

Лекция № 13

Интернет. История. Структура. Сервисы.

История сети Интернет

Слово *Интернет (Internet)* образовано от двух слов: Inter – между и Net – сеть, т.е. межсетевой.

Интернет – это глобальная сеть, состоящая из многих сетей и отдельных компьютеров, поддерживающих связь путем использования общих протоколов.

Его разработку начали американские военные в 60-х годах XX века. К тому времени уже существовали компьютерные сети и информационные сетевые технологии, и основная задача, которая ставилась перед Интернет - это обеспечение компьютерных сетевых коммуникаций территориально распределенных узлов - компьютерных сетей - при нанесении противником ударов по территории, в том числе и ядерных, при которых возникают колоссальные разрушения инфраструктур.

На базе особых правил передачи данных, которые называются сетевыми протоколами - TCP/IP (TransmissionControl Protocol и Internet Protocol), в 1969 в Минобороны США был завершен проект по совместному использованию ресурсов Минобороны, университетов и других правительственных учреждений. Создаваемая система сначала называлась ARPANET - Advanced Research Projects Agency Net, затем, после гигантского разрастания по всему миру получил название Internet. Интернет - по замыслу и существу межсетевая система, система, объединяющая компьютерные сети (по другим источникам первым протоколом сети был NCP²).

В 1971 году появилась первая программа для отправки электронной почты.

В 1973 через трансатлантический кабель к сети подключились первые иностранные государства – Норвегия и Великобритания.

В 70-х годах сеть в основном использовалась для отправки электронной почты, использовались также доски объявлений и новостные группы.

1973 - утверждение протокола FTP для пересылки файлов между компьютерами

1979 - первая многопользовательская игра

1983 - внедрение протокола TCP/IP

В 1984 году разработана система доменных имен (DNS), а также появилась сеть FIDO. С этого года начинается активное развитие сети NSFNet (National Science Foundation Network) – сеть Национального научного фонда США.

1986 - NSFnet (обеспечение доступа ученых к суперкомпьютерам) - 1000 компьютеров в Сети

В 1988 году изобретен протокол IRC – появились первые чаты.

В 1989 году в стенах CERN (Швейцария) родилась идея *Всемирной паутины*. Ее предложил британец Тим Бернерс-Ли, который в течение двух лет разработал протокол HTTP, язык HTML и идентификаторы URL.

В 1990 году ARPANET прекратила свое существование проиграв конкуренцию NSFNet.

В 1991 году WWW стала общедоступной в Интернет, NREN (подключение школ и колледжей) – более 100 000 компьютеров подключенных к Сети

В 1995 году маршрутизацией сети стали заниматься сетевые провайдеры, а не суперкомпьютеры Национального научного фонда США. В этом же году по трафику HTTP обогнал FTP, став основным поставщиком информации в Интернет.

В 90-е годы Интернет объединил в себе большинство сетей, за исключением пожалуй FIDOnet, а также закрытых специализированных сетей.

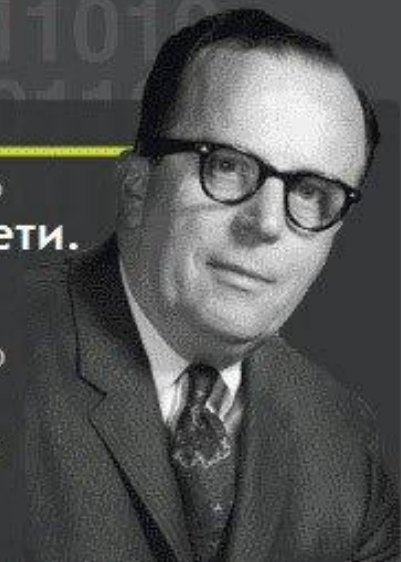
В СССР Интернет стал зарождаться с 1952 года благодаря работам ИТМиВТ АН СССР в рамках работ по созданию автоматизированной системы противоракетной обороны. Сеть использовалась для расчета траекторий полета ракет (также это можно назвать и одной из первых GRID - сетей). С 1972 года работала единая сеть передачи данных в рамках кассовых операций АСУ «Экспресс» в Министерстве путей и сообщений. Выход во всемирную среду осуществился 28.08.1990 благодаря Институту атомной энергетики им. Курчатова и ИПК «Минавтопрома». Домен “ru” зарегистрирован в 1994 году (до этого был su – Soviet Union, который существует до сих пор).

История Интернета

'62

Джозеф Ликлайдер (1915-1990) высказал идею о создании **Всемирной компьютерной сети.**

Он придерживался идеи создания объединения компьютеров в сеть со свободным доступом любого человека к ее ресурсам. Он возглавлял агентство передовых оборонных исследовательских проектов (DARPA) и убеждал своих преемников в необходимости развития компьютерных сетей.



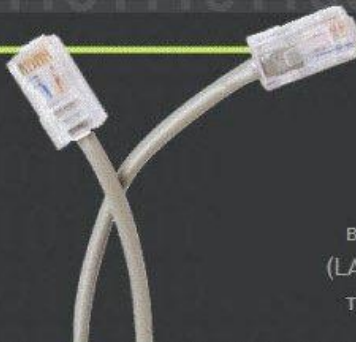
'74

Винтон Серф и Роберт Кан впервые употребляют термин

Интернет

в отчете о разработке набора сетевых протоколов.

'76



Роберт Меткалф из корпорации Xerox **изобретает Ethernet** - коаксиальные кабели для быстрой передачи данных. Ethernet является важной составляющей локальных сетей (LAN), работающих обычно на небольшой территории, в доме, офисе или школе.

'78

Гэри Туэрк осуществил **первую рассылку спама**, разослав четыремстам пользователям сети ARPANET рекламу технических разработок для микрокомпьютеров.



'83

1 января 1983 года **все компьютеры**, подключенные к сети ARPANET **были переведены на протокол TCP/IP** (Transmission Control Protocol/Internet Protocol). Это событие стало началом современного Интернета.

'84

Джон Постел описал **идею создания зон .com, .gov, .org, .edu и .mil** в серии работ, опубликованных в специальной комиссии интернет-разработок (Internet Engineering Task Force, IETF).



'85

Деннис Дженингс основал Национальный фонд науки США. Это определило тот факт, что **протокол TCP/IP стал обязательным** для всей программы NSFNet, объединявшей ВУЗы США.

Когда годом позже фонд возглавил Стив Волф, он решил что научным и исследовательским сообществам потребуется что-то кроме государственного финансирования. Он предположил, что этим будет глобальная компьютерная сеть (WAN), которая соединит локальные сети (LAN).

'95

Пьер Омидьяр запустил онлайн-аукцион AuctionWeb, превратившийся впоследствии в eBay.



Введение ежегодной платы за аренду

доменов, за исключением случаев использования их в правительственной (.gov) и образовательной (.edu) зонах.

'96



Запуск проекта Интернет2, широкополосной компьютерной сети для ВУЗов, рассчитанной на разработки в сфере образования.



Появление почтовой службы Hotmail.

На следующий год Microsoft покупает Hotmail за 400 тыс. долларов.

'98

Google получает 100 тыс. долларов от Энди Бечтольшайма (Sun Microsystems) и основывает компанию Google Technology Incorporated.



'99

Стандартизирован Wi-Fi - технология беспроводного доступа к интернету.



Шон Фаннинг придумал и запустил печально известный Napster. В 18 лет.



'00

Пузырь доткомов лопается,
вызывая обвальное падение индексов NASDAQ,
достигнувшего перед этим высочайшей отметки в
5048,62 пунктов.



'01



Запуск Википедии,
впоследствии сменившей свою доменную
зону с .com на .org.

'03

**Открытие интернет-магазина
iTunes Store компанией Apple** всего
с 200 тыс. композиций. За первые сутки было
продано четверть миллиона песен.



'04



1 апреля 2004 года **Google запускает**
собственный **почтовый сервис Gmail** с
хранилищем для писем объемом 1 Гб. Сначала
это было воспринято как первоапрельская
шутка, так как в то время Hotmail и Yahoo
предоставляли 2 и 4 Мб соответственно.

'05

Появление YouTube,
революционного видео-сервиса в
Интернете. На следующий год
Google покупает YouTube за 1,6
млн. долларов.



'06



Дом Саголла запускает Twitter.



Facebook, рассчитанный изначально только на студентов, становится доступен всем.

'09

Рост данных, передаваемых мобильными устройствами по сравнению с голосовым траффиком каждый месяц.

Впервые подобный траффик по всему миру превышает экзбайт (миллион гигабайт).



Sources:

<http://www.isoc.org/internet/history/brief.shtml> • <http://www.davesite.com/webstation/net-history2.shtml>
http://en.wikipedia.org/wiki/Local_area_network • http://en.wikipedia.org/wiki/J._C._R._Licklider
<http://en.wikipedia.org/wiki/ARPANET> • http://en.wikipedia.org/wiki/Internet_Protocol_Suite
<http://gigaom.com/2009/02/01/a-brief-history-of-twitter/> • <http://www.std.com/> • <http://www.panix.com/>
http://www.zdnet.com.au/50-significant-moments-from-internet-history-339292301.htm?omnRef=http%3A%2F%2Fwww.google.com%2Fsearch%3Fhl%3Den%26safe%3Dof%26q%3Dmobile%2Binternet%2Bhistory%26aq%3Df%26aq%3Dg1%26aq%3D%26oq%3D%26gs_rtl%3D
http://en.wikipedia.org/wiki/Dot-com_bubble#The_bubble_bursts

Design by Jason Powers
Перевёл snarker

 OnlineSchools
www.OnlineSchools.org

Организация объединенных наций 3 июня 2011 г. включила свободный доступ в Интернет в список базовых прав человека. 24 октября 1995 года Федеральный сетевой совет (FNC) одобрил резолюцию, определяющую термин "Интернет".

РЕЗОЛЮЦИЯ: Федеральный сетевой совет признает, что следующие словосочетания отражают наше определение термина "Интернет".

Интернет— это глобальная информационная система, которая:

- Логически взаимосвязана пространством глобальных уникальных адресов, основанных на Интернет-протоколе (IP) или на последующих расширениях или преемниках IP;
- Способна поддерживать коммуникации с использованием семейства Протокола управления передачей/Интернет-протокола (TCP/IP) или его последующих расширений/преемников и/или других IP-совместимых протоколов;
- Обеспечивает, использует или делает доступной, на общественной или частной основе, высокоуровневые сервисы, надстроенные над описанной здесь коммуникационной и иной связанной с ней инфраструктурой.

Организационная структура сети Интернет

Не существует какой-либо одной организации, которая бы владела или управляла сетью Интернет или контролировала бы ее. Эта функция выполняется благодаря слиянию усилий независимых сетей. Сети-участники могут возглавляться президентами или руководителями, однако не существует единого управления сетью Интернет.

Существуют несколько организационных подразделений, отвечающих за развитие Интернет:

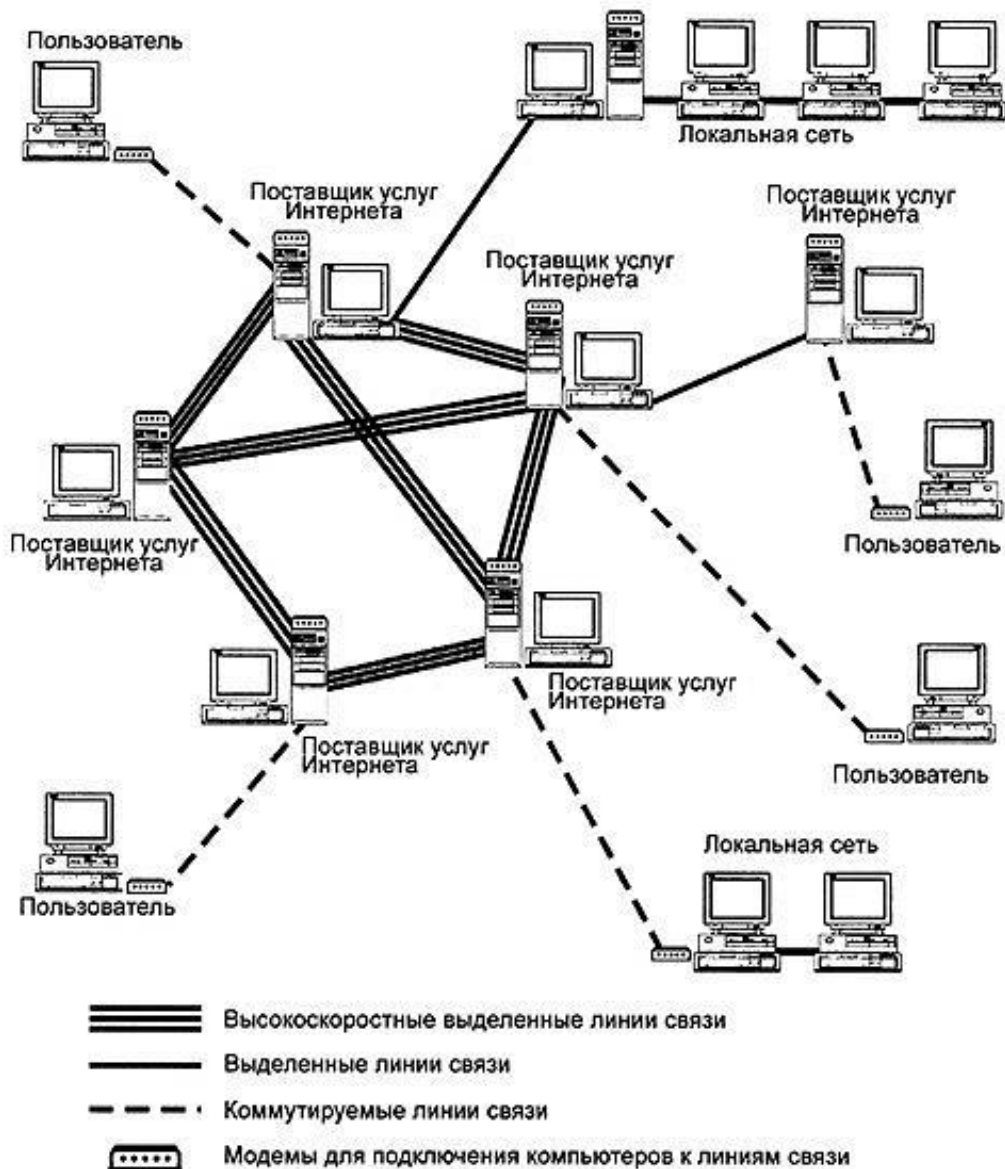
1. Основным из них является Internet Society (ISOC) – профессиональное сообщество, занимающееся вопросами роста и эволюции Интернет, как глобальной коммуникационной инфраструктуры. Под управлением ISOC работает Internet Architecture Board (IAB) – организация, осуществляющая технический контроль и координация работ для Интернет. В IAB входят 2 группы:

- IETF занимается решением технических проблем Интернет;
- IRTF координирует долгосрочные проекты по протоколам TCP/IP.

2. Для наделения региональной сети конкретным IP – адресом существует специальное подразделение Интернет InterNIC (Internet Network Center).

3. Широкое распространение Всемирной паутины вызвало создание W3-консорциума, занимающейся развитием Web.

Структура сети Интернет



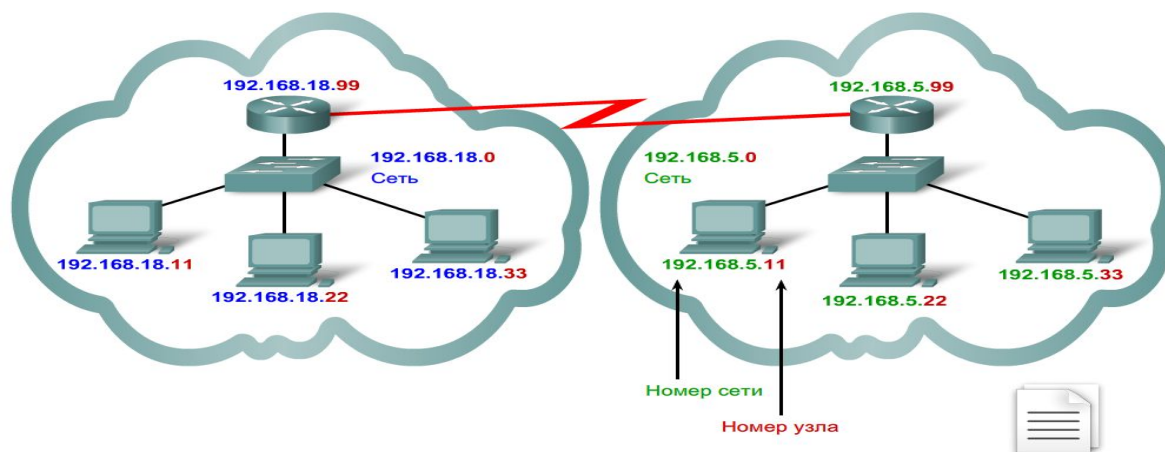
Компьютер пользователя с помощью линии связи подключается к компьютеру провайдера, который, в свою очередь подключен к другому компьютеру сети и т.д. Информация в сети хранится как на компьютерах провайдера, так и на специальных компьютерах, которые называются информационными серверами. Компьютеры, к которым подключаются многие другие компьютеры называют *серверами*. *Провайдером* называется организация, через которую рядовые компьютеры подключаются к глобальной сети.

Адресация компьютеров в сети

IP-адреса

Каждому компьютеру в сети Интернет назначается уникальный числовой адрес – IP-адрес. IP-адрес состоит из 4-х разделенных точками чисел в диапазоне от 0 до 255, т.е. является массивом 4-х байт. Числа слева обозначают адрес сети, а правее адрес компьютера в этой сети. Количество чисел для сети и компьютера может быть от 1 до 3.

Классовая схема IP-адресов (INET)



Бесклассовая IP адресация (CIDR) В этом методе адресации используются маски подсети которые определяют, какая часть IP-адреса относится к сети, а какая – к узлу, и в этом случае граница разделения адреса сети и компьютера может проходить посреди любого из 4-ех байтов. Кроме того, эта система упрощает маршрутизацию.

Система доменных имен (DNS-имен, Domain Name System)

Домен – это совокупность компьютеров в составе сети Интернет, объединенных каким-либо общим признаком, например, принадлежащих одной стране, одной компании и т.п.

Доменная система имен – это метод назначения имен путем передачи сетевым группам ответственности за их подмножество имен.

Каждый уровень этой системы называется доменом.

Домены в именах отделяются друг от друга точками:

site1.russia.ru, site2.ru, microsoft.com.

В имени может быть различное количество доменов, но практически их не больше пяти.

Домен первого уровня (крайний справа) определяет национальную сеть • Россия – ru, Украина – ua, Германия – de и т. д. США: сети образовательных (edu), коммерческих (com), государственных (gov), военных (mil) учреждений, а также сети других организаций (org) и сетевых ресурсов (net), Администрация доменов может создавать домены более низкого уровня.

Например, в домене первого уровня ru может быть создан домен второго уровня krsk.ru, в котором в свою очередь администрация может создать домен третьего уровня education.krsk.ru.

Доменные имена при работе сети Интернет преобразуются в IP-адреса.

Сервисы Интернет

- Электронная почта (E-mail).
- Телеконференции (UseNet).
- Списки рассылки (Mail-list).
- Удаленный доступ (TelNet).
- Служба передачи файлов (FTP).
- Интерактивное общение.
- Передача звука в реальном времени.
 - ✓ Радиовещание в Интернет.
 - ✓ Компьютерная телефония.
 - ✓ Аудиоконференции.
 - ✓ Видеоконференции.

Электронная почта (E-mail)

Электронная почта - это сетевая служба, которая дает возможность пользователям посылать и получать электронные сообщения. Обычно сообщение посылается в виде текста, адресат получает его на своем компьютере через какой-то промежуток времени и читает тогда, когда ему будет удобно.

Почтовая служба основана на двух прикладных протоколах: **SMTP** и **POP3**. По первому происходит отправка корреспонденции с компьютера на сервер, а по второму - прием поступивших сообщений.

Существует большое разнообразие клиентских почтовых программ. К ним относится программа Outlook Express, входящая в состав операционной системы Windows как стандартная.

Чтобы отправить письмо, необходимо знать адрес. Все E-mail-адреса Internet выглядят одинаково. Например, TIV@IS.ISEA.BRATSK.RU. Адрес состоит из двух частей, разделенных символом @. Слева находится имя пользователя (имя почтового ящика), справа - имя домена (имя машины). При рассмотрении имени домена справа налево и разбиении его по точкам на отдельные слова получается набор, описывающий ее положение в сети.

Самое правое слово (в нашем случае RU) называется доменом верхнего уровня и чаще всего (но не всегда) указывает код страны, в которой находится адресат. Домен верхнего уровня может обозначать название сети, в которой находится адресат. BRATSK - регион (может быть название города), ISEA - фирма, IS - подразделение фирмы. Адреса и почтовые ящики обычно получают у администратора сети. Для Internet важна правая часть адреса. Имя домена - это уникальное имя, присвоенное конкретному компьютеру в Internet.

Достоинствами электронной почты являются:

- простота использования (самое сложное здесь - соединение с Internet первый раз);
- высокая скорость передачи сообщений;
- дешевизна (плата берется только за соединение);
- сообщения электронной почты легко хранить (в виде файлов);
- электронная почта практически универсальна (передача через любые сети текстов, графических изображений).

К **недостаткам** электронной почты можно отнести слабую защиту сообщений (возможность доступа третьих лиц).

Телеконференции (UseNet).

Телеконференции - вторая по распространенности служба Интернета, предоставляющая отложенные услуги.

Служба телеконференций состоит из множества тематических телеконференций - групп новостей (newsgroup), поддерживаемых серверами новостей. **Сервер новостей** - это компьютер, который может содержать тысячи групп новостей самых разнообразных тематик.

Группа новостей - это набор сообщений по определенной теме. Новости разделены по иерархически организованным тематическим группам, и имя каждой группы состоит из имен подуровней. Существуют как глобальные иерархии, так и иерархии, локальные для какой-либо организации, страны или сети. Набор групп, получаемых сервером телеконференций, определяется его администратором и их наличием на других серверах, с которыми данный сервер обменивается новостями.

Доступ к группам новостей осуществляется через процедуру подписки, которая состоит в указании координат сервера новостей и выбора интересующих пользователя групп новостей. Следует заметить, что каждый сервер новостей имеет определенный набор конференций, и, если интересующая тематика на нем не найдена, можно попробовать использовать другой сервер. Данная процедура, а также работа с группами новостей осуществляется с помощью программного обеспечения, поддерживающего эти функции, например, широко распространенным приложением компании Microsoft - Outlook Express.

В обсуждении темы телеконференции может участвовать множество людей, независимо от того, где они находятся физически. Обычно, хотя это и не является правилом, за порядком в конференциях следят специальные люди, так называемые модераторы. В их обязанности входит поддержание порядка в конференции в соответствии с установленными в ней правилами поведения и ее тематикой.

Наряду с описанной формой служб телеконференции широкое распространение получили WWW-телеконференции, также называемые форумами. Отличие состоит в том, что они работают через web-интерфейс, и размещаются не централизованно на серверах новостей, а на web-сайтах.

Списки рассылки

Списки рассылки (maillists) - простой, но в то же время весьма полезный сервис Интернет. Это практически единственный сервис, не имеющий собственного протокола и программы-клиента и работающий исключительно через электронную почту.

Идея работы списка рассылки состоит в том, что существует некий адрес электронной почты, который на самом деле является общим адресом многих людей - подписчиков этого списка рассылки. Вы посылаете письмо на этот адрес, и ваше сообщение получают все люди, подписанные на этот список рассылки.

Такой сервис по задачам, которые он призван решать, похож на сетевые новости Usenet, но имеет и существенные отличия.

Во-первых, сообщения, распространяемые по электронной почте, всегда будут прочитаны подписчиком, дождавшись его в почтовом ящике, в то время как статьи в сетевых новостях стираются по прошествии определенного времени и становятся недоступны.

Во-вторых, списки рассылки более управляемы и конфиденциальны: администратор списка полностью контролирует набор подписчиков и может следить за содержанием сообщений. Каждый список рассылки ведется какой-либо организацией, и она обладает полным контролем над списком.

В-третьих, для работы со списком рассылки достаточно доступа к электронной почте, и подписчиками могут быть люди, не имеющие доступа к новостям Usenet или каким-либо группам этих новостей.

В-четвертых, такой способ передачи сообщений может быть просто быстрее, коль скоро сообщения передаются напрямую абонентам, а не по цепочке между серверами Usenet.

Удаленный доступ (TelNet)

Приложения, позволяющие осуществить заход удаленным терминалом, очень популярны в Internet.

TelNet – одна из самых старых информационных технологий Интернета, относящаяся к сервисам прямого обращения. Под TelNet понимают триаду, состоящую из TelNet-интерфейса пользователя, TelNet-процесса и TelNet-протокола. Эта триада обеспечивает описание и реализацию сетевого терминала для доступа к ресурсам удаленного компьютера.

Протокол TELNET позволяет вам подсоединиться к удаленному компьютеру и работать с ним, как будто бы вы используете локальную систему.

Telnet - стандартное приложение, которое присутствует практически в каждой реализации TCP/IP.

Служба передачи файлов FTP

FTP (file transfer protocol) - протокол передачи файлов, но при рассмотрении FTP как службы Интернета имеется в виду не просто протокол, а именно служба доступа к файлам в файловых архивах. Служба FTP имеет свои серверы в мировой сети, на которых хранятся архивы данных. Объемы информации в FTP-архивах очень значительны. Практически любой архив строится как иерархия каталогов. Многие архивы дублируют информацию из других архивов (зеркала).

Протокол FTP работает одновременно с двумя TCP-соединениями между сервером и клиентом. По одному соединению идет передача данных, а второе соединение используется как управляющее. Протокол FTP также предоставляет серверу средства для идентификации обратившегося клиента. Однако существуют FTP-серверы с анонимным доступом для всех желающих.

Преимущества использования FTP:

- Возможность передавать файлы через Интернет (при помощи данного протокола можно скачивать на свой компьютер любые файлы: MP3, ZIP архивы, текстовые и программы).
- Возможность удалённо управлять файлами компьютера.
- Удобство в работе с файлами.

Для работы с FTP необходимы специальные программы, которые и обеспечивают данный протокол.

Есть несколько **типов FTP-программ**.

1. FTP-сервер - это специальная программа, запущенная на компьютере и работающая на нём в фоновом режиме. Она делает из обычного компьютера FTP-сервер, с которого можно скачать файлы или закачать. Она отслеживает все запросы приходящие от других компьютеров, обрабатывает их и даёт ответ. При установке сервера FTP указывается доступная для других программ клиентов директория. Все файлы и директории имеют свои атрибуты, которые ограничивают к ним доступ с удалённых компьютеров. Например, можно сделать так, чтобы один файл был доступен только для чтения, другой только для выполнения, третий только для записи, четвёртый полностью открыт для других машин и так далее. То же самое и с директориями.

2. FTP-клиент - это специальная программа, которая работает с FTP-серверами. Она предназначена для подачи запросов серверам и приёму информации. Фактически она устроена также как и сервер, только управляется она с локальной машины либо автоматически, либо человеком. Управлять клиентом через сеть нельзя. FTP-клиенты могут предоставлять пользователям удобный интерфейс.

Интерактивное общение

Интерактивная услуга - услуга, ориентированная на удовлетворение информационных потребностей пользователей путем предоставления информационных продуктов в диалоговом режиме. В 1998 году разработчики израильской компании Mirabilis создали специальное программное обеспечение ICQ и перенесли возможности вызова абонентов в сеть Интернет. Программа получила название Интернет-пейджера.

Передача звука в реальном времени

Радиовещание в Интернет

Появилось несколько возможных способов применения звуковых потоков в Интернет. Стало очевидным перенесение систем радиовещания из обычного радиоэфира в цифровые сети. Тысячи радиостанций установили у себя потоковые серверы и начали передавать свои обычные «живые» программы в Интернет. Радиостанции так называемого иновещания используют особенный формат, учитывающий условия распространения радиоволн и поясное время в местах приема.

Компьютерная телефония

Самой заметной из последних новинок звуковых коммуникаций в Интернет является ПО, позволяющее географически удаленным пользователям устанавливать телефонную связь через Интернет и платить за нее намного меньше, чем за такой же звонок по обычным телефонным линиям. Ключевым элементом интернет-телефонии является связка «шлюз – Интернет». Шлюз представляет собой компьютер-сервер, дополненный специальными платами расширения и соответствующим программным обеспечением. Он служит интерфейсом между передающим звук устройством пользователя (телефоном, компьютером и т.п.) и сетью Интернет.

Аудиоконференции

Аудиоконференция - это процесс использования электронных каналов связи для организации совещаний между тремя и более участниками, при котором происходит голосовая коммутация (соединение) участников конференции. У каждой конференции есть координатор (ведущий конференции), который управляет ей и следит за тем, чтобы не нарушалась тематика конференции, этикет и т.п.

Видеоконференции и их типы








Видеоконференция - это интерактивный инструмент, который включает в себя аудио, видео, компьютерные и коммуникационные технологии для осуществления связи удаленных территориально собеседников «лицом к лицу» в реальном времени, а также разделения всех типов информации, включая данные, звук, изображение, документы и т.п. Видеоконференции разделяются на персональные (ПВ), групповые (ГВ) и студийные (СВ).

Адресация файлов в Интернете URL (Universal Resource Locator – универсальный указатель ресурсов) –это стандартизированный способ записи адреса ресурса в сети Интернет. URL указывает протокол (метод доступа), компьютер и местонахождение ресурса на компьютере.

В общем случае URL имеет формат:

протокол://адрес компьютера/путь/имя файла

Например, <http://www.bspu.unibel.by/conf2003/sec2/dokl.htm>

Протокол	Сервер	Путь к ресурсу
ftp://	ftp.microsoft.com/	Products/mspress/library/  ANIMAT.ZIP  index.txt  MEMLOG.ZIP  MFCANIM.ZIP  NAVDEMO.EXE  rkhelp.exe
http://	www.home.tltsu.ru/	/~mok/2004/index.html 

WWW (World Wide Web – всемирная паутина) – совокупность взаимосвязанных гипермедийных (гипертекстовых и мультимедийных) документов (Web-страниц).

Мультимедиа – документ включает в себя не только текст, но и графику, видео и звук.

Гипертекстовая ссылка – выделенная часть документа, реализующая переход от одного документа к другому или одной части документа к другой.

Web-страница – документ WWW, который может содержать:

- форматированный текст;
- мультимедийные объекты (графика, звук, видео);
- ссылки на другой файл или на другую Web-страницу.

Web-сайт (site – местонахождение) – группа Web- страниц, связанных вместе единой темой, общим стилем и взаимными гипертекстовыми ссылками.

Web-сервер – компьютер, хранящий Web-страницы.

Браузер (Web-браузер) – программа Web-клиент идля навигации в WWW и просмотра Web-страниц.

HTTP(Hyper Text Transfer Protocol, протокол передачи гипертекста) – протокол, который использует Web-клиент для получения Web-страницы с Web-сервера.

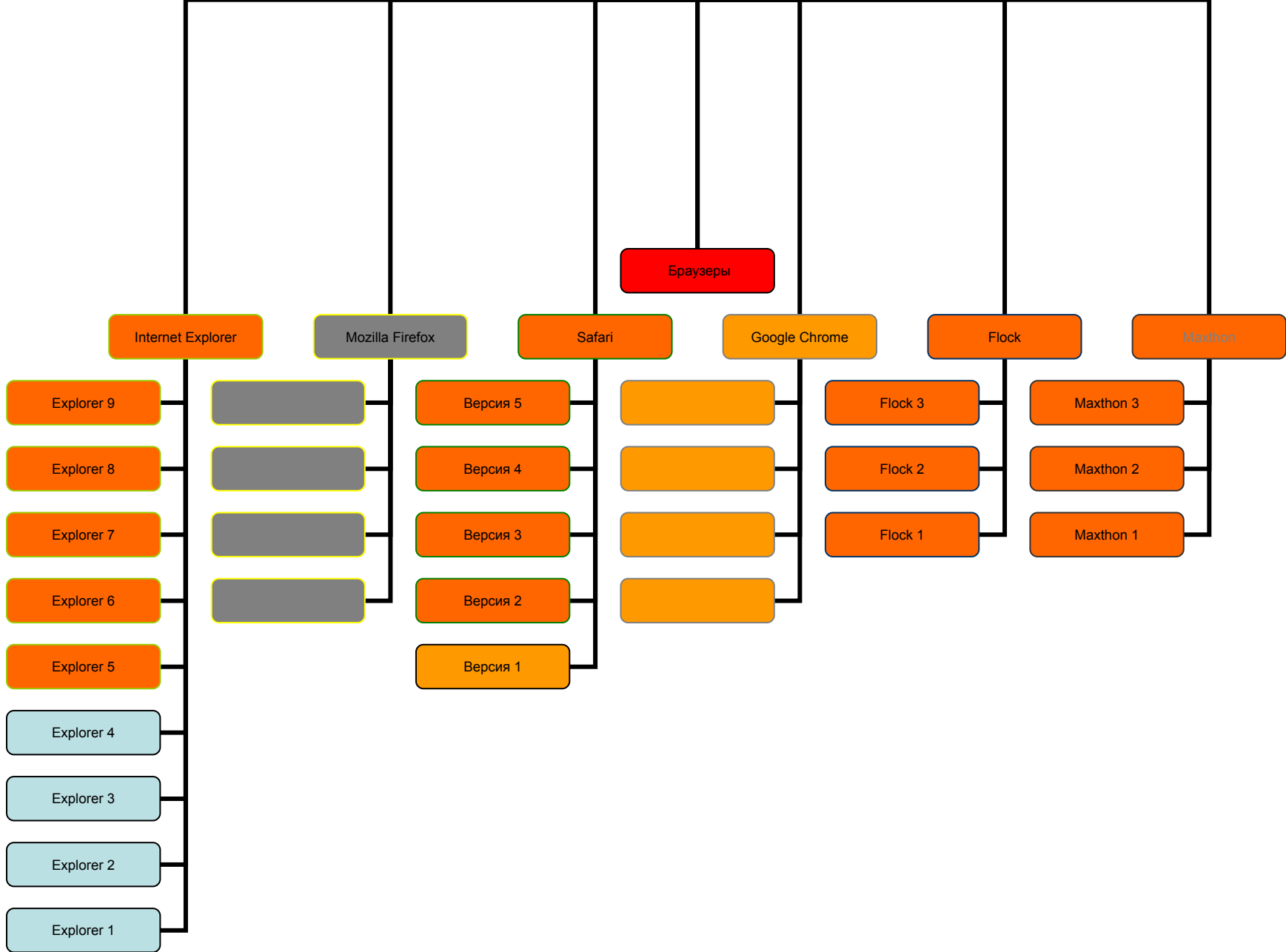
Средства использования сетевых сервисов

Браузеры

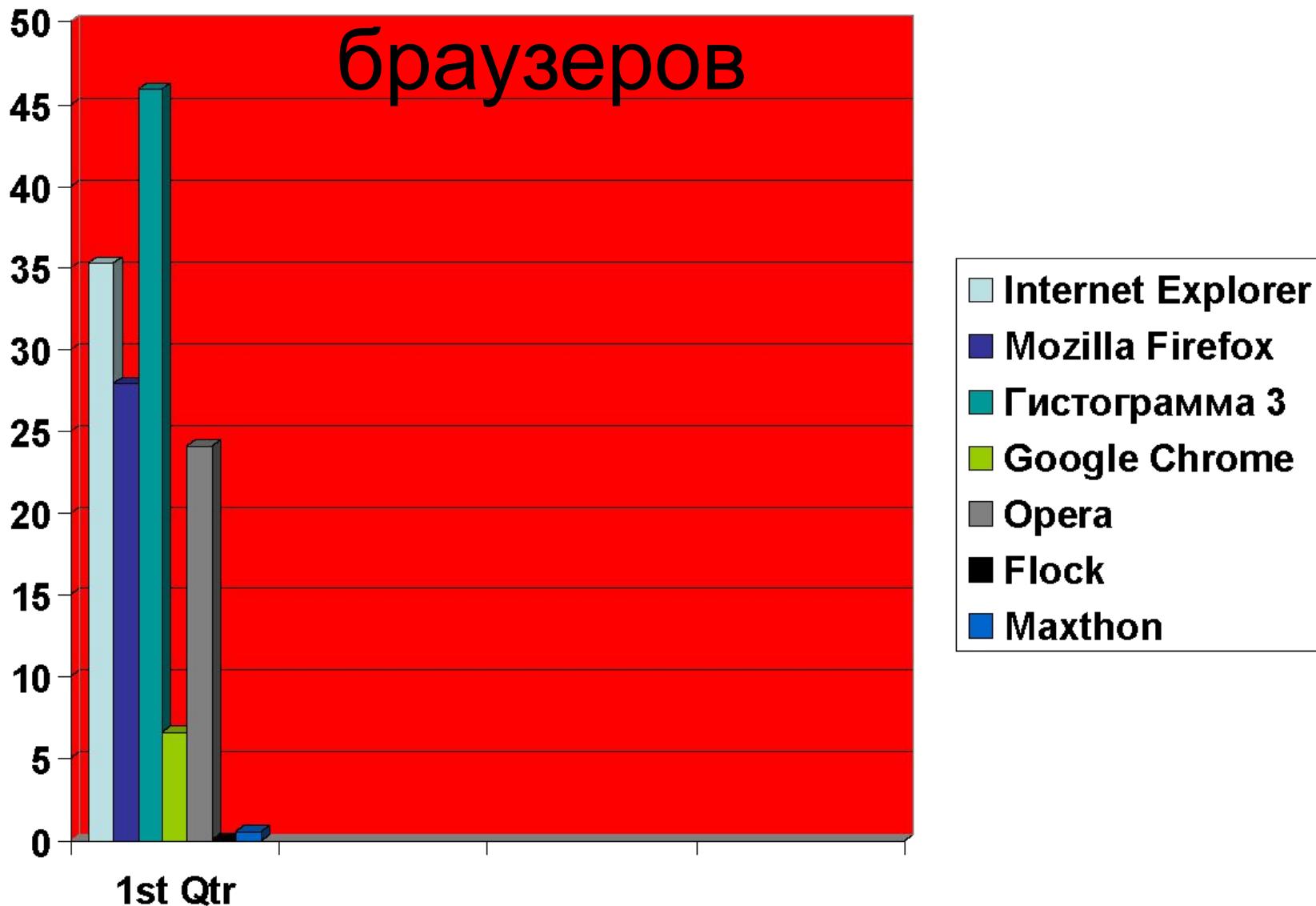


Браузер - это специальная программа, предназначенная для просмотра интернет-страниц. Браузер запрашивает с сервера информацию, интерпретирует ее особым образом и создает веб-страницу на экране пользователя. Создание и интерпретация страниц производится по отраслевым стандартам, утверждаемым международной стандартизирующей организацией.





Статистика использования



Информационно-поисковые системы Интернета

Все поисковые системы объединяет то, что они расположены на специально-выделенных мощных серверах и привязаны к эффективным каналам связи. Поисковые системы называют еще информационно-поисковыми системами (ИПС). Количество одновременно обслуживаемых посетителей наиболее популярных систем достигает многих тысяч. Самые известные обслуживают в сутки миллионы клиентов. В случаях, когда поисковая система имеет в своей основе каталог, она называется каталогом. В ее основе лежит работа модераторов. В основе же ИПС с полнотекстовым поиском лежит автоматический сбор информации. Он осуществляется специальными программами. Эти программы периодически исследуют содержимое всех ресурсов Интернета. Для этого они перемещаются, или как говорят, ползают, по разным ресурсам. Соответственно такие программы называются роботы. Есть и другие названия: поскольку WWW – это аббревиатура выражения Всемирная паутина, то такую программу естественно назвать спайдером по англ. – паук. В последнее время используются другие названия: автоматические индексы или директории. Все эти программы исследуют и «скачивают» информацию с разных URL-адресов. Программы указанного типа посещают каждый ресурс через определенное время. Ни одна поисковая система не в состоянии проиндексировать весь Интернет. Поэтому БД, в которых собраны адреса проиндексированных ресурсов, у разных поисковых систем разные. Тем не менее, многие из них стремятся, по возможности, охватывать в своей работе все пространство мировой Сети. Это универсальные системы.

Итак, работа поисковой системы обеспечивается тремя составляющими:

- Программа «робот» (спайдер). Она анализирует ресурсы и производит их индексацию.
- Индексы поисковой системы. Они формируют создаваемые поисковой системой собственные БД.
- Программа, которая в соответствии с запросом пользователя готовит ему ответ на основе анализа индексов, то есть собственных БД.

Пользователь реально имеет дело только с последней из этих трех составляющих.

Мощные поисковые системы универсального типа созданы для работы на всех основных языках мира. Каждая страна старается создать хотя бы одну собственную поисковую систему. Познакомимся с основными отечественными и зарубежными поисковыми системами.

Yandex (<http://www.yandex.ru>) – самая популярная в настоящее время отечественная поисковая система. Начала работу в 1997 г. Она содержит более 33 миллионов документов, поддерживает собственный каталог Интернет-ресурсов. Также является лучшей поисковой системой для выявления иллюстраций. Англоязычный вариант снабжен справочником ресурсов Интернет. Обладает развернутой системой формирования запроса. В частности, допускается ввод поискового предписания на естественном языке - в этом случае все необходимые расширения производятся автоматически. Более детальный запрос может быть составлен с помощью режима «Расширенный поиск» (знак +), в котором применяется система многоступенчатых меню.

Яндекс - единственная российская поисковая система, индексирующая документы в форматах PDF, DOC, RTF, SWF, PPT и XLS. Актуализация базы осуществляется еженедельно.

Rambler (<http://www.rambler.ru>) – одна из первых российских ИПС, открыта в 1996 году. В конце 2002 года была произведена коренная модернизация, после которой Rambler вновь вошел в группу лидеров сетевого поиска. В настоящее время объем индекса составляет порядка 150 миллионов документов. Для составления сложных запросов рекомендуется использовать режим «Детальный запрос», который предоставляет широкие возможности для составления поискового предписания с помощью пунктов меню.

АПОРТ (<http://www.aport.ru>). На сегодняшний день объем ее базы составляет более 20 миллионов документов. Система обладает широким спектром поисковых возможностей. АПОРТ обладает функцией встроенного переводчика, это дает пользователю возможность формулировать запросы, как на русском, так и на английском языках. Кроме того, АПОРТ имеет специальные режимы для поиска иллюстраций и аудио файлов.

Поисковая система компании Mail.ru начала работать в 2007 году. Объем индексного файла весной 2009 г. составлял более 1.5 миллиарда страниц, расположенных на русскоязычных серверах. Помимо разыскания текстов, системой осуществляется поиск иллюстраций и видеофрагментов, размещенных на специализированных "самонаполняемых" российских серверах: Фото@Mail.Ru, Flamber.Ru, 35Photo.ru, PhotoForum.ru, Видео@Mail.Ru, RuTube, Loadup, Rambler Vision и им подобных. Gogo.ru позволяет ограничивать область поиска сайтами коммерческой направленности, информационными сайтами, а также форумами и блогами. Форма "Расширенного поиска" также дает возможность ограничить разыскания определенными типами файлов (PDF, DOC, XLS, PPT), местом положения искомых слов в документе или определенным доменом.

Наиболее популярными зарубежными поисковыми системами являются Google, Alta Vista, Scirus.

Google (<http://www.google.com>) — одна из самых полных зарубежных ИПС. Объем ее базы составляет более 560 миллионов документов. Отличительной особенностью ИПС Google является технология определения степени релевантности документа путем анализа ссылок других источников на данный ресурс. Чем больше ссылок на какую-либо страницу имеется на других страницах, тем выше ее рейтинг в ИПС Google.

AltaVista (<http://www.altavista.com>) — одна из старейших поисковых систем занимает одно из первых мест по объему документов — более 350 миллионов. AltaVista позволяет осуществлять простой и расширенный поиск. «Help» позволяет даже неподготовленным пользователям правильно составлять простые и сложные запросы.

Основные виды поиска

Поиск информации является одной из составляющих человеческой деятельности. В Интернете ежедневно появляется и исчезает огромное количество ресурсов. Необходимо уметь в них ориентироваться. Чтобы эффективно выполнить поиск, нужно умело составить запрос к поисковой системе. Если ответ ищется в каталоге, то особых сложностей при работе здесь нет. Создатели каталога и модераторы озаботились созданием разделов каталога. Пользователь должен только четко определиться с предметом поиска и затем воспользоваться структурой, которая ему предложена.

В основе полнотекстового поиска лежит умение удачно составить поисковый запрос. При прочих равных условиях следует выбирать для поиска наиболее суженное, мало распространенное слово. В то же время может случиться и так, что поиск по слишком узкому термину не дает результатов. Тогда надо переходить к более широким терминам.

Если поиск по узкому термину не дает результатов, то следует постепенно расширять поисковое понятие или же переходить к смежным терминам.

Поиск по возможности лучше вести по нескольким словам, их сочетаниям, а иногда и по конкретным фразам.

При проведении поиска и оценке его результатов необходимо помнить:

- Слепое доверие к результатам, размещенным в Сети, чревато ошибками.
- Остерегайтесь анонимных сведений, т. к. их источник установить очень трудно.
- В сомнительных случаях не ленитесь проверить найденные сведения запросом в Сети отзывов и мнений по поводу найденных сведений.
- Неожиданно появляющиеся и затем быстро исчезающие ресурсы не надежны.
- В сомнительных случаях необходимо обращаться к таким проверенным источникам, как сетевые энциклопедии, справочники и специальные учебные сайты.

Избегайте «сенсационных» данных, результатов с чрезмерно большой точностью, чисто полемических, рекламных и необъективных материалов.

Приступая к поиску, вы вводите одно или несколько ключевых слов, выбираете вид поиска и нажимаете на клавишу Enter. В ответ выдается список адресов (URL) – список всех индексированных страниц, содержащих любые ключевые слова. Нередко число совпадений при таком поиске огромно. Однако если поисковая система хорошо сортирует результаты по тематике, то нужную страницу можно найти в верхней части списка. Поиск по любому слову может быть удобен в случаях, когда пользователь не уверен в ключевых словах.

Следует иметь в виду, что поисковые машины относятся к словам по-разному. Такие ИПС как «Яндекс» понимают слово во всех его грамматических формах и с учетом этого будут его искать. Им известна морфология русского языка, т. е. если в запросе написать и «человек» и «люди», то ответ будет одним и тем же.

При поиске в Интернете используются следующие логические операторы:

- AND – и (и то и то – два термина вместе);
- OR – или (или тот термин или тот);
- NOT – не (не нужен такой-то термин).

Используя значок *, можно расширить запрос до всех слов, содержащих введенную часть. Например, если ввести электротехни* то в результатах поиска окажутся страницы, содержащие электротехника, электротехнический и т. п. Знак усечения * может быть подставлен в ключ справа, слева и внутри. В Yandex можно запрашивать конкретную форму слова при поиске (исключив другие словоформы), поставив перед ним знак «!».