

Конспект лекции отправлять на почту: tankae@inbox.ru до 21:00

«ТЕХНОЛОГИИ ПОИСКА ТЕМАТИЧЕСКОЙ (ПРОФЕССИОНАЛЬНОЙ) ИНФОРМАЦИИ В СЕТИ INTERNET»

План лекции:

- 1. Написать что такое ИПС**
- 2. Классификация поисковых средств**
- 3. Указать виды операторов и для осуществления поиска информации**

ПРИНЦИПЫ КОНТЕКСТНОГО ПОИСКА

Контекстный поиск - запрос информации в компьютерной системе с использованием набора ключевых слов.

ВИДЫ ПОИСКОВЫХ РЕСУРСОВ

- **Информационные порталы;**
- **Каталоги Интернет-ресурсов;**
- **Поисковые системы**

Информационно- поисковые системы



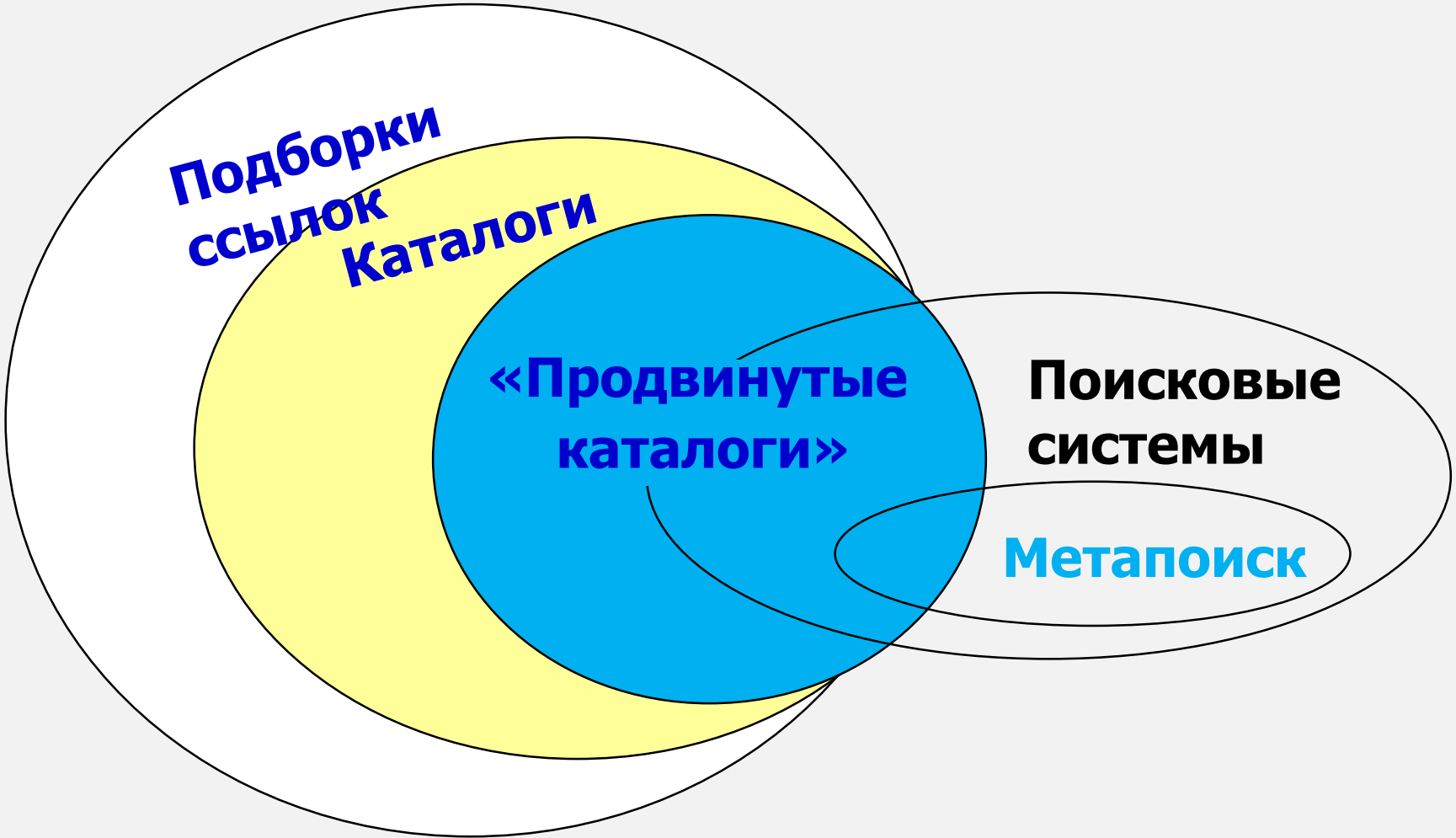
ЧТО ЭТО ТАКОЕ?

(ИПС) — это система, обеспечивающая поиск и отбор необходимых данных в специальной базе с описаниями источников информации (индексе) на основе информационно-поискового языка и соответствующих правил поиска.

Главной задачей любой ИПС является поиск информации релевантной (полной) информационным потребностям пользователя. Очень важно в результате проведенного поиска ничего не потерять, то есть найти все документы, относящиеся к запросу, и не найти ничего лишнего. Поэтому вводится качественная характеристика процедуры поиска — релевантность.



КЛАССИФИКАЦИЯ ПОИСКОВЫХ СРЕДСТВ



ПОИСКОВЫЕ СЕРВЕРЫ

Для поиска информации используются **специальные поисковые серверы**, которые содержат более или менее полную и постоянно обновляемую информацию о Web-страницах, файлах и других документах, хранящихся на десятках миллионов серверов Интернет.

Имеющиеся поисковые системы можно условно разделить на три группы:

- 1) справочники Интернет;
- 2) поисковые системы общего назначения;
- 3) специализированные поисковые системы;



- **Справочник Интернет (поисковый каталог)** является аналогом тематического указателя в библиотеке: он предоставляет перечень наиболее важных документов по заданной теме.

Примеры таких систем –
поисковые серверы:

- <http://WWW.yahoo.com>
- <http://narod.yandex.ru/rubrics/>



Наиболее популярные технологии поиска информации в Интернет

ПОИСКОВЫЕ МЕХАНИЗМЫ

1. **Поисковые машины (searchengines)**

Эти средства поиска в ответ на запрос выдают список страниц, удовлетворяющих заданным критериям.

Примеры поисковых машин:

- **Яндекс (<http://www.yandex.ru>)**
Rambler (<http://www.rambler.ru>)
Google (<http://www.google.com>)
<http://www.google.com.ru>
<http://www.google.ru>

ПОИСКОВЫЕ МЕХАНИЗМЫ

2. Каталоги, в которых сайты упорядочены по категориям специально разработанного дерева-рубрикатора.

Примеры каталогов:

- **Yahoo (<http://www.yahoo.com>)**
Русские каталоги
<http://www.au.ru>
<http://www.stars.ru>

ПОИСКОВЫЕ МЕХАНИЗМЫ

3. Тематические подборки ссылок.

Иногда они содержат рубрикатор и могут рассматриваться как частный случай каталога, ограниченного некоторой темой.

- **Например, сайт alledu.ru содержит мощную систему для поиска информации в области образования.**

Яндекс Google  Rambler

YAHOO!  NiGMA.RU
Интеллектуальная поисковая система

sciseek  LYCOS.
Go get it! bing

gogo  Alexa[™]
The Web Information Company Webalta

SCIRUS
for scientific information only twitter
facebook altavista[™]

About.com  YouTube
Local.com[™]  LIVEJOURNAL msn 

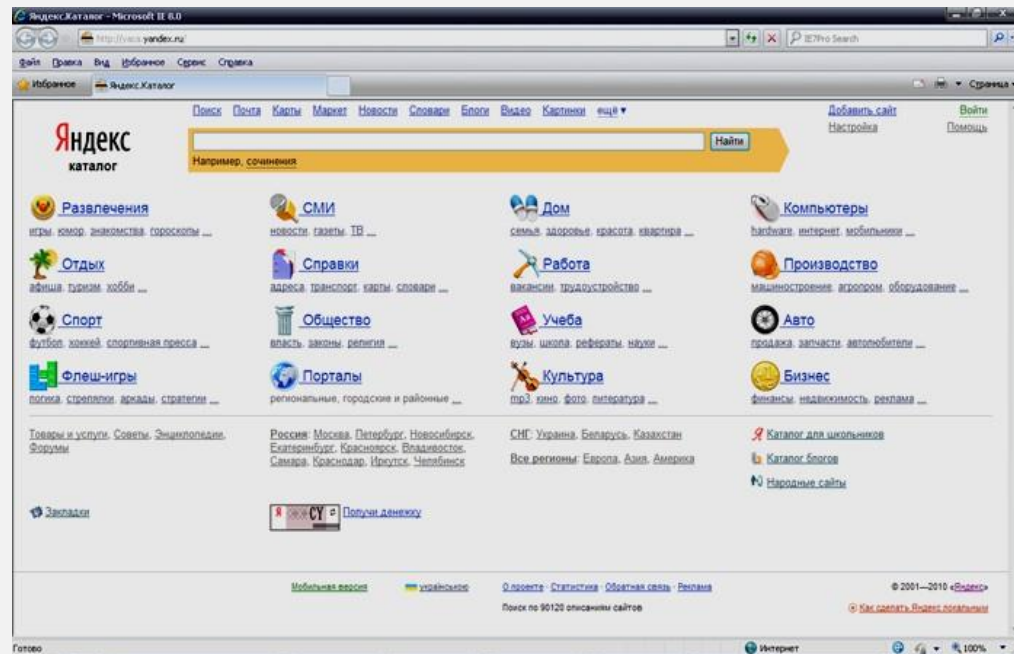
КАК РАБОТАЕТ ПОИСКОВАЯ СИСТЕМА?

- **Первый этап работы поисковой системы** – это индексирование информации, находящейся в Internet. Сетевой робот поисковой системы просматривает огромное количество страниц и заносит адреса и краткое содержание этих страниц к себе в базу данных, точнее в поисковый индекс

КАК РАБОТАЕТ ПОИСКОВАЯ СИСТЕМА?

- **Второй этап работы поисковой системы** – это выдача накопленной и отобранной, отсортированной и классифицированной информации по запросам пользователей.

ЯНДЕКС



ВИДЫ ОПЕРАТОРОВ ДЛЯ ФОРМИРОВАНИЯ ПОИСКОВЫХ ЗАПРОСОВ ДЛЯ YANDEX

«...» - кавычки

Поиск документов, содержащих слова запроса в заданной последовательности и форме.

«Лечебная физкультура»

Будут найдены документы, содержащие данное словосочетание.

- ! – восклицательный знак
- Из списка будут исключены слова других морфологических форм (употребляемые в других падежах, числе и тд)
- !диспансеризация
- Будут найдены документы со словом «диспансеризация» в заданной форме.

| – вертикальная черта

Поиск всех синонимов

Дети|младенец|ребенок

Будут найдены документы, в которых присутствует хотя бы одно из слов запроса: «дети», «младенец» или «ребенок».

- + – знак «плюс»
- Поиск документов, в которых обязательно присутствует выделенное слово.
- **+лечебная гимнастика+физкультура**
- Будут найдены документы, в которых обязательно содержатся слова «лечебная гимнастика» и «физкультура»

- – знак «минус»

Исключение слова из результата поиска. Исключается только слово, перед которым стоит оператор.

+анальгин +(инструкция по применению) – аптека - цена

Будут найдены документы только с инструкцией по применению без цен и предложений аптек

- * – знак «звездочка»
- Поиск по цитате с пропущенным словом (словами). Один оператор * соответствует одному пропущенному слову.
- Работа не * , в лес не *
- Будут найдены документы, содержащие данную цитату, включая пропущенные слова.

~ – знак «тильда»

Поиск документов, в которых заданное слово может присутствовать на странице, но не содержится в одном предложении со словом, указанным до оператора.

Реабилитация ~ инвалидов

Будут найдены документы, в которых слово «реабилитация» может присутствовать, но не в одном предложении со словом «инвалидов».

- ~~ – знак «двойная тильда»
- Поиск документов, в которых отсутствует заданное слово. Исключается только то слово, перед которым стоит оператор.
- **Орёл ~~город**
- Будут найдены документы, в которых нет упоминания о городе-герое

mime:

Поиск по документам в заданном типе файла.

Закон «Об охране здоровья»

mime: doc

Будут найдены только документы в формате .doc, т.е. созданные и сохраненные в программе MS Word 2003

- date:
- Поиск по страницам с ограничением по дате их последнего изменения.
- Диспансеризация date:> 20150810
- Будут найдены документы измененные позднее 10 августа 2015