходимую информацию в Интернете.

- ◆ **ИПС- информационно-поисковые системы.**

Задавая вопрос в Интернете **ПОМНИТЕ** - это вопрос не человеку , а машине!

Машина анализирует информацию в зависимости от заложенных в неё алгоритмов(программ). Эти программы, обрабатывающие запросы называются ИПС. Существует два типа ИПС: словарные и классификационные.

# Классификационные поисковые системы.


- ◆ Это, по сути, электронные аналоги библиотечных каталогов.

Обычно это иерархическое гипертекстовое меню с пунктами и подпунктами, определяющими тематику сайтов, адреса которых содержатся в данном каталоге, с постоянным уточнением темы:

Медицина и здоровье

Альтернативная медицина

Лекарственные препараты...

- ◆ Главное отличие классификационный ИПС – то , что представленные в них данные формируются людьми, осуществляющими поддержку данного сервера.
  - ◆ Основным преимуществом такого средства навигации является чёткое соответствие содержимого сайтов объявленной тематике.
  - ◆ Недостатком является сравнительно небольшой охват существующих ресурсов сети.
- 

# В русскоязычном Интернете примерами классификационных ИПС являются серверы:

- ◆ «АУ» - [www.au.ru](http://www.au.ru)
- ◆ -www.stars.ru

- ◆ На классификационных поисковых системах кроме названия рубрик обычно имеется строка **Поиск**, позволяющая ввести ключевое слово. Тогда машина, обработав все иерархическое дерево, выдаст в ответ ссылки на соответствующие этому ключевому слову web-документы.

# Словарные (автоматические) поисковые системы

- ◆ В отличие от рассмотренных выше, основой автоматической ИПС является размещенная на поисковом сервере база данных, содержащая в себе адреса сайтов, перечень соответствующих размещенным на них WWW-страницам ключевых слов, копий HTML-текстов этих страниц (в ряде ИПС), а также систему управления базой данных, действующую по технологии «клиент — сервер».

- ◆ Посетителю такого поискового сервера предоставляется форма для ввода ключевого слова (слов) или фразы. После щелчка мышью на расположенной рядом с полем ввода кнопке **Поиск** введенная пользователем («клиентом») ключевая фраза пересылается на сервер, обрабатывается, из имеющейся БД извлекаются адреса, соответствующие запросу, и список этих адресов (сгенерированный в виде HTML-документа, строки которого являются гипертекстовыми ссылками на сайты) пересылаются пользователю в качестве результата поиска.



- ◆ При поиске в Интернете важны две составляющие: полнота поиска (т. е. стремление не упустить какую-либо важную информацию) и точность (т. е. отсутствие в результатах поиска лишней информации). Обычно обе эти составляющие называют общим словом *релевантность*, подразумевая под ним максимальное смысловое соответствие результатов поиска указанному запросу. Релевантность также можно рассматривать и как способ сортировки найденных по запросу документов: чем больше документ соответствует запросу, тем ближе к началу списка должен находиться его адрес.

- ◆ Для достижения этого результата поисковая система должна учитывать такие параметры, как количество найденных в тексте ключевых слов, «контрастность» слова (его относительную частоту встречаемости для данного документа), расстояние между словами, положение слова в документе и его подразделах и т. п.
- ◆ Однако не следует забывать, что релевантность — качественная мера. Соответствие ответа ИПС интересам пользователя зависит не только от свойств данной поисковой системы, но и от того, насколько удачно сформулирован запрос, т. е. насколько правильно последний отражает интересы пользователя.

- ◆ Другая особенность работы словарной ИПС касается морфологического анализа обрабатываемых текстов WWW-страниц и ключевых слов/фраз. Это связано с тем, что современные поисковые системы все дальше уходят от устаревшего принципа конструирования ключевых фраз как некоей совокупности связанных операциями булевой логики «шаблонных» слов. Напротив, сегодняшние ИПС способны работать с естественно-языковыми запросами, когда вместо ключевой фразы, по сути, вводится свободное описание желаемой области поиска.

# Поисковые серверы специального назначения.

- ◆ Кроме обычных серверов, ориентированных на поиск текстовых документов (WWW- страниц), в Интернете имеется также ряд серверов, предназначенных для поиска файлов, доступных по протоколу FTP (поиск по именам файлов), подробной информацией о владельцах указанных адресов e-mail и т.п. Некоторые реализованы по технологии WWW, другие представляют собой обособленные виды доступа к сетевым ресурсам (Finger, WHOIS), но доступны при помощи обычного WWW-браузера.

# Метапоисковые серверы.

- ◆ Они представляют собой службы, обеспечивающие не собственно поиск информации по ключевому слову, а только перевод ключевой фразы на другие языки, выделение возможных смысловых форм входящих в неё слов.