



Лектор

Ст. преподаватель Купо А.Н.

Internet – среда медиаобразования

История создания и развития Internet



В 1960-х годах начались эксперименты по соединению компьютеров друг с другом с помощью телефонных линий. Финансовая поддержка фонд Агентства Перспективных Проектов Исследований (**ARPA**) Министерства Обороны США. Разработана принципиально новая технология передачи данных, которая использовала разбиение данных на пакеты, каждый из которых имел свой адрес назначения. Эта технология позволяла нескольким пользователям работать по одной и той же линии связи одновременно.

К середине 70-х годов многие правительственные организации, исследовательские лаборатории и университеты были соединены компьютерной сетью, получившей название **ARPAnet**.

Для объединения сетей, имеющих разные интерфейсы, скорости передачи данных и компьютеры различных изготовителей в 1974 году учеными **Винтом Серфом** и **Бобом Кэном** были предложены новые мощные межсетевые протоколы передачи пакетов данных (прообраз современного протокола **TCP/IP**).

История создания и развития Internet



К середине 80-х годов в сети **NSFnet**, принадлежащей Национальному научному фонду США (**NSF**) и поглотившей к тому времени **ARPAnet**, были объединены пять суперкомпьютерных центров

Начиная с 1986 года можно говорить о становлении глобальной компьютерной сети Internet. Слово **Internet**, является аббревиатурой английских слов **interconnected networks** (соединенные сети)

К концу 80-х годов к Internet подключились множество образовательных учреждений, государственные организации, различные американские и иностранные фирмы

Одновременно с этим событиями развивался и проект объединения компьютерных сетей европейских стран, движущей силой которого была Европейская лаборатория физики элементарных частиц (**CERN**) в Женеве, располагавшая самым мощным в Европе сервером Internet. Результатом этих усилий стало воплощение концепции распределенной компьютерной сети, получившей в 1992 году название **World Wide Web (WWW, Всемирная паутина)**, автором которой считается Тим Бернерс-Ли.

Структура Internet



Internet – это совокупность различных компьютерных систем и компьютерных служб.

Электронная почта (E-mail) – система пересылки электронной корреспонденции;

Телеконференции (Usenet) – система обмена информацией с множеством пользователей;

Файловые архивы FTP – компьютеры, хранящие файлы, доступные для пользователей.

Всемирная паутина (WWW) - гипертекстовая информационная система;

Электронная почта



Электронная почта обеспечивает передачу сообщений от одного адресата другому и выполняет функции обычной почты. Скорость доставки электронного сообщения (в зависимости от состояния связи между источником и адресатом) может занимать несколько часов, минут и даже секунд. С помощью электронной почты можно также получить доступ к базам данных и файлам библиотек. В отличие от обыкновенной почты в текст электронного сообщения можно **вставлять графические и звуковые файлы, а также файлы программ**. Адрес электронной почты (**E-mail**) отдельного лица или организации традиционно используется для идентификации владельца. Поэтому адреса E-mail можно использовать для поиска. Эти адреса активно накапливаются в специальных системах поиска людей и организаций. Серьезное неудобство для поиска по E-mail составляет то, что при получении адреса допускается регистрация пользователя под псевдонимом. Эта практика особенно широко распространена на серверах, предоставляющих бесплатные почтовые ящики.

Телеконференции



Телеконференции были разработаны для обмена новостями между компьютерами по всему миру. В отличие от электронной почты, которая обычно является разговором "одного с одним", **Usenet** - это разговор "многих со многими". Система построена по принципу **электронных досок объявлений**, когда пользователь может разместить свою информацию в одной из тематических групп новостей.

Информация передается пользователям, которые **подписаны на данную группу**.

Телеконференции организованы с помощью **иерархии**. Имена групп новостей начинаются с самой общей темы. За этими общими темами следуют более детализированные.

Данный ресурс наиболее значим для **быстрого накопления информации по узкому вопросу**, а при поиске — чаще для получения частной, неофициальной информации.

Файловые архивы FTP



Файловые архивы FTP — это источники программного обеспечения. В них собрано огромное количество информации, накопленной в FTP-архивах за десятилетия эксплуатации компьютерных систем. Доступ, навигация и передача файлов по FTP отличается простотой.

Если электронная почта предоставляет обмен сообщениями между различными компьютерными системами, то серверы **FTP предоставляют обмен отдельными файлами и программами.**

Они позволяют пользователям не только скачивать информацию, но и **добавлять информацию на сервер.**

Всемирная паутина (WWW)



Всемирная паутина (WWW) и ее технологии на сегодняшний день являются самыми известными средствами доступа к информационным ресурсам Интернета.

Основой этой системы являются **web-страницы**, которые создаются с помощью **языка разметки гипертекста (HTML)**. Гипертекст - текст, содержащий в себе связи с другими текстами, графическими, звуковыми и видеофайлами. Благодаря этому языку **в текст web-страницы можно вставлять ссылки**, позволяющие осуществлять переход к другим частям этой же страницы или к другим документам. Таким образом, можно организовать связь между различными страницами и объединить их в единую систему.

Для просмотра содержимого таких страниц, чтобы не путать с обычным текстом, и перемещения по ссылкам необходима **специальная программа, которая устанавливает соединение с web-сервером**, посылает ему соответствующие запросы, получает от него содержимое страницы и представляет пользователю, обрабатывая ее должным образом. Такая программа называется **web-браузером**, или просто браузером.

Средства доступа к Internet



Для того чтобы получить доступ ко всем ресурсам Интернета необходимо установить **программное обеспечение**, которое взаимодействует с серверами и преобразовывает полученные данные в удобочитаемый вид, а также обеспечить **физическое соединение компьютера** с глобальной компьютерной сетью.

Все службы Internet основаны на использовании технологии “**клиент-сервер**”. Согласно этой технологии на компьютере клиента запускается **программа-клиент**, которая соединяется с сервером и передает или получает файлы, преобразовывая полученную информацию в удобочитаемый вид.

Для работы с электронной почтой наибольшее распространение получили такие почтовые клиенты как **Microsoft Outlook, The Bat**.

Доступ на FTP-сервера может осуществляться с помощью FTP-клиентов разных классов. К ним относятся **специальные программы (например, CuteFTP или WS FTP)**, а также FTP-клиенты, встроенные в файловый менеджер (**например, FAR или Windows Commander**).

Самыми популярными web-клиентами (web-браузерами) являются **Internet Explorer, Mozilla Firefox, Google Chrome**. С помощью этих браузеров можно просматривать содержимое web-страниц, принимать и посылать сообщения по электронной почте, копировать файлы с удаленных компьютеров.

Средства доступа к Internet



Способов физического подключения на сегодняшний момент много. Поэтому пользователь выбирает тот или иной способ подключения в зависимости от целей, которые он перед собой ставит, а также от того, какие **каналы связи** и технологии использует **провайдер Интернет-услуг**.

Модем – устройство, преобразующее цифровые данные, поступающие от компьютера, в аналоговый сигнал для передачи по телефонной линии и осуществляющий обратное преобразование. Слово образовано сокращением слов **модуляция** и **демодуляция**.

ISDN – интегрированная цифровая сеть.

ADSL – асимметричная цифровая абонентская линия.

Шлюз – компьютер, обеспечивающий передачу данных из одной сети в другую.

Хост - компьютер, имеющий самостоятельное подключение к Internet.

Скорость передачи данных характеризуется количеством передаваемых бит за 1 секунду (bit/s).

.

Средства доступа к Internet

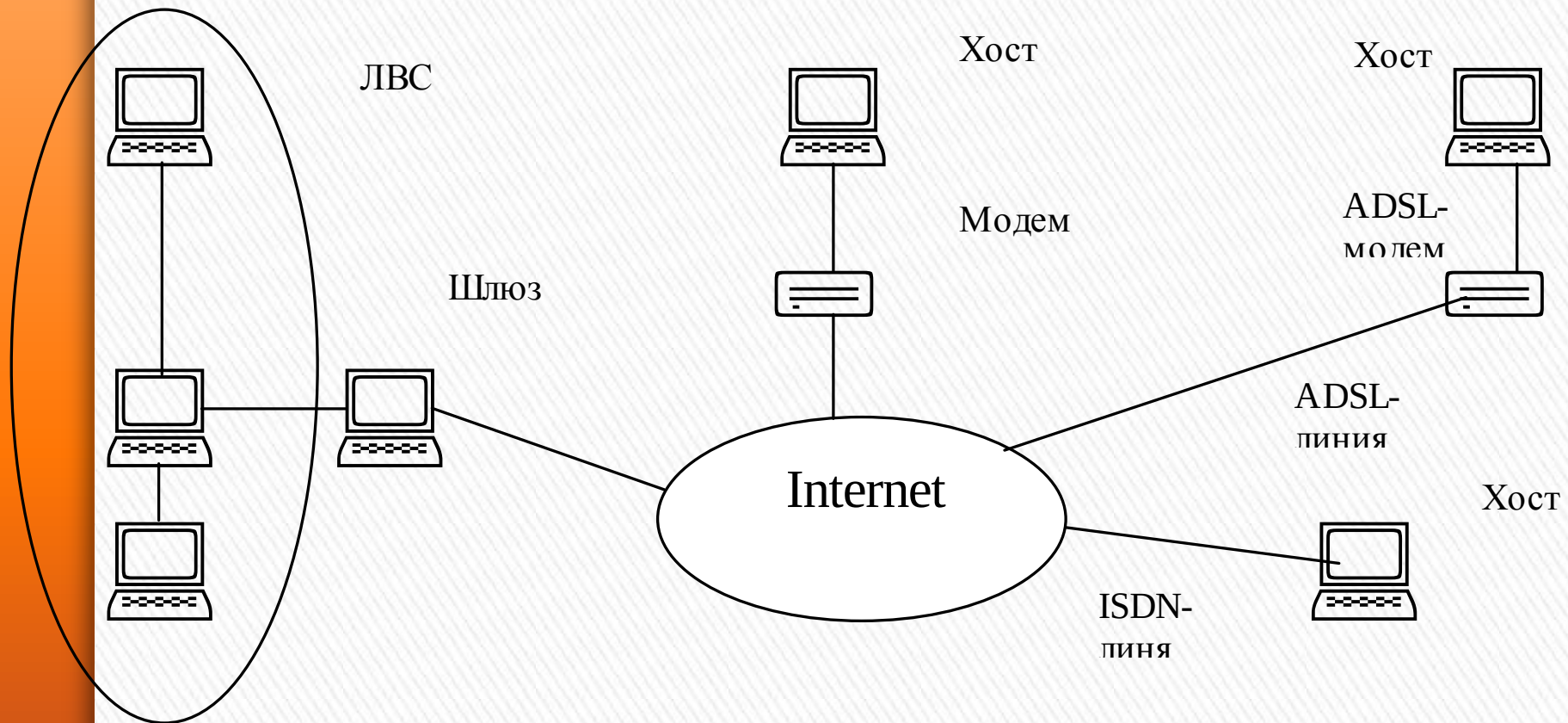


Наиболее характерными являются соединения по **телефонной линии**. Для этого доступа необходимо **наличие модема**. Существует два вида подключения: по **коммутируемой линии** (т.е. одновременно) и по **выделенной** (связь постоянна).

Технология **ADSL** позволяет организовать сверхскоростную передачу данных по телефонной линии с использованием **высокочастотного уплотнения сигналов**. Для этого доступа требуется специальный **ADSL-модем**, а на телефонной станции еще и уплотнительная аппаратура. Такая комбинация оборудования **образует три канала**: канал передачи данных из Интернета в компьютер, канал передачи данных из компьютера в Интернет и канал обычной телефонной связи. Скорость передачи данных варьируется от 32 Kbit/s до 1 Mb/s.

Кроме использования телефонных линий используются и другие способы соединения с Internet. К ним относятся **спутниковая связь**, осуществляющая прием данных с помощью параболической антенны, и **ЛВС (локальные вычислительные сети)**.

Схема доступа к ресурсам Internet >>>



Адресация в Internet



Для обмена информацией между компьютерами используется **способ разбиения информационного сообщения на пакеты** (группа байтов) и их передача. Этот способ был предложен еще в 1974 году. Дальнейшее использование межсетевого обмена информации привели к установлению **определенных правил.**

Протокол – набор правил по передаче и приему информации.

В настоящее время для доступа к Internet используется комплект протоколов, который получил название TCP/IP и включает в себя:

TCP – протокол управления передачей.

ICMP – вспомогательный межсетевой протокол управления сообщениями.

UDP – протокол передачи датаграмм пользователя.

Датаграмма – пакет, доставляемый адресату независимо от других пакетов последовательности, получаемой после разбиения сообщения на пакеты).

PPP – протокол равнодоступной связи.

IP – интернет протокол.

Адресация в Internet



TCP используется для обеспечения надежной, эффективной и достоверной передачи информации в виде потока байтов. **ICMP** поддерживает передачу управляющей информации и сообщений об ошибках между узлами сети. **UDP** предоставляет транспортные услуги процессам прикладного уровня. **PPP** - это протокол конечного пользователя (без привязки к способу передачи информации). **IP** управляет адресацией в сети.

Каждый компьютер в сети TCP/IP имеет адреса двух уровней:

- IP-адрес, состоящий из 4 байт, например, 10.1.20.20. Этот адрес используется на сетевом уровне. Он назначается администратором во время конфигурирования компьютеров. IP-адрес состоит из двух частей: номера сети и номера узла.
- Символьный идентификатор-имя, например, GSTU.GOMEL.BY. Этот адрес назначается администратором и состоит из нескольких частей, например, имени машины, имени организации, имени домена. Такой адрес, называемый также DNS-именем или доменным адресом, используется на прикладном уровне.

DNS (Domain Name System)



DNS (Domain Name System) - это распределенная база данных, поддерживающая иерархическую систему имен для идентификации узлов в сети Internet. Служба DNS предназначена для автоматического поиска IP-адреса по известному символьному имени узла.

IP-адрес преобразуется в 32-битное двоичное число и используется IP-протоколом для передачи или приема данных.

DNS-имя помогает определить местоположения компьютера. Вначале идет имя компьютера, затем имя сети, к которой он относится, которые разделяются точкой.

В отличие от IP-адреса оно читается справа налево. Сперва читается имя сети, а затем имя компьютера.

Каждая компьютерная сеть имеет выход на более обширную сеть. Поэтому имена сетей имеют **древовидную структуру**, называемого доменным пространством имен, в котором каждый **домен (узел дерева)** имеет имя и может содержать **поддомены**.

Имя домена идентифицирует его положение в этой базе данных по отношению к родительскому домену, причем точки в имени отделяют части, соответствующие узлам домена.

DNS (Domain Name System)



Домены верхнего уровня назначаются для каждой страны, а также на организационной основе. Для обозначения стран используются трехбуквенные и двухбуквенные аббревиатуры:

by – Беларусь; **ru** – Россия; **ua** – Украина; **us** – США.

Для различных типов организаций используются следующие аббревиатуры:

com - коммерческие организации;

edu - образовательные;

gov - правительственные организации;

org - некоммерческие организации;

net - организации, поддерживающие сети.

Например:

Адрес **gsu.by** описывает компьютер Гомельского государственного университета, расположенного в городе Гомель республики Беларусь.

Адрес **www.microsoft.com** описывает web-server фирмы Microsoft, относящейся к коммерческим организациям.

URL (Universal resource locator)



В Internet не только каждый компьютер имеет адрес, но и каждая web-страница. Такой адрес называется URL-адрес. **URL** – унифицированный указатель ресурсов. Этот адрес указывает на определенную страницу во всемирной паутине или Internet в целом. Он состоит из **двух частей**. Первая часть представляет собой идентификатор ресурса. Идентификатор ресурса указывает с помощью какого протокола необходимо обращаться к данному ресурсу. Существуют следующие идентификаторы ресурса:

HTTP – протокол передачи гипертекста;

FTP – протокол передачи файлов;

FILE – протокол доступа к файлу на локальной машине.

Вторая часть URL-адреса указывает на месторасположение ресурса, которая также состоит из двух частей. **Первая часть** указывает DNS-имя компьютера, а вторая показывает месторасположения файла на дисковом накопителе этого компьютера. В адресе страницы может отсутствовать часть, указывающая на месторасположение файла. В этом случае будет открыт документ, заданный по умолчанию.

Примеры URL-адресов:

<http://www.gstu.local/univer.html>

<ftp://cacedu.unibel.by>

<file://e:\internet\index.html>

ПРОСТОЙ ПОИСК ИНФОРМАЦИИ В INTERNET >>>

Поиск необходимой информации можно несколькими способами:

- 1) Используя **URL** - адреса, которые указывают на определенные web-сайты или страницы.
- 2) Используя ссылки, находящиеся на уже открытой странице.

Использование поиска по заданному URL-адресу возможна только при знании соответствующего адреса. Адреса наиболее популярных web-сайтов можно найти в компьютерных журналах, а также из телевизионных или печатных рекламных объявлениях. Наибольшую популярность получили так называемые "желтые страницы" Internet, которые есть в каждой библиотеке.

ПОИСК С ИСПОЛЬЗОВАНИЕМ ИНФОРМАЦИОННО-ПОИСКОВЫХ СИСТЕМ



- **Информационно-поисковая система (ИПС)** – система, обеспечивающая поиск и отбор необходимых данных в Интернет на основании входного языка и соответствующих правил поиска.
 - **Запрос на поиск** – строка информационно-поискового языка.
 - **Ключевые слова** – набор понятий и терминов, используемых при поиске.
 - **Информационно-поисковое пространство** – множество документов Интернет, по которым производится поиск.
 - **Результаты поиска** – список документов, найденных ИПС.
 - **Релевантность поиска** – мера соответствия результатов поиска запросам на поиск.
 - **Информационный шум (спам)** – информация, присылаемая пользователю, но не указанная в запросах на поиск (как правило, рекламного характера).
- Основными функциями ИПС являются:**
- получение запроса на поиск от пользователя;
 - обработка запроса;
 - определение точных адресов источников информации;
 - индексирование, т.е. описание источника информации и построение специальной базы данных (индекса);
 - вывод результатов поиска в удобном виде;
 - достижение релевантности при поиске.

ПОИСК С ИСПОЛЬЗОВАНИЕМ ИНФОРМАЦИОННО-ПОИСКОВЫХ СИСТЕМ



Информационно-поисковым пространством является следующий набор компонент WWW:

тексты в документах WWW;
графическая, аудио и видео информация WWW;
адреса e-mail;
группы новостей;
файлы FTP.

Эффективность поиска информации с помощью ИПС оценивается по таким параметрам как: полнота поиска, точность, актуальность и скорость поиска.

Полнотой поиска называется отношение числа найденных документов к общему числу релевантных документов.

Точность поиска — это отношение числа релевантных документов к общему числу полученных документов.

Актуальность ссылок на документы и **скорость поиска** определяются свойствами и возможностями поисковых систем.

ПОИСК С ИСПОЛЬЗОВАНИЕМ ИНФОРМАЦИОННО-ПОИСКОВЫХ СИСТЕМ



Все поисковые средства Интернет можно классифицировать следующим образом:

- 1) **Поисковые системы** (каталоги и поисковые машины).
- 2) **Метапоисковые системы.**
- 3) **Программы ускоренного поиска.**

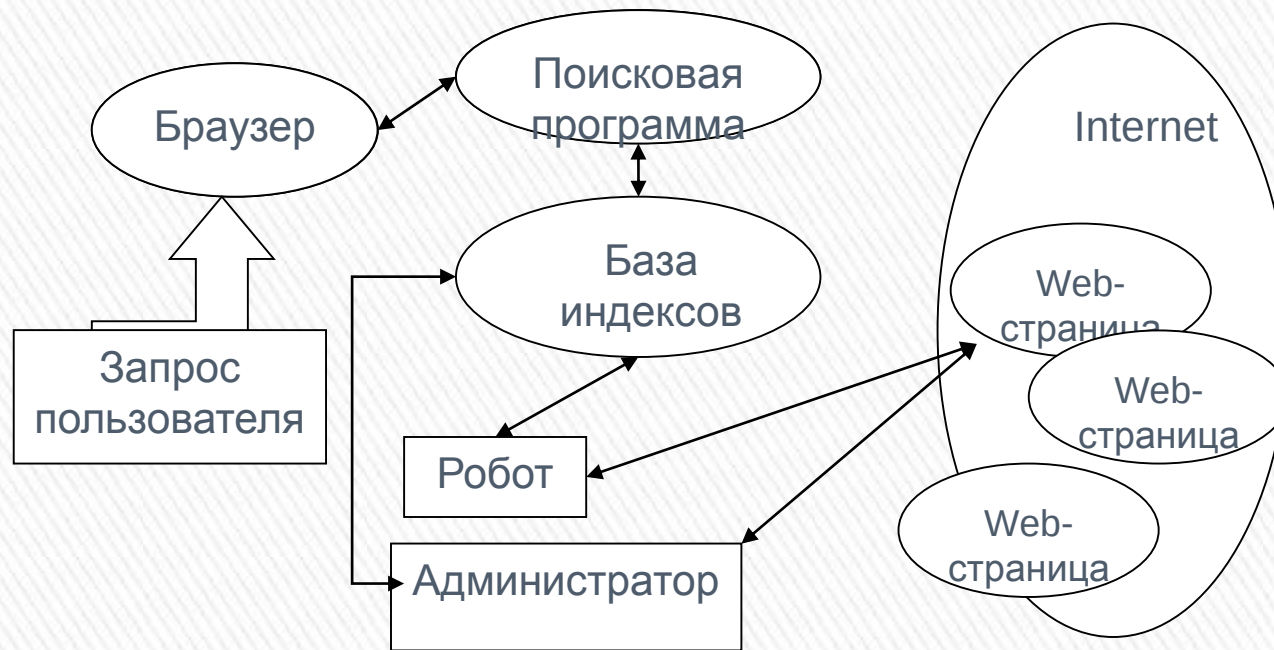
Каталогом называется поисковая система с классифицированным по темам списком аннотаций со ссылками на web-ресурсы. Поиск с помощью системы-каталога выполняется через последовательность уточненных тем или по ключевым словам. Наиболее распространенными каталогами являются **Yahoo** (адрес – www.yahoo.com), **List.ru** (адрес – www.list.ru), Созвездие Internet (адрес – www.stars.ru) и т.д.

Поисковой машиной называется **ИПС**, которая содержит специальную программу-робот, предназначенную для формирования базы данных, включающей информацию об объектах поиска (web-страницах, статьях Usenet и т.д.). Ниже приведены наименования и адреса наиболее известных поисковых машин.

Google – www.google.com; **Rambler** – www.rambler.ru; **Yandex** – www.yandex.ru; **AltaVista** – www.altavista.com; **Lycos** – www.lycos.com; **Tut.by** – www.tut.by

Метапоисковыми системами называются системы, способные послать запросы пользователя одновременно нескольким поисковым серверам, объединить полученные результаты и представить их пользователю в виде документа со ссылками. Наиболее популярными метапоисковыми системами являются **MetaCrawler** (адрес – www.metacrawler.com), **Search.com** (адрес – www.search.com).

Структура и принципы работы ИПС



Поисковая программа служит для перевода запросов пользователя, написанных на информационно-поисковом языке, в формальные запросы системы, а также для поиска ссылок на информационные ресурсы Интернет в базе индексов и выдачи результатов поиска. **Браузер** обеспечивает интерфейс пользователя с поисковой системой, просмотр результатов поиска. **Робот-индексировщик** – программа, которая служит для сканирования Интернет и поддержки базы индексов в актуальном состоянии.

Структура и принципы работы ИПС



Пользователь, работая с ИПС, может придерживаться определенных стратегий поиска, одна из которых имеет вид:

- Выбрать ИПС
- Прочитать справочную информацию
- Определить основные понятия, описывающие предмет поиска
- Выбрать ключевые слова
- Определить тип операторов поиска
- Осуществить поиск
- Просмотреть результат на предмет количества найденных документов и правильность синтаксиса
- Модифицировать запрос в зависимости от результата
- Попробовать другую ИПС

Структура и принципы работы ИПС

При поиске с помощью поисковых машин результат представлен в **виде последовательности**:

- список найденных документов по несколько единиц на странице с перечнем номеров таких страниц в последней строке и возможностью перехода к нужной странице;
- заголовок и URL-адрес найденного документа;
- степень релевантности документа;
- несколько первых предложений или фрагментов текста документа с выделенными ключевыми словами;
- дата обновления, размеры и язык документа.

Не все запросы на поиск в Интернет удовлетворяются, иногда возникают ситуации так называемого **нулевого поиска**, когда в ответ на запрос пользователя не выбрано ни одного документа. **Причин нулевого поиска** может быть несколько.

- Документ отсутствует в Internet.
- Документ отсутствует в той части Internet, которую обслуживает ИПС.
- Документ не проиндексирован по времени.
- Документ запрещен к индексированию.
- Документ находится в закрытой зоне.

Язык запросов на поиск



Для большинства поисковых машин можно провести поиск по шаблону или по маске, когда вместо окончания слова ставится **символ *** или поиск по усеченным словам, когда **метасимвол ?** используется в середине слова, заменяя любой символ.

Например, запрос Технолог*

позволит найти все ссылки, содержащие слова «технология», «технологический», «технологичный» и т.д.

«+» , «-» - **операторы** обязательного включения или исключения слова при поиске.

Например, запрос + математические – компьютерные системы

означает, что следует найти все документы, обязательно включающие слово «математические», но исключающие слово «компьютерные»

\$title – **оператор** поиска в заголовках документа.

Например, запрос \$title(Уравнения)

позволит найти все документы, у которых в заголовке есть слово «Уравнение».

Язык запросов на поиск



\$anchor – оператор поиска в тексте ссылки.

Например, запрос **\$anchor(Dialogic)**

позволит найти документы, в тексте ссылок которых есть слово «Dialogic».

#hint=(выражение) – оператор поиска в тексте подписи к изображениям.

Например, запрос **#hint=(Чертеж)**

позволит найти документы, в подписи под изображением, которых есть слово «Чертеж».

#url="значение" – оператор поиска документов на сайте, адрес которого задан.

Например, запрос **#url="www.exponenta.ru"**

позволит найти документы на сайте www.exponenta.ru.

#link="значение" – оператор поиска документов, содержащих ссылка на сайт, адрес которого задан в поле “значение”.

Например, запрос **#link=" www.exponenta.ru ”**

позволяет найти документы, ссылающиеся на сайт с адресом www.exponenta.ru.

Язык запросов на поиск



При формировании сложного запроса на поиск в поисковых машинах используются знаки логических операций.

Примеры поиска с применением сложных запросов.

В Yandex:

Спорт (NOT футбол) – позволяет найти документы, в которых присутствует слово «спорт», но отсутствует слово «футбол».

Футбол | Хоккей – позволяет найти документы, в которых присутствует либо слово «футбол», либо слово «хоккей».

Волейбол & баскетбол – позволяет найти документы, содержащие одновременно и слово «волейбол », и слово «баскетбол».

В Rambler:

(современные OR информационные) & технологии – позволяет найти документы, содержащие слова «современные» «технологии» или слова «информационные» «технологии».