

ПРОГРАМИРАНЕ И ИЗПОЛЗВАНЕ НА КОМПЮТРИ 3

доц. д-р Дора Димитрова

катедра ПКТ, ФКСУ

каб. 2300А

тел. 965 2192

dimitrova@tu-sofia.bg

pct.tu-sofia.bg/dd/pik3



● Определение

- приложение на **компютри и компютърни програми** за събиране, съхраняване, обработване, предаване и извличане на **информация**

● Множество аспекти

- инсталиране
- проектиране
- програмиране
- администриране
- управление
- ...



● Технологична среда

- изчислителна мощност на компютрите
- мултимедии
- виртуална реалност
- Интернет – информационна вселена без граници

● Социална среда

- разширяване на приложните области
- повишена технологична култура на потребителите
- повишени изисквания към функционалността на системите



Съвременни приложения

Компютърна
обработка и
управление на
документи

MS Word & html

Анализ и
визуализация на
данни и
информация

MS Excel

Моделиране и
симулация на
системи и
процеси

MatLab

Информационни
системи и
приложения

MS Access & MS SQL Server

Разпространение
на информация
в мрежова среда

Internet

Потребителски
интерфейси



- Представяне на документи в електронна среда
 - съдържание
 - форма
- Обработка на електронни документи
 - на съдържанието
 - на формата
- Представяне на електронни документи в Интернет
- Търсене и извличане на информация от електронни документи в Интернет



Представяне на документи в електронна среда

● Съдържание на документ

- компоненти
- структура

● Форма на документа

- параметри на формата



Компоненти на съдържанието

Статични

текст

линеен

хипертекст

графика

векторна

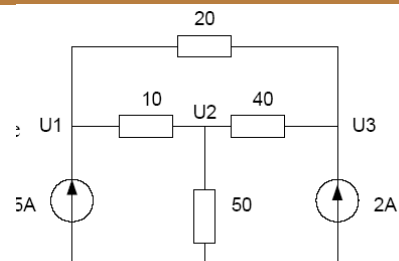
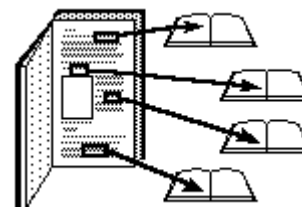
растерна

Динамични

аудио

анимация

видео

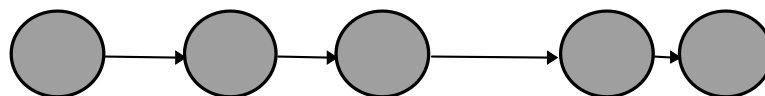


Документ

страници

параграфи

СИМВОЛИ

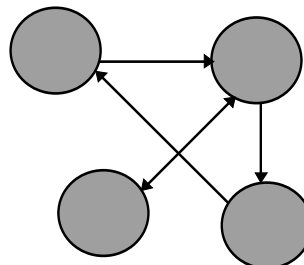
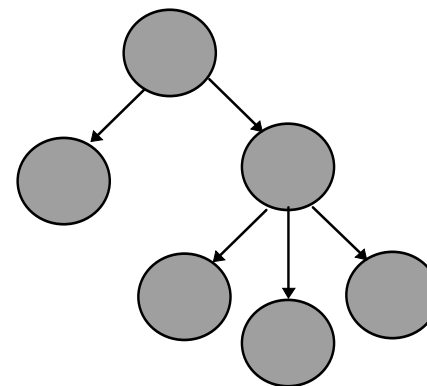


Структури

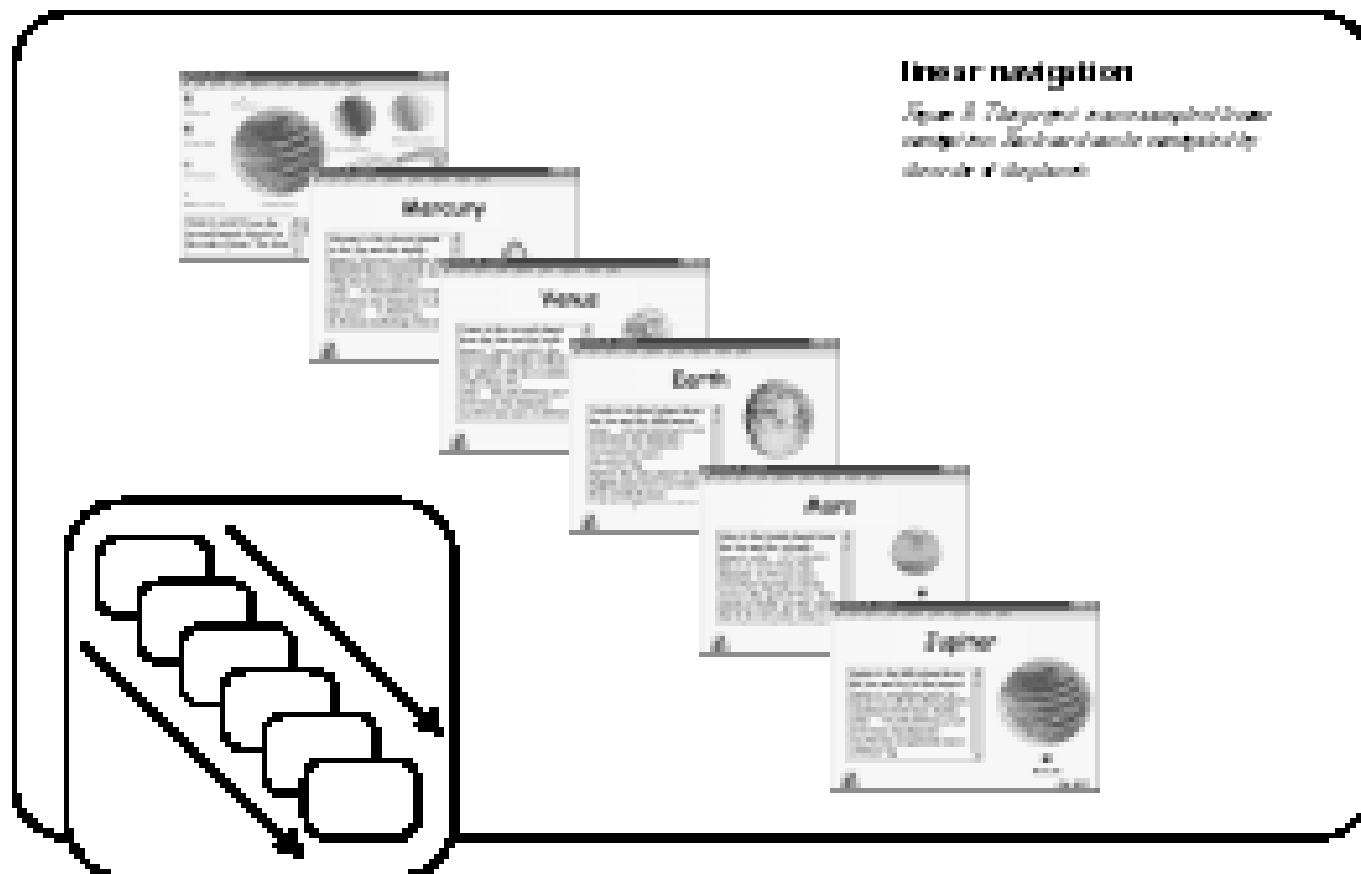
линейна

йерархична

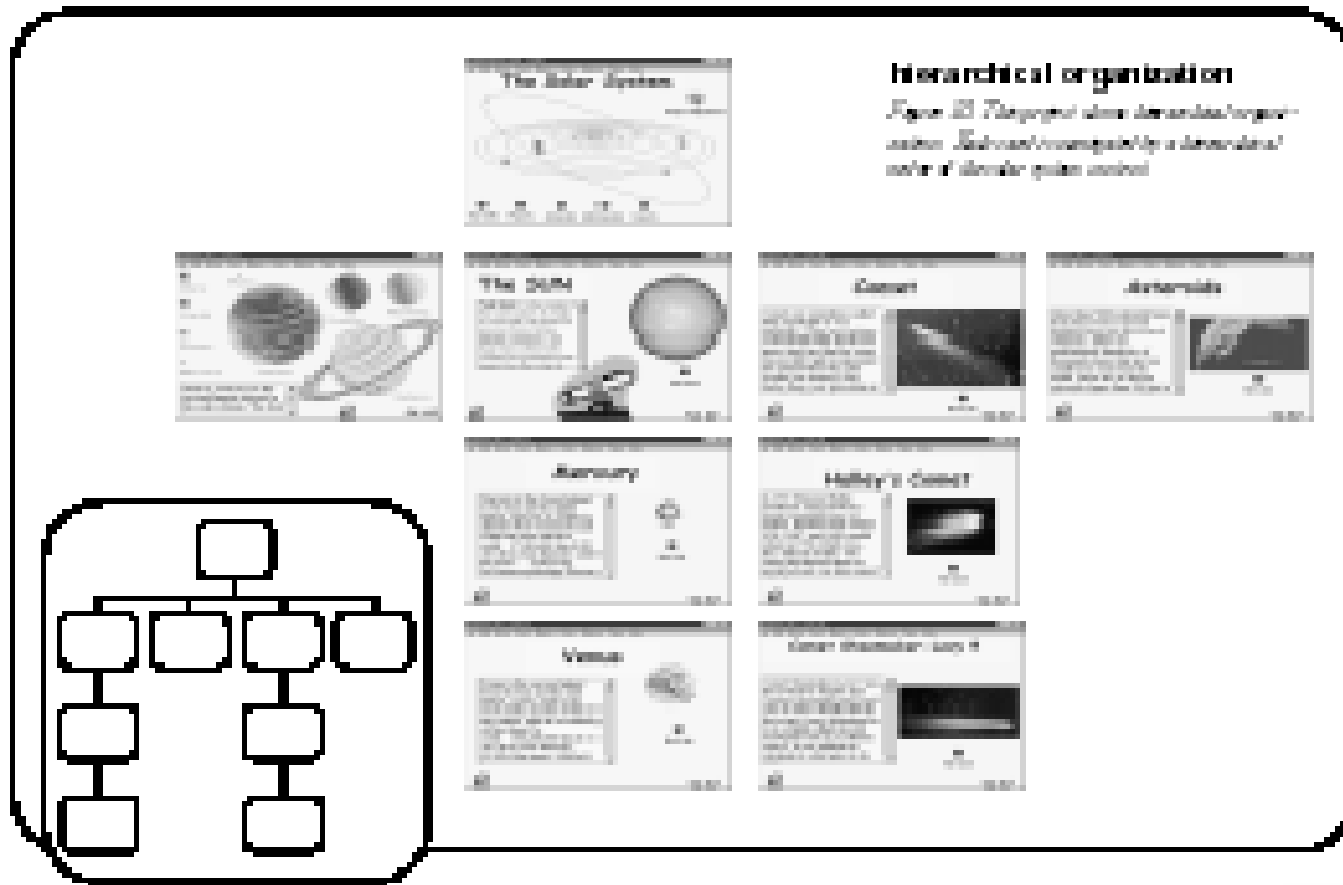
мрежова



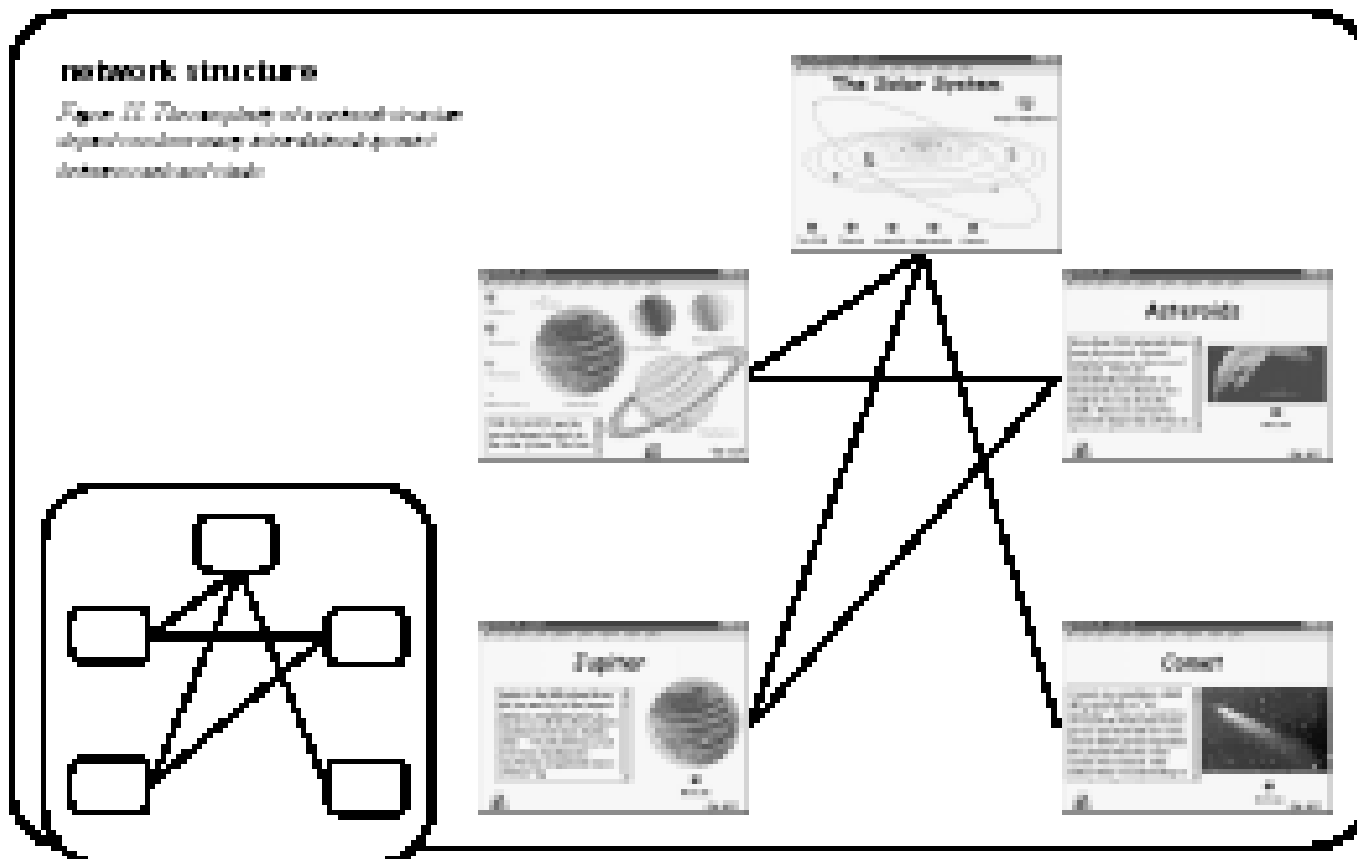
Линейна структура



Йерархична структура



Мрежова структура



Обработка на компонентите

ТЕКСТ

- редактиране
- форматиране
- търсене
- сортиране
- шифроване
- лингвистична обработка

графика

- Редактиране
- Оцветяване
- Осветяване
- Операции с пикселите
- Филтриране
- Композирание
- Зрителни ефекти
- Геометрични трансформации
- Координатни трансформации

на съдържанието

аудио

- Редактиране
- Търсене
- Филтриране
- Конвертиране
- Звукови ефекти

ВИДЕО

- Редактиране
- Трансформации
- Конвертиране

анимация

- Създаване
- Възпроизвеждане



Форма на документ

● Визуализация

● Параметри

- количествени
- пространствени
- светлинни
- времеви

● Форматиране – задаване на стойности на параметрите за обектите на структурата

- символи
- параграфи
- страници
- документ



Представяне на електронни документи в Интернет

● Видове документи

- статични
- динамични
- активни

● Представяне на структурата

- XML

● Представяне на формата

- HTML, XHTML



- Специализиран програмен език за описание на документи от WWW
- Задава инструкции на клиентската програма за визуализиране на документа
- Основан на стандарта **SGML**, *Standard General Mark-up Language*
- Правилата на езика се задават в спецификация **DTD**, *Document Type Definition*



Представяне на документ чрез HTML

● Текстов файл, който се съхранява на сървера

- създава се с текстов редактор или HTML редактор
- ASCII файл с разширение към името .htm или .html

● Към основните елементи от съдържанието на документа се добавят команди, опашки (*tags*)

- описват (специфицират) форматирането
- използват се по двойки – отваряща и затваряща

● Формат на файла

- `<head>`
 - данни за документа
- `<body>`
 - съдържанието на документа



Формат на файла

```
<html>  
  <head>  
    <title>Window Title</title>  
    данни за документа  
  </head>  
  
  <body>  
    съдържание на документа  
  </body>  
</html>
```



Типове команди

● За форматиране

- документ
- параграф
- символи

● За опериране

- хипервръзки
- вграждане на обекти
- контроли

● Формат на командите

`<start_tag>`Обект, който ще се появи на екрана`</end_tag>`

`<start_tag param_name="name",param_value="value"> ... </end_tag>`



Пример

```
<html>
  <head>
    <title>Project Web Home</title>
  </head>
  <body background="images/backgrnd.gif">
    <hr>
    <p>
      <font size="4" face="Ariel"> Welcome to my new</font>
      <font color="#004080" size="4" > home page</font>
    </p>
    <br>
    <p>
      <font size="2" face="Ariel">It is created by
      <font color="#800080">me </font>
    </font>
    </p>
    <hr>
  </body>
</html>
```



Команди на HTML

<i>команда</i>	<i>бележка</i>	<i>пример</i>
<P>	Normal paragraph style	
 	Line break	 - Blank Line
<H1>	Heading 1 style	
<H2>	Heading 2	
<H3>	Heading 3	
<H4>	Heading 4	
<H5>	Heading 5	
<HR>		to separate parts of a document
	unordered bulleted lists The style: •Item one •Item two •Item three	 Item one <L2> Item two <L3> Item three
	ordered (numbered) lists The style: 1.Item one 2.Item two 3.Item three	



Команди на HTML

	<i>форматиране на символи</i>	
	Emphasis style	<i>italic</i>
	Strong style	bold
<U>	Underline style	<u>underline</u>
<STRIKE>	StrikeThrough	strike through
<PRE>	PreFormatted style R1C1 R1C2 R1C3R2C1 R2C2 R2C3	to display a block of text in a fixed-width font
<TT>	Typewriter style	fixed-width typewriter font
<SAMP>	Sample style	monospaced style
<ADDRESS>	<i>Address style Box 211 1212 Third Ave. Someplace XX</i>	displays text in an italic format.
<DL>	<i>Definition List</i> First Term Term Definition Second Term Term Definition	display a term and its definition



Команди на HTML

<APPLET>	CODE , size, CODEBASE,	

<A	anchor, specify a hypertext reference	HREF= http://alice.bg
<FORM>	параметри на заявката, които ще въведе потребителя	
<PARAM	TYPE = "... " NAME = "... " VALUE = "... "	>
<Script>	JavaScript	



CSS *Cascading Style Sheet*

- Правила за прилагане на **СТИЛ** към документ
- Използва се при форматиране на HTML и XML документи
- CSS оператори, които идентифицират порции от документа, който се форматира в определен стил
- Установяват връзка между HTML или XML с набор от стилове
- Два начина на прилагане
 - document-level – вътре в документа
 - external – в отделен текстов файл **xx.css**
- Реферира се в HTML и XML чрез инструкции



Видове CSS

● Три вида

- Element type
- Attribute class
- Attribute ID

● Примери

```
p {
  display: block;
  margin-bottom: 10px;
  font-family: Times, serif;
  font-size: 12pt;
}
p.empha {
  display: block;
  margin-bottom:
10px;
  color: green;
  font-family: Times,
serif;
  font-size: 12pt;
  font-style: italic;
}
```

```
#feelingblue
{
  color: blue;
  font-family: Ariel;
}

<h1 id="feelingblue"> My Blues
Collection Comes Here
</h1>
```



XML *eXtensible Markup Language*

- Мета-език за описание на структурата на документ
- Няма фиксирани опашки, позволява на потребителите да дефинират свои
- Позволява дефиниране на нови езици, специфични за конкретни типове приложения, технологии или бизнес области
- Позволява интегриране на данни от различни източници
- Позволява дефиниране на различни изгледи на един и същ документ
- Формата на изобразяване на документа се дефинира в отделни файлове *XSL eXtensible Stylesheet Language*



Спецификация на XML

- Прави се спецификация на езика, която се използва за контрол и валидиране
- Две технологии:
 - чрез Document Type Definition (DTD);
 - чрез XML Schema Definition (XSD).
- Дефинират се
 - структурата на документа
 - допустимите стойности на атрибутите на обектите



- XML документи се формират чрез описание на XSL
- XSL е друг дескриптивен език
- Състои се от две части
 - XSL Transformation (XSLT)
 - XSL Formatting
- XSLT дефинира преобразования от един XML код в друг XML или в HTML или в друг дескриптивен език. По този начин се създават условия за визуализиране на документа в различни среди, за различни потребители, задачи и платформи
- XSL Formatting предоставя типографски елементи за визуализация



● Мотивация за развитие

■ Internet

- универсално хранилище на човешкото знание и култура
- безплатен източник на информация

● Проблеми

- нерегламентирани заявки за извличане
- разнородни информационни източници
- огромно количество
- продължително търсене



● Причини за проблемите

- липса на единен модел на данните, съдържащи се в Интернет
- липса на качествени описания и структури
- изоставане на софтуерните технологии в сравнение с потребителските фактори

● Съвременни изследвания

- моделиране на данни
- методи за филтриране, класификация и категоризация на документи
- потребителски интерфейси
- езици за формиране на заявки
- визуализация на информацията
- системни архитектури
- и др.



Извличане на данни и извличане на информация

● DR Data Retrieval

- намиране на документи, съдържащи определени ключови думи
- добре дефинирана семантика
- нетърпимост към грешки

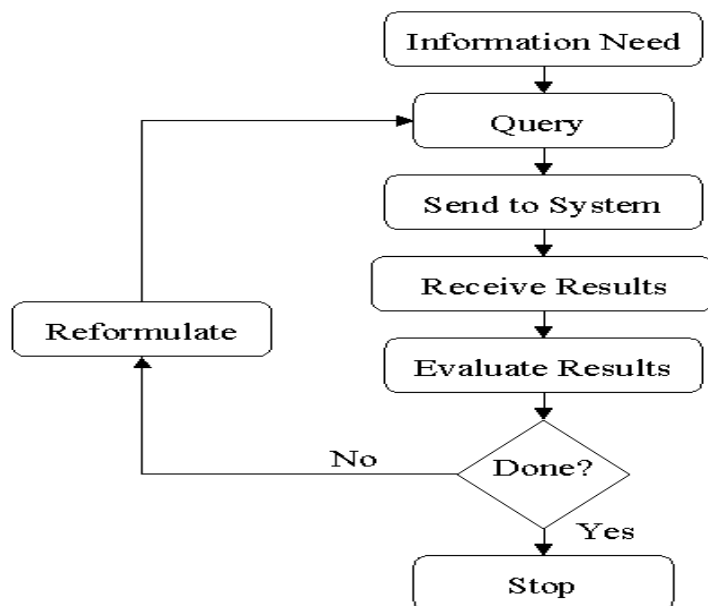
● IR Information Retrieval

- информация относно тема
- недефинирана семантика
- толеранс на грешките

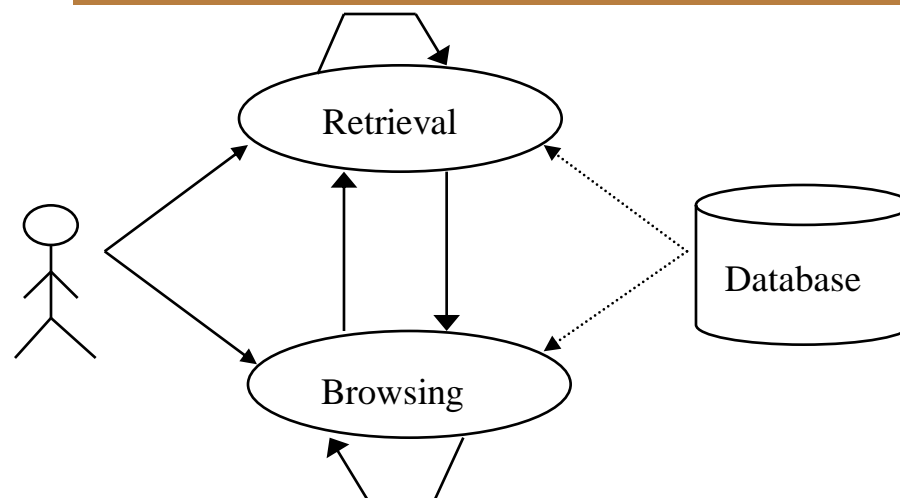
● Информационна система IR

- интерпретира съдържанието на намерените документи
- ранжира намереното по отношение на потребителски критерии
- оценява приложимостта на намерените документи





Модели на търсене



Прост интерактивен модел

- дълги списъци от намерени документи без оценка на приложимостта
- статична цел на потребителя
- итеративно-настройващи се заявки
- използва се от web search machines

Самообучение на потребителите

- използване на хипервръзки и навигация в търсенето
- 'berry-picking' model
 - актуализация на целите на потребителя
 - обобщаване на резултати от множество под-заявки
 - главен резултат: акумулираното знание, получено по време на търсенето, не в крайната извадка



Видове търсене

● разглеждане

- ненасочено изследване на информационни структури
- следва се от селекция

● запитване

- получаване на ново (под) множество от информационни елементи, не събирани заедно досега
- следва се от разглеждане

● селектиране

- избор от организирана информация
- използва се като резултат или за формулиране на запитване

● сканиране

- целенасочено изследване на заглавия, термини, категории и др.

● навигация

- сканиране + селектиране



Формулиране на заявки

● Методи

- Избор на ресурсна колекция, метаданни или информационно множество, отговарящо на критериите за приложимост
- Специфициране на думи, фрази, дескриптори за сравнение

● Средства

- команден език
- попълване на формуляри
- избор от меню
- директна манипулация
- естествен език



● Дума и фраза

- дизюнкция или конюнкция ?
- приложимостта на резултата зависи от алгоритъма на ранжиране
 - статистически
 - тегло
 - вероятност
 - процент на срещане на думата

● Естествен език

- потребителят контролира приложимостта
- необходимо е да познава разпознаваемите команди



● Въпроси и отговори

- от синтаксиса на въпроса се извлича информация за свойствата на отговора (напр. каква част от изречението е)

● FAQ finder

- установяване на съответствие между двойки въпрос и отговор

● Предефинирани типове въпроси

- идентифициране на типа въпрос на потребителя и последващо перефразиране (доуточняване)

● Свободен текст на въпросите

- алгоритъм за разделяне на използваните думи и оценяване на тяхното значение



● Digital Libraries

- Многоезични документи
 - речници
- Мултимедийни документи
 - синхронизация на потоци информация
 - намиране по метаданни
 - визуални езици за заявки
 - *Query By Image Content*
- Структурирани документи
 - приложени една или повече структури на данните
 - описания чрез SGML
- Разпределени документи
 - във физическо или логическо пространство



Машини за търсене

● Действия

- проучване на мрежата
- индексирание на документи
- обработка на заявки
- класиране на резултати

● Оптимизация на търсенето *SEO Search Engine Optimization*

- *white hat* - позволени средства за оптимизация, напр.
 - коректно описание в метаданни
 - ключови думи
 - повече връзки към документа
 - улеснения за търсещите програми
 - карта на документа
 - малка дълбочина на йерархия
 - текст вместо/заедно с графика
 - статични вместо динамични страници
 - без фрейми
- *black hat* – непозволени средства, напр. некоректни мета данни, скрити съобщения, огледални сайтове и др.

