

Сергей Кузнецов

**УДК 681.3**  
**ББК 32.973.26-018.2**  
**К891**

# **Внутри Internet: Методы поиска информации**

**Кузнецов С.Д.**  
**К891** Внутри Internet: Методы поиска информации.  
- М.: Бук-пресс, 2006. - 224 с. -  
(Хитрости и тонкости).

**ISBN 5-8321-0177-4**

Поиск отдельных людей и организаций является часто возникающей для многих пользователей задачей. Специфика текущего момента состоит в том, что в Internet сегодня можно найти информацию не только о лицах и учреждениях, в той или иной степени причастных к Сети, но и о тех, которые не имеют к ней прямого отношения. Об этом и рассказывается в данной книге.

**УДК 681.3**  
**ББК 32.973.26-018.2**

**ISBN 5-8321-0177-4**

© Кузнецов С.Д., составление, 2006

© Бук-пресс, 2006

## Содержание

---

# Содержание

### Поиск информации

Поиск отдельных людей и организаций	3
Тестирование системы	8
Как найти информацию в Web	18
Методы поиска	43
Стратегия поиска	55
Проверка метода	59
Систематизация поиска	65
Поиск в стиле ДИСКo	101
Особенности русского языка в Internet	105
Оффлайн браузеры	113
Вопросы и ответы	130
Поиск работы через Internet	141

### Приложения

Глоссарий наиболее частых фраз и слов, на которые вы, вероятно, «наткнетесь» в Internet	167
Использованные материалы	217

# Поиск информации

## Поиск отдельных людей и организаций

Поиск отдельных людей и организаций является часто возникающей для многих пользователей задачей. Специфика текущего момента состоит в том, что в Internet сегодня можно найти информацию не только о лицах и учреждениях, в той или иной степени причастных к Сети, но и о тех, которые не имеют к ней прямого отношения. Это связано с нарастанием объемов и числа баз данных с соответствующими ресурсами, которые выкладываются на сервера свободного или ограниченного доступа и содержат сведения об огромном количестве людей, нередко даже не подозревающих о своем присутствии на них. Речь идет о широком спектре ресурсов от телефонных баз данных отдельных регионов и компаний до тех, которые содержат поименный список жителей целых государств.

Все эти службы имеют один серьезный недостаток — они не представляют собой единую кем-либо администрируемую систему, а являются лишь хаотически с точки зрения стороннего наблюдателя пополняемым набором информационных узлов.

Доступ ко многим из них можно получить через глобальные поисковые системы, такие, например, как AltaVista и Yahoo, активизировав соответствующую ссылку «**People Search**». Именно таким образом большинство пользователей и узнает о принципиальном существовании в Сети сервисов поиска людей.

Использование этих служб вполне приемлемо, однако для того, чтобы вести поисковые работы этого направления эффективно, и в случае неуспеха иметь хотя бы какие-то гарантии того, что вероятность присутствия информации о заданном объекте в Сети действительно мала, необходим более систематичный подход к этой проблеме. При этом, требуется следующее:

- владеть набором адресов наиболее солидных поисковых служб Сети;

- обладать информацией о специфике баз данных и особенностях информационно-поисковых языков соответствующих систем;
- обладать информацией об их рейтинге, особенно в случае рассмотрения идентичных по многим показателям служб.

Получение набора Internet-адресов поисковых служб людей и организаций является достаточно легко разрешимой задачей. Однако в случае, если перечень пунктов в таком списке переваливает за сотню или даже десяток, к реальному успеху в решении поисковой задачи может привести только свободное ориентирование в возможностях этих сервисов. Главной задачей при этом становится разумно спланировать последовательность использования поисковых служб. А именно это, по сути, и является самым сложным.

Несмотря на существующее многообразие сервисов, базовыми сегодня можно с уверенностью считать службы, размещаемые на Web-серверах. Исходя из многих источников наибольшей

популярностью в Сети среди пользователей Европы и Северной Америки пользуется служба поиска адресов электронной почты Four11. Расположенная в портале Yahoo служба имеет характерный для каталогов этой системы простой и удобный интерфейс, который реализован в виде шаблонов простого и расширенного поиска. Four11 является глобальной системой, но по понятным историческим причинам, как и в большинстве подобных служб преобладает информация о людях, проживающих в США и Канаде, затем идет Европа, затем третий мир и, увы, Россия.

Несмотря на простоту шаблона остаются некоторые неясные на первый взгляд места, а также неточности. Во-первых, у службы два отделения: поиск адресов электронной почты (E-mail Search) и поиск телефонов (Telephone Search).

Обсудим для примера первую возможность. Во-первых, достаточно уникальна ситуация, что система может искать адресата только по собственно имени (First Name), и указание фамилии (Last Name) не является обязательным, то

есть, у этого поля нет пометки «required». Этот факт можно использовать для поиска людей с очень редкими именами. Видимо, шаблон разработчикам системы кажется настолько простым, что размещение help-ссылки о его использовании где-нибудь поблизости представляется им необязательным. Тем не менее, как вы думаете, можно ли использовать в этой службе при построении запроса маску «\*», во всех ли полях это можно делать, что собственно следует при необходимости ввести в поле «Domain», можно ли попытаться разыскать в этой системе людей из Петербурга, можно ли искать адресата, если вы располагаете только частью его e-mail-адреса? Добавьте к этому свои собственные вопросы.

Сказать, что все это уже заранее нам известно, и мы можем поделиться информацией с читателем — это значит не сказать ничего, поскольку ему придется столкнуться с десятками других, часто, еще гораздо более проблемных интерфейсов. Нам бы хотелось показать, как можно получить эту информацию методом тестирования.

### Тестирование системы

Можно ли использовать в этой службе при построении запроса маску (\*), что крайне удобно, если известна только часть имени или фамилии?

Вводим в поле **Last Name** любую распространенную фамилию, например, **Smith**. Кликаем на шаблоне клавишу **Search**. Отклик получен, система работает!

Вводим в поле **Last Name** часть фамилии вместе с маской: **Sm\***. Предполагается, что не только ранее найденные **Smith** должны откликнуться снова, но и **Sme**, **Smile** и т.д. Отклик получен, маска работает. Ту же операцию можно проделать для поля **First Name**. Работает. Рассмотрение того, работает ли маска в поле **Domain**, ненадолго отложим.

На **Four11**, маска, как мы убедились, применяется в виде звездочки. Следует иметь в виду, что во многих системах она также поддерживается, но не явно, а отрабатывает автоматически тогда, когда в запрос вводится только часть слова.

Что, собственно, следует при необходимости ввести в поле «**Domain**»?

Разумеется домен, в котором, как вы предполагаете, находится почтовый сервер вашего адресата. Если адрес e-mail выглядит как `tala@server.citmgu.ru`, то все, что стоит после значка @ является доменами: **ru** — домен верхнего уровня, **citmgu** — домен второго уровня, **server** — домен третьего уровня. Можно использовать сразу несколько доменов разного уровня, разделенных точками. На `yahoo.com` большое количество бесплатных почтовых ящиков. Можно ли узнать всех адресатов, размещаемых на Yahoo, т. е. с элементом адреса `@yahoo.com`?

Вводим «`yahoo.com`» в поле **Domain**, все остальные поля оставляя пустыми. Работает, но в списке отклика появляются не только адреса типа `jdalva@yahoo.com`, но и `ziadstar@yahoo.com` и даже `asangulo@hotmail.com`. Очевидно, система реагирует некорректно, и использовать такие запросы следует с осторожностью, либо произвести дополнительный анализ.

Применим только что найденный нами и, естественно, зарегистрированный в службе Four11 адрес `ziadstar@yahoo.com` для очередного тестового запроса, дерзко введя его в поле

**Domain.** В отклике появляется целый список адресов, в основном с домена `yahoo.com`, при этом адрес `ziadstar@yahoo.com` в нем также присутствует, то есть, искать людей только по адресу e-mail без дополнительной информации система корректно не позволяет, что, в общем, неудивительно, поскольку мы ввели в поле **Domain** не только домен, но и имя почтового ящика.

Проверим, можно ли использовать только домен верхнего уровня для локализации почтовых ящиков России. Вводим отдельно в поле **Domain** запросы с доменом `ru`, а затем отдельно с доменом `su`. Если известен широко распространенный домен города, например, `spb` (Санкт-Петербург), то можно ввести `spb` или `spb.ru`, однако должно быть понятно, что далеко не все пользователи Санкт-Петербурга имеют почтовые ящики на серверах в этом домене.

Проверим, можно ли использовать маску в поле **Domain**. Запрос в виде **yahoo\*.com** должен показать нам уже найденные ранее адреса Yahoo, в существовании которых в системе теперь

мы уверены. Нулевой оклик, значит маска здесь не применима.

При этом, если бы шаблон требовал обязательного присутствия в запросе элементов фамилии, то для сохранения корректности запроса можно было бы использовать вместе со всеми вышеуказанными запросами и поле **Last Name**, куда ввести, например, «a\*», предполагая, что на любом почтовом сервере с большой вероятностью найдется пользователь с фамилией, начинающейся на букву «А».

Тестирование, аналогичное тому, которое мы выполнили выше, позволяет ответить на подавляющее количество вопросов, относительно любой поисковой службы, в том числе и выявить часто неприятные для ее разработчиков «недокументированные» особенности.

В целом же, для того, чтобы получить исчерпывающее представление о работе поисковой службы, целесообразно начать с анализа не простого, а расширенного поискового шаблона, на который можно выйти по ссылке **Advanced** из шаблона простого поиска. Не останавливаясь так же подробно на его

деталях, рассмотрим только дополнительные возможности. Прежде всего, приятным сюрпризом является то, что поля **State/Province** и **Country** в расширенном шаблоне даны как ссылки на страницы, где подробно объяснено, что, собственно, следует в них вводить, а именно, никак не название государства или региона, а специальный аббревиатурный код, например, «ru» для Российской Федерации в поле **Country**, в данном случае код совпадает с именем домена. К сожалению, в других поисковых системах, эти обозначения, знакомые всем северо-американцам, особенно когда речь идет об отдельных штатах и некоторых административных территориях, подробно не расшифровываются.

Привлекательным, особенно для русскоязычного пользователя, кажется использование в системе сервиса **SmartNames**, позволяющего перебрать в запросе все допустимые формы имени на основе введенного, например, Bob = Robert, или Rick=Richard. Нужно только не забыть поставить в шаблоне галочку.

Для неискушенных пользователей следует заметить, что источником

пополнения базы данных адресов электронной почты поисковых служб является прежде всего система телеконференций, затем идет персональная регистрация заинтересованных лиц по собственному желанию. При этом значительная часть информации вводится людьми самостоятельно, ее достоверность никем не контролируется. Поэтому на запрос на имя президента компании Microsoft Билла Гейтса (Bill Gates), вы сможете найти большей частью не однофамильцев этого известного в компьютерном мире человека, а шутников, склонных к мании величия.

Ссылка «**People Search**» — не единственная, которая может встречаться на Web-порталах и иметь отношение к поиску людей и организаций. Это могут быть ссылки «White pages» и даже «Yellow Pages». Часто, когда речь идет о таком поиске приходится сталкиваться с понятиями так называемого white- и yellow-поиска. Под white-поиском обычно понимается поиск адреса конкретного адресата по его заданному имени, то есть, человека — по фамилии, организации — по названию. При этом объект поиска

заранее известен, есть уверенность в его существовании и используемое для поиска имя до определенной степени уникально, так что вероятность успеха даже на первом шаге поисковой процедуры оказывается значительной. White pages (Белые страницы) — это фактически используемые в обычных телефонных справочниках перечни людей или организаций по алфавиту, и именно такую ссылку следует разыскивать на узле.

Если поиск по Белым страницам не приводит к немедленному успеху, то возникает предположение об ошибке в исходных данных и необходимость обратиться к yellow-поиску. Yellow-поиск предполагает розыск адресатов не только и не столько по их собственному имени, сколько по другим признакам, которые также тем или иным образом классифицируются. При этом исходное имя адресата является либо неполным, либо неоднозначным и позволяет лишь отнести обозначаемого им абонента к какой-то группе, классу и т.д. На практике информационные системы Yellow Pages (Желтые страницы) фактически сразу включают в себя и White Pages — каждый адресат идентифицирован его телефоном и



почтовым адресом. Кроме того, некоторые Yellow Pages позволяют искать также и по алфавитному индексу (white-поиск). С другой стороны, White pages также содержат элементы yellow-поиска, поскольку кроме задания собственного имени они обычно позволяют указать название города, штата и другие, сужающие охват поиска, данные (что необходимо, например, в случае однофамильцев). Вероятно, именно с этим и связано то, что многие телефонные справочники, представленные в Сети и выполняющие, фактически white-поиск, называют себя Yellow pages. Под желтыми страницами понимают также и обширные каталоги ресурсов Internet, организованные по тематическому признаку.

Продолжая наш разговор о системах поиска людей и организаций, перечислим другие наиболее значительные службы Сети. Итак, начать все-таки предпочтительней со службы Four11 и затем воспользоваться ниже перечисленными: детально познакомиться с их шаблонами, особенно расширенного поиска, если таковой присутствует, учесть специфику своей задачи, при

необходимости протестировать эти системы и выстроить приемлемую для себя схему их использования на будущее.

- InfoSpace
- Switchboard
- WhoWhere?
- Worldpages
- BigBook
- BigFoot
- BigYellow
- LookSmart
- PeopleFinder

При поиске людей в Сети нередко приходится проявлять талант подлинного сыщика. Естественен поиск человека по его следам — по научным работам, журнальным статьям, причастности к деятельности той или иной организации. Не сразу может прийти в голову ввести в поисковый шаблон, например, в AltaVista запрос в виде фамилии. При этом, например, могут откликнуться документы с оглавлением соответствующих журналов, в которых представлен данный автор и в результате уже через несколько минут вы будете располагать адресом редакции,

через которую можно попробовать получить дополнительную информацию.

При розыске организаций, название которых состоит из одного слова или укладывается в односложную аббревиатуру, разумно использовать поиск web-страниц, в URL (сетевой адрес) которых входит данный термин, поскольку, как известно, односложное название компании, обычно присутствует в доменном имени сервера или названии каталога на нем. В этом случае нужна поисковая машина, поддерживающая поиск по URL, такая как AltaVista или Рамблер. Если название компании двусложное, целесообразно искать в поисковых системах те страницы, где это название входит в заголовок (title). В URL такое название тоже может звучать, часто как простое слияние двух слов, слияние слов со вставкой дефиса между ними или слияние с урезанием конечных частей одного из входящих в название слов или обоих. Достаточно эффективным может быть поиск в локальных базах данных отдельных газет и журналов, но обслуживание такой информацией все чаще становится платным.

Хочется отметить постоянно растущую популярность службы коммуникации пользователей в реальном времени ICQ, поисковая система которой сегодня становится все более развитой и в русскоязычной Сети наряду с сервером Эрос (база данных адресов электронной почты русскоговорящих пользователей) приводит к успеху намного быстрее других сервисов.

### Как найти информацию в Web

Поиск информации на отдельном Web-узле — это та задача, которую приходится решать каждому пользователю Internet. Если вы связываетесь с Сетью через модем, то очевидно, что, чем больше вы тратите времени на поиски, тем дороже стоит получаемая информация. В этом случае ее изучение прямо в Сети становится непозволительной роскошью. Многие предпочитают этому ознакомительный просмотр и быстрое копирование необходимых материалов, а затем их более предметное изучение в автономном режиме. Перенос содержимого сервера целиком на локальный компьютер, а для этого

существуют специальные средства, в большинстве случаев также менее предпочтителен, чем выборочное копирование. Следовательно, самым полезным навыком становится умение быстро разобраться в структуре узла и способах навигации (т. е. путей перемещения с одной web-страницы узла на другую). Даже при подключении по выделенной линии, если вы решаете какую-либо поисковую задачу, и требуется просмотреть десятки узлов из списка отклика поисковой машины, вопрос о скорости освоения информации остается одним из определяющих. Можно ли здесь помочь пользователю? Мы предлагаем рассмотреть два важных для разрешения этой проблемы вопроса.

- Типовые структуры размещения информации на Web-узле и возможности навигации
- Логика «третьего» уровня и приемы применения автоматических поисковых средств.

### **Типовые структуры размещения информации на Web-узле и возможности навигации**

Структура современного Web-узла может быть различной и тесно связана со способами навигации по его страницам, которых не так уж много, а именно:

а) путем начального задания адреса вручную в строке URL или выбора документа из списка истории браузера (программы просмотра web-страниц), если таковая уже накоплена (мы будем использовать только два основных браузера, или программы: MS Internet Explorer — сокращенно IE, и Netscape Navigator — NN);

б) по гипертекстовым ссылкам;

в) по каталогам узла с помощью обрезания строки ранее введенного адреса (URL)

Каждый раз вы теряете время, если делаете неверный выбор.

Вариант а) знаком всем и каждому по клавишам «Назад» и «Вперед» на верхней панели браузеров. Не все, тем не

менее, знают что у последних поколений этих программ они несколько поумнели: при удержании на клавише нажатой кнопки мыши можно не только перейти к предыдущей или последующей странице, но и выбрать нужную из списка истории посещений за последний сеанс работы.

Вариант б) является самым неоднозначным. Именно он требует хорошего знания структуры информационного узла.

Вариант в) чаще применяется более опытными пользователями. Во многих случаях он удачно работает тогда, когда вы приходите на узел по ссылке, но указанный в ней файл не найден. Браузер выдает сообщение типа «File Not Found» или, если узел более ухожен, обрабатывает специальная программа-скрипт, приносящая извинения разработчика сайта (вместо слова «узел» часто также употребляется термин «сайт», от английского «site») и предлагающая пользователю другие возможности найти информацию самостоятельно.

Пусть вы обратились к ресурсу по URL:  
<http://server.citmgu.ru/internet/courses/search1.html>

Искомый файл search1.html должен находиться на сервере с доменным именем server.citmgu.ru в каталоге courses, который в свою очередь является подкаталогом каталога Internet, лежащего в корне сервера. Предположим, что ресурс не найден. После чтения в окне браузера надписи типа «File Not Found» многие возвращаются клавишей возврата браузера назад к предыдущей странице, полагая что информация недоступна, и это напрасно. Сервер работает, — это главное, а сам документ мог быть перемещен, переименован или заменен аналогичным по содержанию. Выход прост: попытаться выйти на головную (домашнюю) страницу узла и разыскать его самостоятельно или с помощью локальной поисковой системы. Для этого в URL следует откусить часть адреса с правой стороны и нажать на «Ввод», то есть, ввести:

<http://server.citmgu.ru/internet/courses>

а в случае повторной неудачи:

<http://server.citmgu.ru/internet>  
и так вплоть до собственно доменного имени сервера:

<http://server.citmgu.ru>.

При таких обращениях к каталогам сервера, о которых мы узнаем по длинному адресу, серверная программа может отобразить как список файлов указанного каталога, так и конкретную web-страницу, относящуюся к этому каталогу и предназначенную для загрузки по умолчанию. При этом переходить сразу к имени сервера, минуя промежуточные обращения, не всегда целесообразно, поскольку, если на сервере, скажем, размещается несколько домашних страниц отдельных пользователей или компаний, то всякая логическая связь между ними, вполне вероятно, отсутствует. Корневой каталог сервера может при этом наполнять еще одна более крупная компания или организация, которая не имеет никакого отношения к документам, лежащим в глубине дерева каталогов и никаких ссылок на эти документы. Следовательно, при отсутствии дополнительной информации лучше идти мелкими шагами, последовательно поднимаясь от каталога к каталогу вверх к корню сервера.

Так как основные перемещения в Web-пространстве приходится все-таки делать по гипертекстовым ссылкам (случай б)), то и выбор, который стоит перед пользователем, это либо продолжить движение по ним, либо ознакомиться со структурой каталогов сервера путем усечения адреса (URL-навигации), либо использовать клавиши браузера для возврата к уже просмотренной странице.

Фактически речь идет о том, что одновременно существует два логических уровня организации информации — путем размещения ее определенным образом в иерархии каталогов, а также путем ее связывания с помощью гипертекстовых ссылок.

На качественно скроенном информационном узле, как правило, оба уровня несут смысловую нагрузку, то есть, например, перейти к просмотру предлагаемых учебных курсов по Internet можно как по ссылке «Курсы об Internet» на домашней странице узла, так и при обращении по URL:

<http://server.citmgu.ru/internet/courses>

Вообще говоря, различия в базовых структурах Web-узлов во многом строятся на существовании такой двухуровневой системы навигации.

В некоторых источниках принято различать плоскую, линейную, древовидную и комбинированную структуры. Попробуем взглянуть на них глазами не Web-мастера, а пользователя-навигатора.

Плоская структура предполагает, что в центр узла ставится головной документ, с которого имеются ссылки на все остальные документы, те в свою очередь также могут ссылаться друг на друга и на головную страницу. Ясно, что точкой отсчета для очередной процедуры поиска-просмотра в этом случае является головной документ и требуется определить наиболее быстрый доступ к нему с любой страницы, например, по специальной ссылке на каждой странице, по закладке в браузере или по URL.

Структура каталогов стоит здесь на втором плане.

Линейная, или последовательная структура связывает ряд документов, в

каждом из которых предусмотрены только ссылки вперед- назад. В чистом виде она встречается редко, и в этом случае перескочить сразу через несколько пунктов — но только назад, позволяет прямой выбор документа из списка просмотренных страниц из истории браузера.

Следующая — древовидная структура гипертекстовых ссылок узла полностью повторяет логическую организацию его каталогов, хотя это, разумеется, не означает, что каталоги и ссылки будут иметь совершенно одинаковые названия. При этом эффективность URL-навигации заметно возрастает.

Комбинированная структура является самой распространенной и предполагает совместное использование упомянутых выше структур. Если переходы по гипертекстовым ссылкам после двух-трех первых десятков просмотренных узлов ни у кого не вызывают затруднений, то URL-навигация редко попадает в поле зрения пользователя, хотя при наличии древовидной структуры документов она может превосходить по эффективности все другие виды перемещений (например,

каталоги Yahoo). На Yahoo, если вы получаете отклик на поисковый запрос, а затем по гиперссылке попадаете в нужный раздел, усечение адреса позволяет быстро переходить к более высоким уровням.

Для того, чтобы воспользоваться URL-навигацией на незнакомом узле, необходимо обратить внимание на то, существует ли там какое-либо соответствие названий каталогов тематическим разделам, заявленным в ссылках на домашней странице, то есть, при очередном шаге по ссылке следует отследить изменение URL в адресном поле браузера.

Безусловно, такие исследования оказываются практичными далеко не всегда, однако готовность воспринять информационный узел в виде узнаваемой структуры помогает сэкономить десятки секунд при работе с наборами документов, что в итоге выливается в серьезное повышение производительности труда.

### **Логика «третьего» уровня и приемы применения автоматических поисковых средств**

Следующие наши наблюдения будут связаны с «третьим» уровнем логики — дизайном узла и типовыми названиями разделов, в которых размещается информация заданного типа, а также со средствами автоматического поиска, применяемыми на отдельной web-странице и на узле. Все эти вопросы тесно связаны между собой.

Домашняя страница узла в большинстве случаев содержит в левой или верхней части экрана основное меню, с которого идут ссылки на информационные разделы. Желание разработчика оставлять такое меню видимым из любой точки нередко приводит к использованию фреймов — специальных прямоугольных кадров на экране, в каждом из которых может отображаться свой документ. При переходе по соответствующей ссылке новый документ может загружаться в отдельный фрейм, а меню при этом остается на экране нетронутым в своем

фрейме. Если материал во фрейме не помещается целиком на экране, то по умолчанию с правой стороны от него формируется полоса прокрутки. В этом случае фрейм воспринимается как отдельное окно. Всякий раз полезно обратить внимание на размер движка этой полосы, т.к. чем он меньше, тем большая доля информации осталась за пределами экрана, а исходя из этого можно принять то или иное решение об очередном шаге.

Если информации хватает места, то полоски сбоку не возникает и выделить фрейм из экрана становится трудно. При этом у многих пользователей появляются проблемы с поиском и сохранением информации на локальный компьютер, поскольку для этого приходится использовать простые, но все-таки специальные возможности браузеров, а не те, что применяются при отсутствии фреймов.

Итак, на сегодняшний день широко распространены два вида домашних страниц, а именно презентационная и информационная.

**Презентационная** страница, как правило, насыщена графикой, имеет

небольшое количество ссылок и помещается в один экран. Ее задача — представить компанию или учреждение в Internet, которые хотят заявить о сфере своей деятельности и указать возможности контактов.

**Информационная** страница, напротив, сводит к минимуму присутствие графических объектов и дает максимальное текстовое наполнение, что увеличивает ее объем иногда до двух-трех и более экранов.

Оба типа страниц могут предвзраться еще одной, так называемой страницей-заставкой, обычно, почти чисто графической, которая загружается перед домашней страницей. Ранее она часто использовалась для выбора кодировки на русскоязычных страницах, однако теперь, когда эта проблема решается браузером автоматически, необходимость в ней отпала, и она просто становится дополнительным препятствием на пути к информации. Задача пользователя — быстро пройти через нее, убедившись в ее минимальной смысловой нагрузке.

Один из самых эффективных способов ускорения работы с web-



страницей — это активное использование средств автоматического поиска, конечно, если вы знаете, что вам нужно. Такой подход особенно практичен для многоэкранных страниц с информационными узлами, когда визуальное ознакомление с материалом становится слишком трудоемким.

Поиск на странице можно произвести по терминам, введенным в специальный поисковый шаблон, который активизируется в браузерах клавишами **CTRL-F** или через главное меню «**Правка ⇨ Поиск на этой странице**» (MS IE) или «**Edit ⇨ Find in Page**» (NN).

Мало кому известно, что работа этих шаблонов несколько отличается в разных браузерах и имеет некоторые особенности.

Для того чтобы поиск по странице был успешным, к сожалению, предлагаемой справочной системой браузера инструкции, недостаточно. Сама инструкция выглядит достаточно прозаично.

### Поиск текста на активной странице

- В меню **Правка** выберите пункт **Найти на этой странице**.
- Введите текст, который вы хотите найти.
- Задайте условия поиска.
- Нажмите кнопку **Поиск далее**.

Тем не менее вряд ли у новичка не возникнет вопросов в связи с использованием этого шаблона, значимость которого по понятным причинам крайне велика — вы просто можете пройти мимо значимой информации. В чем же проблема? Необходимо помнить о следующем.

1. Поиск всякий раз проводится вверх или вниз по странице в зависимости от указания направления в шаблоне, начиная с начала (если вниз), или с конца документа (если вверх), независимо от того, какая часть страницы отображается на экране на момент начала поиска. Однако это справедливо только в том случае, если на странице нет выделенных областей (на кстати это указано в справке к NN).

2. Допустимо введение в шаблон не только единичного термина, но и фразы, что делается одной строкой без использования специального синтаксиса. Специальная пометка в шаблоне позволяет искать с учетом регистра символов.

3. Найденное слово или фраза выделяются в тексте, и происходит автоматическое перемещение к их местоположению, однако выделенное поле не всегда можно наблюдать. Сам шаблон, остающийся на экране во время поиска может загоразивать его (в этом случае надо сместить шаблон мышью), кроме того оно может быть отображено на экране не целиком, а только его кромкой (следует использовать полосу прокрутки). Другими словами, если нет сообщения о том, что поиск на странице завершен, то обязательно необходимо обнаружить выделенное в тексте поле, иначе информация будет потеряна.

4. Если при старте поиска уже есть выделенная область текста, то поиск начинается именно с нее в заданном в шаблоне направлении, само содержимое выделенного поля участия в поиске уже не принимает, также как и оставшаяся часть

страницы. Отметим, что всякий раз, когда поисковая процедура закончена, на странице остается выделенная область текста, соответствующая последнему совпадению. Если необходимо выполнить поиск с уже новыми терминами, то следует сначала снять уже существующее выделение кликом мыши в любой точке текста, иначе в новом поиске будет участвовать только часть страницы вверх или вниз от выделенной области в зависимости от направления, заданного в шаблоне.

5. Надписи, выполненные в графике, не откликаются на поисковые запросы.

6. Поисковый шаблон браузера NN не имеет возможности задать поиск термина как целого слова, как это позволяет сделать IE. По умолчанию, если не выставлен нужный флаг, введенный в шаблон термин, обнаруживает совпадения со всеми словами, в которые входит указанный в шаблоне фрагмент, то есть, для термина «поиск» будут найдены совпадения со словами «поиска», «поисковый» и т.п. Таким образом, поиск по странице в Internet Explorer имеет дополнительную полезную функцию.

7. Если при проведении поиска экран разделен на фреймы, то обсуждаемые нами браузеры работают по-разному. Браузер Internet Explorer объединяет документы из разных фреймов в единое текстовое поле, в котором и производит последовательный поиск после активизации пункта меню «**Найти на этой странице**», то есть, автоматически переходит из фрейма во фрейм. Браузер NN имеет специальную возможность поиска внутри заданного фрейма. При этом не следует забывать о влиянии выделенных областей на работу программ.

В целом для работы с фреймами NN кажется несколько удобнее. Так, например, узнать в нем о присутствии фреймов, если они не просматриваются явно, можно, заглянув в меню **File**, где сразу же становятся доступными функции работы с фреймами (frames), также можно получить более предметную информацию о фреймовой структуре и при просмотре источника документа, однако для этого нужно знать язык HTML. Для операции с документом во фрейме бывает нужно активизировать соответствующий фрейм щелчком (кликом) левой кнопки мыши по любой

неактивной точке фрейма (то есть не по гипертекстовой ссылке, что спровоцирует немедленный переход). Как же это сделать, если фрейм визуально иногда не отличим от остальной части экрана? Просто следует выполнить клик где-то в области того самого материала, который вас интересует.

Основное меню узла, которое обычно расположено слева или сверху экрана, довольно часто содержит надписи, которые выполнены в виде графики, то есть, представляют собой рисованные объекты, а не алфавитно-цифровой набор символов, введенных с клавиатуры.

Поэтому, как было указано выше, нельзя локализовать термины из графического меню с помощью функции поиска по странице.

Однако, чтобы не лишать пользователя такой возможности, часто разработчики дублируют заголовки графического меню в самом низу домашней страницы в символьном режиме.

Обратиться к финальным ссылкам, дублирующим графическое меню, бывает

полезно и тогда, когда само меню выглядит слишком вычурно и неудобно для чтения.

Глоссарий терминов, соответствующих определенному типу информации, естественно, может отличаться в зависимости от профиля изучаемых ресурсов, поэтому пользователь, постоянно решающий поисковые задачи, должен следить за обогащением своего словарного запаса специфичной сетевой лексикой.

Неординарность этой лексики, базой для которой является естественный язык, связана с тем, что в Сети доминируют авторы, имеющие техническое образование или, по крайней мере, «технократический» образ мышления, которым свойственны лаконичность, колорит и, увы, порочное тяготение к сленгу, во многих случаях совершенно неоправданному и создающему мнимый ореол элитарности вокруг тех, кто им пользуется. Разумеется, чем ближе профиль материала к техническим проблемам и чем фамильярней изложение, тем выше доля сленговой лексики. Так, например, если сайт предназначен для

представления программного обеспечения, то раздел, содержащий новые поступления, скорее будет называться не «Новые программы», а «Свежий софт».

На многих серверах предусмотрена страница, которая предлагает еще более детальное изложение его содержания, чем основное меню. Такая страница называется «Карта сервера» («Site map»). Если ссылка на нее присутствует на домашней странице и приготовлена она не в виде графики, а в виде обычного текста, то быстро найти ее можно с помощью фразы-запроса «Карта сервера» («Site map»), или иногда просто «карта» («map»). Если ссылка на нее не текстовая, а графическая, то располагается она обычно либо в начале, либо в конце страницы, поэтому в случае нулевого отклика на текстовый запрос, следует просмотреть именно начало и конец страницы, а не приступать к немедленному чтению.

Аналогично используют функцию поиска по странице для того, чтобы найти ссылку на локальную поисковую машину, если она организована разработчиком узла. Тогда после нажатия **CTRL-F** следует ввести в шаблон слово «поиск» («search»),

и ссылка будет найдена в течении секунды.

Для более специализированных узлов целесообразно выработать собственную тактику выбора значимых терминов. Еще раз хочется подчеркнуть, что важно не просто знать как то, что вы разыскиваете, называется по-русски или по-английски, а как это называется в Internet.

Хороший принцип, стараться не делать лишних движений, — предполагает продуманность каждого шага поиска и навигации.

Отдельные слова стоит сказать о специальных программах, которые могут быть внедрены прямо в загружаемую вами страницу и написаны на языках Java, JavaScript, VBScript и других. Обычно при старте таких программ в строке состояния браузера внизу экрана возникает сообщение типа «Starting Java...». Об этом должен знать пользователь, поскольку с помощью таких программных средств нередко задается основное меню узла, что, вообще говоря, является не самым лучшим тоном.

Остановимся лишь на одном аспекте меню такого типа, которое обескуражило не одного новичка. Представьте, что зайдя на сайт крупной компании вы не видите даже намека на то, чтобы познакомиться с оглавлением узла. Страница почти пустая! После минутного размышления вы случайно проводите мышью по ничем не выделенному полю экрану, даже не нажимая ее кнопки, и вдруг — вот оно, меню с названиями разделов, плавающими по экрану. Мораль очень проста: указанные программы могут активизироваться по некоторому событию, происходящему на экране, например, по движению мыши (даже без щелчка ее кнопки), так что на многих «молчаливых» поначалу страницах не лишено смысла исследовать поля мышечувствительности, заглянув в каждый уголок.

Другой неприятностью, которую они могут доставить, является возможность автоматического открытия еще одного или нескольких окон браузера, например, в случае, клика по какой-либо гиперссылке, когда документ начинает загружаться в новое окно. Его появление может остаться незаметным, если не обращать внимание на панель задач, если вы работаете в

Windows 95/98/NT, которая немедленно реагирует на открытие нового окна. Его пугающей многих особенностью является неактивность клавиши «Назад». Но это вполне объяснимо, т. к. у новорожденного окна еще нет истории, которая сохраняется по-прежнему в спрятанном теперь родительском окне. Не следует допускать большого количества одновременно открытых окон, поскольку они быстро поглощают ресурсы оперативной памяти системы и приводят к некорректной работе браузеров, вплоть до ошибок, делающих необходимой перезагрузку системы.

Известно, что при обращении к Web-серверу по протоколу http, если загрузка происходит слишком долго, можно попытаться клавишей браузера «Останов» или «Stop» прекратить загрузку, а затем снова возобновить ее и добиться от сервера более быстрого обслуживания. Если до остановки загрузки часть информации с узла отобразилась, все гиперссылки в этом случае оказываются работающими. Если вы не дождалась окончания загрузки и нажали на появившуюся гиперссылку, то текущая загрузка прерывается автоматически, но

только в случае если новый документ открывается в том же окне. Если же он открывается в новом, то происходит одновременная загрузка двух документов: нового и старого, что замедляет передачу каждого из них. Поэтому если старый документ перестал вам быть интересен, необходимо нажать на «Останов» («Stop») до активизации очередного перехода.

Еще одно замечание сделаем относительно возможности еще до нажатия на гиперссылку отследить адрес (URL), по которому она осуществит переход. Когда указатель мыши встает на ссылку (без нажатия), то в строке состояния браузера появляется соответствующий адрес. Эту информацию можно использовать для предварительной оценки целесообразности такого перехода, она также полезна и в случае применения разработчиком специальной графической карты гипертекстовых ссылок (UsemapClient Side), когда отдельные фрагменты сомкнутой воедино картинки, могут являться ссылками на различные ресурсы.

### Методы поиска

На первых курсах института меня всегда удивляло умение преподавателя безошибочно находить ошибки в курсовой работе. Как бы вы ни пытались запутать решение и подогнать ответ, седому профессору достаточно было лишь заглянуть на первую и последнюю страницы проекта, и он возвращал работу нерадивому студенту. Только на старших курсах мы научились обходить препоны. Секрет оказался прост: нужно было подогнать не только ответ под условие, но и условие под ответ. Это знает любой старшекурсник. Попробуем употребить приобретенные в институте умения для поиска информации в сети Internet. Подгоним запрос под результат. Для этого возьмем документ и, зная, как поисковая система собирает его обрабатывать, составим запрос. Вот тут-то и возникает маленькая загвоздка. Необходимо хотя бы в общих чертах знать, как система функционирует (наверное, именно на этом парадоксе построено все высшее образование). Ничего не поделаешь, придется разбираться с основами работы поисковой системы.

### Автоматический анализ текстов

Оказывается, все созданные человеком тексты построены по единым правилам! Никому не удастся обойти их. Какой бы язык ни использовался, кто бы ни писал — классик или графоман, — внутренняя структура текста останется неизменной. Она описывается законами Зипфа (G.K. Zipf). Зипф предположил, что природная лень человеческая (впрочем, это свойство любого живого существа) ведет к тому, что слова с большим количеством букв встречаются в тексте реже коротких слов. Основываясь на этом постулате, Зипф вывел два универсальных закона:

#### Первый закон Зипфа «ранг — частота»

Выберем любое слово и посчитаем, сколько раз оно встречается в тексте. Эта величина называется частота вхождения слова. Измерим частоту каждого слова текста. Некоторые слова будут иметь одинаковую частоту, то есть входить в текст равное количество раз. Сгруппируем их, взяв только одно значение из каждой группы. Расположим частоты по мере их убывания и пронумеруем. Порядковый

номер частоты называется ранг частоты. Так, наиболее часто встречающиеся слова будут иметь ранг 1, следующие за ними — 2 и т.д. Ткнем наугад в страницу и определим вероятность встретить слово, на которое пал выбор. Вероятность будет равна отношению частоты вхождения этого слова к общему числу слов в тексте.

$$\text{Вероятность} = \frac{\text{Частота вхождения слова}}{\text{Число слов}}$$

Зипф обнаружил интересную закономерность. Оказывается, если умножить вероятность обнаружения слова в тексте на ранг частоты, то получившаяся величина (C) приблизительно постоянна!

$$C = \frac{\text{Частота вхождения слова} \times \text{Ранг частоты}}{\text{Число слов}}$$

Если мы немного преобразуем формулу, а потом заглянем в справочник по математике, то увидим, что это функция типа  $y=k/x$  и ее график — равнобочная гиперболой. Следовательно, по первому закону Зипфа, если самое распространенное слово встречается в тексте, например, 100 раз, то следующее по частоте слово вряд ли встретится 99 раз. Частота вхождения второго по популярности слова, с высокой долей

вероятности, окажется на уровне 50. (Разумеется, вы должны понимать, что в статистике ничего абсолютно точного нет: 50, 52 — не так уж и важно.)

Значение константы в разных языках различно, но внутри одной языковой группы остается неизменно, какой бы текст мы ни взяли. Так, например, для английских текстов константа Зипфа равна приблизительно 0,1. Интересно, как выглядят с точки зрения законов Зипфа русские тексты? Они не исключение. Анализ хранящихся в моем компьютере файлов с русскими текстами убедил, что закон безупречен и тут. Для русского языка коэффициент Зипфа получился равным 0,06-0,07. Хотя эти исследования не претендуют на полноту, универсальность законов Зипфа позволяет предположить, что полученные данные вполне достоверны.

### **Второй закон Зипфа «количество — частота»**

Рассматривая первый закон, мы отмахнулись от факта, что разные слова входят в текст с одинаковой частотой. Зипф установил, что частота и количество слов, входящих в текст с этой частотой,



тоже связаны между собой. Если построить график, отложив по одной оси (оси X) частоту вхождения слова, а по другой (оси Y) — количество слов в данной частоте, то получившаяся кривая будет сохранять свои параметры для всех без исключения созданных человеком текстов! Как и в предыдущем случае, это утверждение верно в пределах одного языка. Однако и межъязыковые различия невелики.

На каком бы языке текст ни был написан, форма кривой Зипфа останется неизменной. Могут немного отличаться лишь коэффициенты, отвечающие за наклон кривой (в логарифмическом масштабе, за исключением нескольких начальных точек, график — прямая линия).

Законы Зипфа универсальны. В принципе, они применимы не только к текстам. В аналогичную форму выливается, например, зависимость количества городов от числа проживающих в них жителей. Характеристики популярности узлов в сети Internet — тоже отвечают законам Зипфа. Не исключено, что в законах отражается

«человеческое» происхождение объекта. Так, например, ученые давно бьются над расшифровкой манускриптов Войнича. Никто не знает, на каком языке написаны тексты и тексты ли это вообще. Однако исследование манускриптов на соответствие законам Зипфа доказало: это созданные человеком тексты. Графики для манускриптов Войнича точно повторили графики для текстов на известных языках.

Что дают нам законы Зипфа? Как с их помощью извлечь слова, отражающие смысл текста? Воспользуемся первым законом Зипфа и рассмотрим график зависимости ранга от частоты. Как уже упоминалось, его форма всегда одинакова. Исследования показывают, что наиболее значимые слова лежат в средней части диаграммы. Это и понятно. Слова, которые попадают слишком часто, в основном оказываются предлогами, местоимениями, в английском — артиклями. Редко встречающиеся слова тоже, в большинстве случаев, не имеют решающего смыслового значения.

От того, как будет выставлен диапазон значимых слов, зависит многое. Поставь широко — нужные термины

потонут в море вспомогательных слов; установи узкий диапазон — потеряешь смысловые термины. Каждая поисковая система решает проблему по-своему, руководствуясь общим объемом текста, специальными словарями.

### Весовые коэффициенты

До сих пор мы рассматривали отдельно взятый документ, не принимая во внимание, что он входит в базу данных наряду с множеством других документов. Если представить всю базу данных как единый документ, к ней можно будет применить те же законы, что и к единичному документу. Для того, чтобы избавиться от лишних слов и в тоже время поднять рейтинг значимых слов, вводят инверсную частоту термина. Значение этого параметра тем меньше, чем чаще слово встречается в документах базы данных. Вычисляют его по формуле:

Инверсная частота термина  $i = \log$   
(количество документов в базе данных  
/ количество документов с термином  
 $i$ )

Теперь каждому термину можно присвоить весовой коэффициент, отражающий его значимость:

Вес термина  $i$  в документе  $j =$   
частота термина  $i$  в документе  $j$   $\times$   
инверсная частота термина  $i$ .

Современные способы индексирования не ограничиваются анализом перечисленных параметров текста. Поисковая машина может строить весовые коэффициенты с учетом местоположения термина внутри документа, взаимного расположения терминов, частей речи, морфологических особенностей.

В качестве терминов могут выступать не только отдельные слова, но и словосочетания. Джорж Зипф опубликовал свои законы в 1949 году. Пять лет спустя знаменитый математик Беноит Мандлеброт (Benoit Mandelbrot) внес небольшие изменения в формулы Зипфа, добившись более точного соответствия теории практике. Без этих законов сегодня не обходится ни одна система автоматического поиска информации. Как видите, математический анализ позволяет машине с хорошей точностью, без участия человека распознать суть текста.

### Представление базы данных

Итак, мы разобрались, как машина «понимает» суть текста. Теперь необходимо организовать всю коллекцию документов так, чтобы можно было легко отыскать в ней нужный материал. База данных должна взаимодействовать с пользовательским запросом. Запросы могут быть простыми, состоящими из одного слова, и сложными — из нескольких слов, связанных логическими операторами. Простой запрос оправдывает свое название. Пользователь вводит слово, машина ищет его в списке терминов и выдает все связанные с термином ссылки. Структура такой базы данных проста. Взаимодействие со сложными запросами требует более изощренной организации.

Наиболее простой способ представить элементы базы данных в форме, удобной для многовариантного поиска — создать матрицу документ-термин.

Предположим, база данных имеет 8 документов (Д1, Д2, Е, Д8), в которых содержатся 12 терминов. Если термин входит в документ, в соответствующей клеточке ставится единица, в противном

случае — ноль (в реальной системе все сложнее: помимо прочего, учитываются еще и весовые коэффициенты терминов).

Составим, например, такой запрос: корабли в бутылках. Система обработает запрос: удалит стоп-слова и, возможно, проведет морфологический анализ.

Останется два термина: корабль и бутылка. Система будет искать все документы, где встречается хотя бы один из терминов. Посмотрим на матрицу. Указанные в запросе термины есть в документах: Д1, Д2, Д4, Д7, Д8. Они и будут выданы в ответ на запрос. Однако нетрудно заметить, что документы Д4 и Д7 не удовлетворяют нашим чаяниям — они из области виноделия и никакого отношения к постройке моделей кораблей в бутылках не имеют. Впрочем, система все сделала правильно, ведь, с ее точки зрения, термины корабль и бутылка равноценны.

### Пространственно-векторное представление базы данных

Пространственно-векторная модель позволяет получить результат, хорошо

согласующийся с запросом. Причем документ может оказаться полезным, даже не имея 100% соответствия. В найденном документе может вовсе не оказаться одного или нескольких слов запроса, но при этом его смысл будет запросу соответствовать. Как достигается такой результат?

Все документы базы данных размещаются в воображаемом пространстве (это может быть многомерное пространство, представить которое весьма трудно). Координаты каждого документа зависят от структуры терминов, в нем содержащихся (от весовых коэффициентов, положения внутри документа, от расстояния между терминами и т.п.). В результате окажется, что документы с похожим набором терминов разместятся в пространстве ближе друг к другу.

Предположим, мы хотим найти документы, касающиеся постройки моделей кораблей в бутылках. Составим запрос, например, такой: корабли в бутылках. Получив его, поисковая система удалит лишние слова, выделит термины и вычислит вектор запроса в пространстве

документов. Установив некоторый диапазон соответствия, система выдаст документы, повествующие о необычных увлечениях — хобби, классическом судомоделизме и т.п. В них может вовсе не оказаться некоторых слов запроса, однако документы останутся достаточно релевантными. Термины, относящиеся к вину, будут группироваться в другой точке пространства, и запрос их не затронет. Как видите, «уровнировку» терминов удалось преодолеть. В пространственно-векторной модели термины взаимодействуют друг с другом, что повышает релевантность документов. Понятно, что пространственно-векторная модель лучше воспринимает запросы, составленные на естественном языке, чем матричная.

К сожалению, догадаться, по какой схеме работает та или иная поисковая система Internet, очень трудно. Как правило, создатели держат ее в секрете. Мы в простой форме изложили лишь основы работы поисковой системы. В реальности механизм индексации и структура базы данных значительно сложнее. Однако полученных знаний уже достаточно, чтобы попытаться выработать

оптимальную стратегию поиска информации в сети Internet.

### Стратегия поиска

Итак, мы знаем, как система выделяет ключевые слова. Воспользуемся этим знанием, чтобы сформировать оптимальный запрос. Прежде всего оговорим некоторые исходные предпосылки. Допустим, мы имеем некий текст-источник и хотим найти в сети Internet документы схожего содержания. Откуда возьмется текст-источник? Поскольку сама задача поиска не могла возникнуть из ничего, где-то непременно должна существовать информация, возбудившая интерес к проблеме. Может быть, это журнальная статья, книга, web-страница и т.п. Именно эту информацию и нужно упорядочить и привести в форму, удобную для анализа. Если задача существует только у вас в голове, попробуйте написать небольшое сочинение, изложив свое видение проблемы, — оно и станет текстом-источником. Если бы нам удалось препарировать текст-источник так же, как это делает поисковая машина, по идее, мы

могли бы получить результаты с максимально высокой релевантностью. Попробуем. Возьмем текст-источник и проанализируем его. Если текст-источник — файл на диске вашего компьютера, укажите программе путь к нему — она сама вычислит все необходимые параметры. В противном случае, например, когда текст-источник — страница в журнале, анализ придется сделать вручную.

Последовательность действий такова:

- Подбираем текст-источник. Чем четче описание проблемы в тексте-источнике, тем качественнее и точнее окажется результат. Размытый и путанный текст-источник выудит из поисковой системы столь же бестолковые документы.
- Удаляем из текста стоп-слова (их можно просто вычеркивать).
- Вычисляем частоту вхождения каждого термина. Причем делаем это без учета морфологии слов. Так, слова

ship и ships будут разными терминами. Не нужно учитывать и регистр, все буквы считаем строчными.

- Выписываем на отдельный лист термины в порядке убывания их частоты вхождения (первыми должны идти те, которые встречаются чаще).
- Выбираем диапазон частот. Он должен лежать где-нибудь посередине. Не нужно брать слишком часто или, наоборот, слишком редко встречающиеся термины. Выбор диапазона субъективен. Вам следует ориентироваться на конкретный смысл текста. Необходимость выбирать диапазон вручную не должна смущать, ведь теперь вы выбираете термины не из текста, а из построенного по определенному закону упорядоченного списка.
- Из выбранного диапазона выписываем термины. В

большом тексте в диапазоне может оказаться довольно много слов. Все их применить вряд ли удастся. Достаточно взять 10-20 терминов. Их следует выбирать, руководствуясь, в первую очередь, здравым смыслом. Причем не стоит ограничиваться только характерными терминами, даже если они кажутся наиболее удачными. В список должны попасть и общие слова (их лучше выбирать из средней части диапазона).

- Составляем запрос, располагая отобранные слова в порядке их следования в списке терминов. Запрос должен пониматься машиной как слова, связанные логическим оператором ИЛИ. Это очень важное требование. Для того, чтобы результат не исказился, следует изучить особенности синтаксиса запросов конкретной поисковой системы.

- Отправляем запрос поисковой системе.

В ответ вы можете получить несколько миллионов ссылок. Но не пугайтесь. Если поисковая машина ранжирует результаты (а это еще одно необходимое условие), на первых страницах окажутся практически стопроцентно релевантные документы.

Самое любопытное, что документ — источник запроса (если его аналог существует в Internet) вовсе не обязательно будет возглавлять список. Он может оказаться и на задворках.

### Проверка метода

Разумеется, предлагаемый метод поиска нельзя назвать универсальным. Далеко не все поисковые машины воспримут его с одинаковым восторгом. Какими же свойствами должна обладать поисковая система, чтобы применение метода было оправдано? Множество факторов оказывают влияние на результат. Это и общий объем базы данных, и механизм индексации, структура данных и так далее, и тому подобное. Но наиболее

важными являются два умения поисковой системы: способность понимать запросы, составленные на разных языках (для нас — на русском), и мощное ранжирование результатов. С русским языком все понятно — без него нам в Internet не интересно. Но почему так важно ранжирование? Мы договорились вводить запрос с логикой ИЛИ. Это сильно увеличивает количество возвращаемых поисковой машиной документов. Без ранжирования всякий поиск теряет смысл. Наилучшие результаты дает ранжирование по схеме:

точное соответствие – все слова  
запроса – все слова, кроме  
последнего, – все слова, кроме двух  
последних, – ... – все слова, кроме  
n последних, – первое слово (плюс,  
разумеется, ранжирование по  
количеству терминов в тексте)

Алгоритм может быть и более мощным, но даже при такой последовательности мы можем быть уверены, что, сколько бы документов найдено ни было, наиболее удачные окажутся впереди.

Наилучшие результаты в поиске по предлагаемому методу продемонстрировала система AltaVista. Хотя на тестовый запрос система выдала более 5 миллионов ссылок, для англоязычного запроса на первых трех страницах все ссылки оказались абсолютно релевантными! (Причем документ-источник появился только на третьей странице.) Для русского текста из десяти ссылок на первой странице точными оказались только первые восемь. Однако при ближайшем изучении выяснилось, что это все, что есть в Internet на искомую тему.

Что происходило на других поисковых системах? Картина сложилась пестрая. Одни справились не хуже лидера, другие не справились вовсе. Прежде чем «перемывать косточки», хочу, чтобы вы поняли: неудача говорит не о несовершенстве той или иной поисковой системы или метода, а лишь о неприменимости выбранного метода поиска для данной поисковой машины.

Начнем с зарубежных поисковых систем. Помимо AltaVista очень хороший результат в поиске на английском языке показал HotBoot. Увы, запрос на русском

языке поставил его в тупик. Очевидно, русские буквы для сервера HotBoot — непреодолимое препятствие. Оценить Yahoo! не удалось из-за характерной формы вывода результатов. На тестовый запрос был получен объемистый список каталогов, копаться в которых показалось бессмысленным. Более скромные поисковые машины Northern Light, Excite, Infoseek и другие хотя и выполнили задание, но обилием релевантных ссылок не поразили (возможно, просто из-за того, что их базы данных не столь велики). Для поиска на английском языке на первых страницах оказалось 40-60% релевантных ссылок (впрочем, не такой уж плохой результат). При обработке запроса на русском языке эти системы проявили любопытное единодушие. Было найдено множество документов, но, как мне показалось, никакого ранжирования не было проведено вовсе. В итоге на первую страницу могло попасть, случайно, от силы один-два релевантных документа.

Самостоятельное применение пользователем для решения той или иной задачи любого осмысленного метода требует от системы отсутствия в ней излишней опеки. Автомобиль с



автоматической коробкой передач, которая все делает за вас, вещь хорошая, но вряд ли целесообразно отправляться на нем в ралли Париж — Дакар. То же и в поисковой системе. Чрезмерное увлечение морфологической обработкой слов может лишить поиск гибкости. Именно этим грешат некоторые российские поисковые системы. Спору нет, удобно ввести в поле запроса фразу на естественном языке и получить список документов, которые (по мнению поисковой машины) этому запросу удовлетворяют. Такой поиск дает неплохой результат в среднем. Однако любое отклонение в сторону от утвержденной схемы может резко снизить эффективность поиска. Все эти соображения первоначально вызывали серьезные сомнения в применимости метода на отечественных поисковых серверах. Что же получилось в реальности?

В целом опасения подтвердились. Для системы «Апорт!» выбранная методика оказалась полностью чужда. Rambler представил хорошие результаты только после того, как логика запроса была изменена на И. На первой странице все документы, как для русского, так и английского поиска, оказались полностью

релевантными. Увы, логика И неизбежно ведет к потере весомой части релевантных документов.

С поиском на английском языке отлично справился Яндекс — стопроцентная релевантность на первой странице для англоязычного запроса. Однако русский запрос был обработан заметно слабее. Изучение отклика российских поисковых систем привело к парадоксальному (и крамольному) выводу: морфологическая обработка не обязательно увеличивает число релевантных документов! Разумеется, это утверждение не бесспорно. Для других методик поиска морфологический анализ может оказаться незаменимым; в предлагаемой же нами — он явно лишний. Недаром так хорошо справилась с задачей AltaVista — в ней даже английский текст морфологически не обрабатывается. Все слова для нее, за редким исключением (имеется в виду стоп-лист, но он создается только для англоязычных и близких к ним текстов), лишь последовательность символов.

Предлагаемая методика поиска информации в сети Internet хорошо

подходит для исчерпывающего обзорного поиска. Обзорный поиск незаменим, когда нужно найти как можно больше документов на заданную тему. Анализ текста-источника вручную — весьма трудоемкое и скучное занятие.

## Систематизация поиска

### Средства и способы поиска

Для решения стандартных, часто встречающихся поисковых задач, сеть имеет стандартные же средства — мощные инструменты, разработанными именно для этой цели. Эти **основные средства поиска** работают автоматически, а значит отвечают на запрос достаточно быстро и им не надоедает множество запросов от одного абонента.

**Дополнительные способы поиска** используют возможности, предоставляемые некоторыми другими службами сети, ее персоналом, а также ее абонентами, которые (эти возможности) могут как-то способствовать поиску. Эти способы являются дополнительными, поскольку они:

- либо не предназначены для массового использования;
- либо не являются универсальными (накапливают адреса в недостаточном объеме или по узким направлениям);
- либо не являются стандартными или обязательными для того, кто их предоставляет (то есть, вообще говоря, нет гарантии, что на запрос хоть как-то ответят).

Характерным примером такого дополнительного способа поиска является публикация запроса («крика о помощи») в соответствующей телеконференции — есть некоторая вероятность, что кто-нибудь ответит по существу, но могут и промолчать.

### White- и yellow- поиск

Следует различать White-поиск и Yellow-поиск. Понимание того, что является исходными данными и что — результатом, весьма полезно как само по себе, так и для ориентации среди многочисленных поисковых средств Internet.

Под White-поиском понимается поиск адреса одного конкретного адресата по его достаточно определенному имени (человека — по фамилии, организации — по названию). «Достаточно определенное имя» означает, что объект поиска заранее известен: есть уверенность в его существовании и предлагаемое для поиска его имя вполне уникально. Поэтому есть некоторая вероятность, что искомые данные будут найдены на первом же шаге поиска. White pages в обычных телефонных справочниках — это просто перечни людей или организаций по алфавиту. Если поиск сразу не удался, то сразу возникает предположение, что в исходных данных ошибка, и можно попробовать Yellow-поиск.

Yellow-поиск важен, прежде всего, сам по себе, а не как продолжение неудачного White-поиска. Он позволяет искать адресатов не только и не сколько по их собственному имени, но по иным признакам, охваченных какой-либо классификацией. Формально говоря, под Yellow-поиском понимается поиск адреса одного или нескольких адресатов (какого-нибудь из некоторой группы адресатов) по недостаточно определенному или вообще

неопределенному его имени. «Недостаточно определенное имя» означает неполное, неоднозначное или в чем-то неверное имя, не позволяющее найти абонента средствами White-поиска. «Неопределенное имя» означает либо просто отсутствие какого-либо конкретного имени, либо такое его значение, которое позволяет лишь отнести обозначаемого им абонента к какому-то весьма широкой группе абонентов (например, задача найти адрес «какого-нибудь госпиталя в США»).

Кратко сформулировать разницу между этими двумя видами поиска можно так:

- White-поиск ищет адрес конкретного, заранее известного адресата — ищется адресная информация по одному, но очень важному признаку — по собственному имени адресата. Этот признак — первичный, персонифицирующий. У человека или организации могут меняться почтовый адрес, номер телефона, но

собственное имя остается неизменным.

- Yellow-поиск ищет адрес заранее неизвестного адресата или группы адресатов — сперва ищется собственное имя (имена) по каким-то дополнительным признакам (по роду деятельности, по географическому признаку), а затем уже по найденному собственному имени уже можно отыскать и адресную информацию.

Обычно Yellow Pages системы фактически сразу включают в себя и White Pages — у найденного адресата сразу видны его телефон и почтовый адрес. Кроме того, некоторые Yellow Pages позволяют искать просто в алфавитном списке своих абонентов (white-поиск).

С другой стороны, White pages также содержат элементы yellow-поиска — кроме задания собственного имени они обычно позволяют указать название города, штата и другие, сужающие поиск, данные (что необходимо в случае многих однофамильцев). Возможно, именно

поэтому многие on-line телефонные справочники, выполняющие, фактически white-поиск, называют себя Yellow pages.

### **Yellow-поиск по иерархическому классификатору и по ключевым словам**

Перед помещением в свой каталог, Yellow Pages система предлагает адресату заполнить карточку, где кроме полей с адресной информации ему надо заполнить поля, соответствующие тем признакам, по которым расклассифицирован каталог (например, поля «вид деятельности», «выпускаемая продукция»). При этом возможны два похода: адресат может выбрать готовое значение признака из предлагаемого списка (например, из классификатора отраслей), либо он может написать это значение произвольно, «своими словами». В первом случае последующий поиск будет проводиться по тому же самому классификатору (обычно иерархического типа), во втором случае поиск будет проводиться по «ключевым словам», которые будут сравниваться (необязательно до полного совпадения) с

теми, которыми характеризовал себя адресат.

Поиск по иерархическому классификатору происходит до точного совпадения образца с каким-либо из содержащихся там значений, поиск же по ключевым словам может проводиться до разной степени совпадения. Обычно условие совпадения бывает «мягким» — достаточно, чтобы предлагаемый образец оказался фрагментом в сравниваемом тексте, однако иногда поисковые системы позволяют усиливать это условие.

Обычно поиск по признаку «отрасль промышленности» проводится в иерархическом классификаторе (либо в стандартном, используемом статистическими службами, либо в «самодельном», меньшим по размерам, но более удобном), а поиск по признаку «выпускаемая продукция» происходит по ключевым словам, хотя реально для этого используется тоже иерархический классификатор (обычно стандартный кодификатор продукции и товаров).

В отличие от Yellow pages организаций, которые могут быть собраны без опроса самих адресатов (по

отраслевым спискам или по спискам государственных статистических органов), Yellow pages по частным лицам малочисленны и не являются исчерпывающими. Включение в такие базы данных является делом сугубо добровольным, при заполнении карточки предлагается возможность заполнить на выбор по желанию поля типа «место работы», «профессия», «хобби и интересы», «место учебы», «прежнее место жительства» и т.п.

### **Обратный white- и yellow-поиск**

Некоторые телефонные White pages имеют возможность обратного поиска — нахождения адресата по заданному телефонному номеру. Yellow pages также имеют похожую возможность — для адресата, найденного по ключевым словам или по иерархии значений признака, приводятся и значения признаков из других иерархий (найдя, например, изготовителя алюминиевой фольги, можно тут же увидеть, что это — предприятие авиационной промышленности).

### Поиск e-mail адресов

Основные задачи:

- white-поиск e-mail адреса конкретного человека;
- white-поиск e-mail адреса конкретной организации;
- white-поиск e-mail адреса в конкретной организации;
- white-поиск e-mail адреса узла сети;
- обратный white-поиск абонента по e-mail адресу;
- yellow-поиск e-mail адреса организации;
- частный случай: поиск e-mail адресов учебных заведений и их студентов.

Для поиска e-mail адресов людей, организаций и узлов сети специально предназначены справочная служба Netfind (доступна в on-line режиме) и автоматические адресные сервера типа Whois (доступны и в on-line, и по e-mail). Свои адресные базы они наполняют также автоматически, постоянно и в массовом порядке прямо из адресных баз

крупных узлов сети (или ищут прямо в них), которые содержат достаточно полный набор записей о своих абонентах.

Основные средства поиска:

- автоматическая справочная служба Netfind;
- автоматические адресные сервера типа Whois;
- автоматические адресные сервера других типов;
- справочные системы Yellow pages, содержащие и e-mail адреса.

Дополнительные способы поиска:

- специальные запросы к автоматическим серверам какой-либо массовой службы (телеконференции или списки рассылки) или получение каким-либо другим путем списка ее абонентов;
- запрос к абонентам какой-либо массовой службы;
- запрос к персоналу узла сети;
- запрос к автоматическому адресному серверу или к

персоналу какой-либо организации, предоставляющей публичный доступ к адресам своих корреспондентов (клиентов, партнеров и т.п.).

### Справочная служба Netfind

Netfind по одному запросу производит White-поиск во многих адресных базах данных и в других справочных службах. Для использования этого сервиса на своем компьютере нужно иметь клиентскую программу Netfind, при ее отсутствии можно заходить telnet-ом (logname «netfind») на какой-либо из Netfind-серверов в мире:

- archie.au
- bruno.cs.colorado.edu
- dino.conic.it
- ds.internic.net
- lincoln.technet.sg
- macs.ee.mcgill.ca
- malloco.ing.puc.cl
- monolith.cc.ic.ac.uk
- mudhoney.micro.umn.edu
- netfind.oc.com

- netfind.vslib.cz
- nic.nm.kr
- nic.uakom.sk
- redmont.cis.uab.edu

Кроме того, Netfind доступна и через WWW-сервера:

- <http://www.internic.net>
- <http://www.nova.edu/Inter-Links>

### Whois-сервера

Whois-сервера производят White-поиск в больших базах адресов крупных узлов или целых сетей. Работа с whois-сервером происходит по определенным правилам (whois-протокол), поэтому желательно пользоваться специальной клиентской программой (whois-клиент). В UNIX-е такая программа так и называется: whois, ей можно задавать адрес любого из whois-серверов для поиска на нем. Whois-сервера доступны также и через telnet.

- <http://www.internic.net>
- <telnet://nic.ddn.mil>  
<mailto:service@nic.ddn.mil> HELP

- `telnet://rs.internic.net`  
`mailto:mailserv@internic.net`  
HELP
- `telnet://whois.internic.net` — «Главный» whois-сервер Internet, содержащий адреса людей и узлов.
- `telnet://whois.ripe.net` — Whois-сервер RIPE — организации, объединяющей крупные Internet-сети в Европе.
- `ftp://sipb.mit.edu/pub/whois/whois-servers.list`
- `ftp://rtfm.mit.edu/pub/whois/whois-servers.list` — Список известных whois-серверов.
- `http://www.nova.edu/InterLinks/cgi-bin/whois.pl` — «Обратный» Whois-поиск организации по ее домену (части e-mail адреса).

### **Обратный поиск по e-mail адресу (Finger)**

Поисковая служба Finger позволяет найти конкретного человека по его полному e-mail адресу. Ей можно пользоваться, находясь в операционной

системе UNIX и искать абонентов на других UNIX-машинах (благо что Internet состоит именно их них). Finger может показать регистрационное имя абонента на удаленной машине (login name), некоторые важные данные о нем, если он их указал (название организации или подразделения, где он работает и его телефон), а также время последнего чтения им своей электронной почты (можно определить, пользуется ли он еще своим e-mail адресом). Многие UNIX-машины имеют finger-сервер, принимающий запросы не только о конкретных e-mail адресах, но и позволяющий делать различные поиски среди своих пользователей.

### **Справочная служба X.500**

Глобальная справочная служба X.500 объединяет локальные справочные службы, выглядящие как один большой иерархический справочник. Для доступа требуется специальная программа (локальный X.500 клиент), однако если ее нет, то можно воспользоваться любым из имеющихся в мире общедоступных удаленных X.500 клиентов, доступ к которым возможен по telnet, через



## Поиск информации

---

некоторые Gopher- и WWW-сервера и даже через e-mail:

- <http://www.dante.net:8888/>
- <telnet://jethro.ucc.su.oz.au> fred Австралия
- <telnet://elem4.vub.ac.be> dua Бельгия
- <telnet://x500.denet.dk> de Дания
- <telnet://login.dkuug.dk> ds Дания
- <telnet://login.dkuug.dk> dua Финляндия
- <telnet://x500.tu-chemnitz.de> x500 Германия
- <telnet://ashe.cs.tcd.de> de Германия
- <telnet://x500.ieunet.ie> de Ирландия
- <telnet://jolly.nis.garr.it> de Италия
- <telnet://zoek.nic.surfnet.nl> zoek Нидерланды
- <telnet://elcl.mat.torun.edu.pl> de Польша
- <telnet://chico.rediris.es> directorio Испания
- <telnet://hypatia.umdc.umu.se> de

## Поиск информации

---

Швеция

- <telnet://nic.switch.ch> dua Швейцария
- <telnet://dir.ulcc.ac.uk> dua Англия
- <telnet://paradise.ulcc.ac.uk> dua Англия (проект Paradise)
- <mailto:Directory@uninett.no> HELP

### Системы Yellow Pages

На Yellow Pages обычно используются классификации по по роду деятельности, по выпускаемой продукции и оказываемым услугам, по географическому признаку. Иногда они дополнены white-поиском (просто по алфавиту).

- <http://www.four11.com>
- <http://www.lookup.com>
- <http://www.whowhere.com>
- <http://sunsite.oit.unc.edu/~masha>
- <http://okra.ucr.edu/okra>

### **Запросы к серверам телеконференций и спискам рассылок**

У массовых служб типа систем телеконференций или списков рассылки обычно предусмотрен свободный доступ к спискам их абонентов.

- <mailto:news@kiae.su> USTAT
- <mailto:news@demos.su> USTAT

Сервера новостей, раздающих новости USENET, имеют команду USTAT, позволяющие получить список абонентов, подписанных на этом сервере на ту или иную телеконференцию.

- <mailto:listserv@bitnic.bitnet>

### **Помощь из конференций Usenet**

#### **comp.mail.maps**

В этой USENET-овской телеконференции регулярно публикуются списки UUCP станций, которые затем сохраняются на многих файловых серверах.

#### **soc.net-people**

Размещение объявлений-запросов адресов.

#### **soc.net-people comp.mail.misc**

Регулярно публикуется FAQ по поиску адресов.

#### **soc.college**

Регулярно публикуется FAQ по поиску адресов в колледжах и университетах.

#### **relcom.maps**

Регулярно публикуется FAQ по поиску адресов.

#### **comp.mail.misc**

Регулярно публикуется FAQ по поиску адресов.

#### **soc.college**

Регулярно публикуется FAQ по поиску адресов в колледжах и университетах.

#### **relcom.maps**

#### **ukr.maps**

Телеконференции сети Relcom, предназначенные для публикации

запросов адресов, а также для размещения разного рода материалов, способствующих поиску (например, списков узлов сети Relcom и списки абонентов некоторых узлов).

### Поиск www-адресов в Internet

#### WWW-адрес, WWW-ссылка

Для некоторого единообразия терминологии далее вместо термина URL (Universal Resource Locator) используется термин «WWW-адрес». При необходимости будут различаться WWW-адрес сервера и WWW-адрес документа. В качестве синонима будет использоваться термин «ссылка на WWW-сервер» или «ссылка на WWW-документ».

#### Search engine

Поисковый WWW-сервер, обрабатывающий запросы на поиск WWW-адресов документов. В WWW-технологии каждый WWW-документ при его создании может быть снабжен набором ключевых слов по усмотрению автора. Поисковые сервера считывают эти ключевые слова, находят такие же в своих больших словарях и добавляют ссылку на данный WWW-документ в список уже имеющихся

ссылок для каждого такого слова. Кроме поиска по ключевым словам, все большие поисковые сервера имеют универсальные иерархические классификаторы, охватывающие все отрасли знаний, сферы деятельности, области интересов, стороны общественной жизни и т.п.

Некоторые поисковые сервера содержат внутри себя Yellow и White Pages WWW-адресов серверов.

#### Yellow Pages WWW

Поисковый WWW-сервер, обрабатывающий запросы на поиск WWW-адресов серверов (Home Page организаций). Все большие Yellow Pages WWW системы позволяют находить не только WWW-адрес искомой организации, но вместе с ним показывают сразу и телефон, факс, телекс, обычный и e-mail адрес. Иногда дается и краткое описание профиля организации.

#### Index sites

WWW-сервера, содержащие большое количество ссылок на другие WWW-сервера или на WWW-документы. Могут быть специализированными или универсальными. Если предметная область

сложна или универсальна, то ссылки обычно иерархически расклассифицированы, в простом случае они просто упорядочены по алфавиту. От Search engines и Yellow Pages WWW отличаются ограниченным числом ссылок (часто это круглое число: 100 или 1000), их тщательным подбором и отсутствием поиска по ключевым словам.

### Reference Sites

Собирательное название для Search engines, Index sites, Yellow Pages WWW и других похожих источников WWW-адресов.

### Задачи, средства и способы поиска WWW-адресов

Поиск WWW-адресов интересен как сам по себе (поиск WWW-документы на заданную тему), так и как задача поиска организаций в сети. В свою очередь, найденный WWW-сервер конкретной организации может быть интересен как сам по себе, так и в качестве источника адресной информации (телефоны и e-mail адреса самой организации и ее подразделений).

Типичные поисковые задачи:

- поиск WWW-адреса сервера конкретной организации или конкретного человека;
- поиск WWW-адреса сервера какой-то организации по определенным условиям;
- поиск новых WWW-адресов серверов;
- поиск новых WWW-адресов документов.

Основные средства поиска (reference sites):

- большие универсальные поисковые WWW-системы (search engines);
- поисковые WWW-системы по нескольким search engines;
- остальные универсальные и специализированные search engines и index sites;
- справочные системы Yellow Pages по WWW-адресам серверов;

- MetaReference sites (каталоги search engines, index sites, Yellow Pages WWW).

Дополнительные способы поиска:

- персональные и тематические подборки ссылок на WWW-сервера;
- newsletters и guides на WWW-серверах, объявления в телеконференциях, списки рассылки
- запрос к абонентам какой-либо массовой службы;

Следует отметить, что число даже основных поисковых средств (различных reference sites) достигает нескольких сотен, если не больше. Кроме широко известных references sites, перечисленных ниже, Ниже перечислены либо наиболее известные из них, либо те, сообщения об существовании или появлении которых просто привлекли внимание по той или иной причине. Порядок их перечисления, а также само присутствие в списке никоим образом не отражает их полезность, оценить которую невозможно при беглом обзоре.

Ответ на вопрос — почему их столько, очевиден, на вопрос — зачем их столько, и какими из них стоит пользоваться — пока совершенно не ясен. Понятно только, что пользоваться ими всеми по очереди можно только от отчаяния. Желаящие обозреть все величие проблемы приглашаются в недра иерархии Reference поискового сервера Yahoo.

### **Большие универсальные поисковые WWW-системы (search engines)**

Поисковые WWW-сервера позволяют найти отдельные WWW-документы (WWW-страницы), относящиеся к заданным тематикам или снабженные заданными ключевыми словами или их комбинациями. На больших поисковых серверах имеются оба эти способа поиска (по иерархии понятий и по ключевым словам).

Наполнение поисковых серверов происходит либо автоматически (они непрерывно обзорывают все WWW-сервера в мире), либо вручную (с отбором материала).

Поисковый сервер обычно имеет ссылки и на остальные большие поисковые сервера, но не передает им автоматически запрос на поиск.

Наиболее известные поисковые WWW-сервера:

**<http://www.yahoo.com>**

Один из первых и самых известных search engine, отличающийся развитой иерархической классификацией. Наполняется вручную, имеет наименьшее количество ссылок по сравнению с другими search engines (по некоторым темам примерно в 10 раз меньше Webcrawler-a, но по другим — больше его).

Кроме WWW-адресов, содержит ссылки на ftp- и gopher- ресурсы. Имеется иерархический классификатор и поиск по ключевым словам, поддерживающий операции «и», «или», но только одного типа в одном запросе. Найденные по ключевым словам ссылки снабжаются еще и указанием их положения в иерархической классификации сервера. Для сокращения диапазона поиска предусмотрена возможность искать в пределах текущей подтемы классификатора. На первом уровне

иерархии имеется раздел «References», содержащий много ссылок на разного рода reference sites.

Имеет развитой сервис новостей.

**<http://www.lycos.com>**

Один из самых известных и самых больших по количеству ссылок search engine.

Имеется иерархический классификатор и поиск по ключевым словам, поддерживающий операции «и», «или», но только одного типа в одном запросе. Операция «и» вообще не работает, операция «или» работает неверно — возвращает количество ссылок, равное максимальному из количеств ссылок по отдельным аргументам).

**<http://www.webcrawler.com>**

Относительно небольшой search engine, по-видимому, наполняется вручную — примерно в 10 раз меньше Lycos-a.

Имеется иерархический классификатор и поиск по ключевым словам, поддерживающий операции «и», «или», «не» и их комбинации, которые,

однако, отрабатываются не всегда верно (находятся документы, совершенно не подходящие под условия поиска).

**<http://www.inktomi.com>**

Новый и, вероятно, самый большой по количеству ссылок сервер, наполняется автоматически и хранит все найденные ссылки (примерно в 5 раз больше Lycos-a).

Имеется иерархический классификатор и поиск по ключевым словам, поддерживающий операции «и», «или», но только одного типа в одном запросе.

Кроме собственно поиска документов имеет Yellow Pages по разным категориям.

**<http://www.infoseek.com>**

Средний по количеству ссылок search engine (по некоторым темам чуть больше Webcrawler-a, но по некоторым — больше Lycos-a).

Имеет иерархическую классификацию и поиск по ключевым словам, не поддерживающий операций «и», «или».

Поиск может производиться: по всему WWW-пространству, лишь среди серверов, отобранных хозяевами данного сервера (в этом случае каждая найденная ссылка снабжена краткой аннотацией), лишь среди серверов компаний (Yellow Pages поиск, каждая найденная ссылка снабжена кратким описанием профиля компании), среди статей USENET, среди e-mail адресов, среди свежих новостей (сервер имеет дополнительный очень развитый сервис новостей). Результат поиска дополнительно снабжается и списком подходящих тем из иерархической классификации (общий список тем для данного запроса, а не конкретные иерархические пути для каждой найденной ссылки, как у Yahoo).

**<http://www.altavista.com>**

Большой search engine (немного меньше Lycos-a по числу ссылок). Имеет лишь поиск по ключевым словам, поддерживающий только операцию «и». Поиск может проводиться по всему WWW-пространству или среди статей USENET.

Для проведения «тотального» поиска сразу по нескольким большим серверам

имеются специальные средства — сервера, транслирующие запрос на поиск в формы, подходящие для каждого из охватываемых поисковых серверов.

- <http://www.w3com.com/fsearch>  
FrameSearch
- <http://www.miyawaki.pair.com>  
Bingo!
- <http://www.albany.net/~wcross/allsrch.html> All-in-One-Search-Page
- <http://www.postnet.se/internet/SuperSearch.htm> InterNet Stockholm
- <http://members.aol.com/markwelch/search.htm>

### **Системы Yellow Pages по WWW-адресам серверов**

В отличие от поисковых WWW-серверов, системы Yellow Pages с WWW-адресами, содержат ссылки не на отдельные WWW-документы, а на WWW-сервера различных организаций. Сами же организации подробным образом рассортированы по роду деятельности, по

выпускаемой продукции и оказываемым услугам, по географическому признаку, просто по алфавиту (White Pages).

### **<http://www.gnn.com>**

Одна из первых и наиболее известных систем Yellow Pages для WWW-адресов. Содержит ссылки на 2500 WWW-серверов. Имеет универсальную иерархическую классификацию, а также алфавитный список тем и алфавитный список серверов (White Pages). Каждая ссылка снабжается комментарием примерно в один абзац.

### **<http://www.yellow.com>**

Известная система Yellow Pages. Каждая ссылка снабжена комментарием с названием организации, почтовым адресом, телефоном, факсом, телексом.

### **<http://www.directory.net>**

Поиск по паре ключевых словам, поддерживающий операции «и», «или», либо использующий эту пару как одну строку.



### **Персональные и тематические подборки ссылок на WWW-сервера**

Очень многие WWW-сервера, содержат подборки WWW-ссылок по какой-либо конкретной тематике, либо по нескольким темам (обычно лежат на персональных Home pages). Эти подборки не претендуют на полноту (часто ограничиваются круглым числом 1000), так как являются результатом ручного отбора, но как раз по этой причине они могут быть особенно полезны для начального ознакомления с какой-либо темой.

Некоторые из таких подборок:

- [http://www.myanmar.com/myanmar/mother\\_of\\_all\\_link.html](http://www.myanmar.com/myanmar/mother_of_all_link.html)
- <http://www.hookup.net/~jmorris/cool>
- <http://www.tias.com/dealers.html>
- <http://www.island.net/~rednikki>
- <http://www.vnet.net/users/voyager/CybernetVoyager.html>

- <http://homepages.enterprise.net/kevinc/addrand.html>
- <http://www.skylink.net/%7Eanytime/cool.html>
- <http://www.dubleclick.com/coolLinks.htmlLinks>
- <http://cent.com/abetting/Links.html>
- <http://btdqs.usgs.gov/index/coollink.html>
- <http://www.n-vision.com/Quazimoto/bookmark.htm>
- <http://cactus.org/%7Ewoan/hotlist.html>
- <http://totalweb.totalpc.com/cool.html>
- <http://www.craftstore.com/links.htm>
- <http://www.dancon.com/cool.html>
- <http://www.gil.com.au/%7Etedb/fav.html>
- <http://www2.ecst.csuchico.edu/%7Ebeej/hotlist.html>

- <http://felix.scvnet.com/%7Etim/bkmk1.html>

### Поиск телефонов и почтовых адресов в Internet

Основные задачи:

- white-поиск телефона и почтового адреса конкретного человека;
- white-поиск телефона и почтового адреса конкретной организации;
- обратная задача — поиск абонента по его телефону и почтовому адресу;
- yellow-поиск человека (с нахождением его телефона и почтового адреса);
- yellow-поиск организации (с нахождением ее телефона и почтового адреса);

Основные средства поиска:

- on-line универсальные телефонный справочники White Pages;
- on-line универсальные системы Yellow Pages.

Дополнительные способы поиска:

- поиск e-mail или www-адреса — возможно, телефон или почтовый адрес будут указаны вместе с ними;
- запрос к справочной службе организации для поиска телефонов и почтовых адресов ее сотрудников или организаций, как-то с ней связанных (подразделений, филиалов, партнеров).

### Телефонные справочники White Pages

<http://www.c2.net/~buttle/tel>

Метасправочник телефонных справочников по всему миру. Указывает преимущественно на большие универсальные телефонные справочники, поддерживаемые телефонными компаниями, а не на небольшие yellow-pages.

<http://www.switchboard.com>

Телефонный справочник США (White Pages).

<http://www.yahoo.com/search/people>

Yahoo People Search — телефонный справочник США (White Pages).

[gopher://gopher.nd.edu:70/11](http://gopher://gopher.nd.edu:70/11)

Phonebooks — телефонный справочник организаций США (государственных органов, учебных заведений, научных организаций).

<http://www.procd.com>

Directory of Directories — содержит более 100 ссылок на телефонные справочники колледжей и университетов по всему миру.

### Телефонные справочники Yellow Pages

В Yellow Pages для поиска обычно используются признаки:

- **business type, activity sector** — область (вид) деятельности, отрасль, например: «машиностроение», «торговля», «юридические услуги».
- **products & services** — производимая продукция, оказываемые услуги, например:

«автомобили», «зерно», «услуги адвоката».

- **business name, company** — собственное имя (наименование) предприятия, фирмы, организации.

<http://www.telephonebook.com>

Телефонный справочник бизнес-организаций США (Yellow Pages).

<http://www.databaseamerica.com>

Телефонный справочник (Yellow Pages).

<http://www.yellow.com>

Телефонный справочник и справочник www-адресов (Yellow Pages).

<http://bigyellow.com>

NYNEX US Business Yellow Pages. Позволяет искать по типу и по наименованию бизнеса. Задание типа бизнеса очень неудобно — он не выбирается из готового списка (хотя тот и имеется), а задается произвольно. Делается попытка поиска и если такого типа бизнеса в списке нет, только тогда он появляется и в нем можно выбрать.

<http://www.europages.com>

Справочник по Европе. Позволяет искать по типу бизнеса и по наименованию бизнеса, а также по выпускаемой продукции (оказываемым услугам).

### **Платные телефонные и адресные справочники White и Yellow Pages**

<http://www.searchamerica.com>

Платный телефонный справочник США (White pages). Позволяет поработать на 1 \$ бесплатно, при условии введенного номера кредитной карточки.

## **Поиск в стиле ДИСКo**

Если нам требуется что-то найти в Internet, мы отправляемся на поисковый сервер и запрашиваем нужную информацию. Можно также послать свой запрос сразу на несколько серверов с метапоисковой страницы, такой как All-in-One ([www.allonesearch.com](http://www.allonesearch.com)), Inference Find ([www.infind.com](http://www.infind.com)) или наш российский «Следопыт» (<http://www.medialingua.ru/www/wwwsearc.htm>), который, помимо прочего, умеет запрашивать заданные слова во всех

формах, а также переводить их с русского на английский и обратно. Еще один вариант — воспользоваться метапоисковой клиентской программой, которая будет отправлять запросы непосредственно с вашей машины, например Copernic ([www.copernic.com](http://www.copernic.com)) или Inforia Quest ([www.inforia.com](http://www.inforia.com)).

В конце прошлого года полку метапоисковых клиентов прибыло: компания «ДИСКo» разработала своего «Искателя». Как и другие продукты «ДИСКo» (диспетчер файлов «Командир» и автономный браузер «Качалка»), «Искатель» распространяется компанией «Арсеналь». Что же умеет эта программа?

Она отправляет запросы параллельно на несколько популярных серверов; поддерживаются три российских сервера — Rambler, Яндекс и «Апорт» — и четыре американских — AltaVista, Lycos, Yahoo! и HotBot. Перед началом поиска вы указываете, какие серверы вам нужны (например, если в запросе есть русские слова, лучше отключить Yahoo! и HotBot, которые не воспринимают кириллицу). Можно послать параллельно несколько запросов, каждый со своими параметрами.

Определенную проблему представляет то, что на каждом поисковом сервере используется собственный язык запросов (правда, у этих языков много общего, и даже самые изощренные из них относительно просты). «Искатель» позволяет применить к словам, составляющим запрос, логическую операцию «и» (искать документы, в которых встречаются все перечисленные слова) либо «или» (искать документы, в которых есть хотя бы одно из них), заказать поиск введенной строки целиком (попробуйте сделать это, например, на Rambler без «Искателя») и задать слова, которых не должно быть в найденном документе. Кроме того, можно сформировать сколь угодно изощренный запрос и отправить его на сервер «как есть».

Реализована проверка корректности найденных ссылок. Как ни часто происходит обновление индексов на поисковых серверах, оно не успевает за изменениями в Сети, так что многие зарегистрированные в индексах документы оказываются недоступны. Чтобы не разочаровываться, получив вместо нужного документа отказ, можно еще в «Искателе» выяснить, существует ли он. Программа предусматривает как автоматическую проверку всех ссылок, так и индивидуальную, по выбору пользователя.

Результаты поиска представляются в виде очень компактного, легко обозримого дерева, которое перешло в «Искатель» из «ДИСКо Качалки». Деревья, кстати, вплотную связаны с новой технологией поиска XTree — совместной разработкой компаний «ДИСКО» и «Арсеналь», о которой мы сообщали в прошлом году.

«Искатель» объединяет ответы от разных серверов. Ссылки на один и тот же документ представляются как одна, а разные копии одного и того же документа (если программе удастся это установить), в том числе и варианты, различающиеся кодировкой русских букв, размещаются под одной вершиной дерева поиска.

Программа имеет и ряд дополнительных функций, причем некоторые из них составляют предмет особой гордости разработчиков: таковы, в частности, возможность запоминания параметров поиска и его результатов для последующего уточнения и обновления, а также экспорт ссылок в закладки Internet Explorer и импорт закладок. Мне они почти никогда не бывают нужны, но многим, вероятно, пригодятся.

## Особенности русского языка в Internet

Что ни говори, а, заходя на русскоязычные Web-страницы, мы сейчас чаще видим обычные русские слова, чем, к примеру, текст на ломаном английском с предложением выбрать кодировку. Совместными, хоть и не всегда согласованными усилиями разработчики браузеров, администраторы серверов и Web-мастера в целом решили проблему поддержки русского языка в WWW; досадные исключения встречаются, но это именно исключения. А отечественные поисковые системы ищут слова из запросов в разных кодировках, причем некоторые («Яндекс», «Апорт») еще и в разных формах.

### Русские адреса

Казалось бы, чего еще желать? На этот с виду чисто риторический вопрос разработчики из компании «МедиаЛингва» дали вполне осмысленный ответ: пожелайте, чтобы адреса страниц тоже задавались по-русски. Сами URL (Uniform Resource Locators — единообразные определители местонахождения ресурсов), которые незабвенный Май Иванович Мухин в своих обзорах «Перелетные мухи» называл

«урлами», русифицировать, конечно, нельзя — по стандарту в них допустимы только латинские буквы, цифры и некоторые специальные символы. Однако можно надстроить над URL еще один уровень адресов, на котором таких ограничений не будет.

Эта идея была реализована в созданной весной этого года Национальной службе имен (НСИ). Служба, как уже упоминалось, разработана компанией «МедиаЛингва», а поддерживает ее компания «100%». На сервере НСИ, который находится по адресу <http://www.names.ru>, хранится база адресов, позволяющая искать Web-страницы по их «русским именам» — названиям фирм и организаций, именам владельцев, характерным ключевым словам, рекламным девизам. Скажем, по запросу «Открытые системы» будет найдена страница <http://www.osp.ru>, по запросу «МПС» — <http://www.css-mps.ru>, по запросам «Яблоко» и «Явлинский» — <http://www.yabloko.ru>, а по запросу «машинный перевод» — <http://www.ets.ru>.

На том же сервере (и на нескольких других) имеется бесплатный драйвер, установив который можно вводить запросы к базе русских имен прямо в поле адреса;

получив запрос, драйвер связывается с сервером НСИ, находит в базе нужный URL и производит автоматическое переключение на соответствующую страницу. Таким образом, с точки зрения пользователя вводимые им в качестве адресов русские слова и словосочетания будут полностью эквивалентны стандартным URL. Драйвер существует для Windows 3.x и 95/NT. Те же, кто работает в другой операционной системе или по какой-либо причине не хочет устанавливать драйвер, могут пользоваться Web-интерфейсом НСИ: он также есть на нескольких серверах, в том числе, разумеется, на самом <http://www.names.ru> и на <http://www.infoart.ru>.

### **База имен**

База имен состоит из двух частей: имена фирм и организаций и имена частных лиц. Первые должны быть уникальными, т. е. коль скоро издательство ЭТС зарегистрировало для своей страницы имя «машинный перевод», то фирме ПРОМТ это сделать уже не удастся (зато на страницу, посвященную ПРОМТ, — правда, не на ее собственную, а на принадлежащую компании «Агама», — <http://russia.agama.com/promt.htm>, — отсылает запрос «программы перевода»); несколько адресов могут оказаться

результатом поиска, но лишь в случае, когда запрос не имеет точного соответствия в базе. В обратную сторону, как мы видели, таких строгих ограничений нет: одному URL могут соответствовать и несколько имен; регистрационная форма предусматривает, помимо основного имени, до четырех синонимов. С именами частных лиц все наоборот: требование уникальности, естественно, отсутствует (человек не должен страдать из-за того, что у него есть однофамильцы), а синонимы — прозвища, псевдонимы — не допускаются. Имена обоих типов можно снабдить краткой аннотацией, с которой НСИ будет работать как обычная поисковая система.

Регистрация в базе производится только по заявкам, причем сотрудники НСИ предварительно проверяют соответствие имен и аннотаций страниц их реальному содержанию. Если частное лицо регистрирует только свой адрес электронной почты (что допускается), то проверить правильность сообщенных сведений, очевидно, невозможно. По вновь зарегистрированному адресу посылается уведомление — оно гарантирует, что владелец адреса действительно хотел зарегистрироваться, — а вот гипотетическую ситуацию, в которой кто-

то присваивает себе имя другого человека, чтобы перехватить часть его переписки, предотвратить нельзя.

Регистрация с 1 октября сделана платной, она стоит 15 долл. в год для частных лиц и 40 долл. в год для фирм. Можно также бесплатно зарегистрироваться в демонстрационном режиме.

### **Перспективы и проблемы**

Как уже упоминалось, для частных лиц в базу заносится информация об адресе не только домашней страницы, но и электронной почты. В дальнейшем предполагается обеспечить работу с русскими именами также в почтовых программах и открыть Национальную почтовую службу, естественным образом дополняющую Службу имен.

Две компании — «Телепорт-ТП» и «МТУ-Информ» — уже подписали со «100%» договор об установке дополнительных серверов НСИ, а со временем появятся и другие, так что запросы к службе (двум службам) будут обрабатываться быстро даже тогда, когда их число многократно возрастет. Драйвер НСИ должен быть помещен на русский «портал» Internet Explorer 5 и войти в состав русской версии Netscape Communicator

5. Вероятно, многие провайдеры включат его в свои пакеты программ для работы с Internet. Так что планы развития системы, с помощью которой мы сможем «стопроцентно по-русски» общаться со Всемирной сетью, как видим, весьма обширные.

Несомненно, по мере своего расширения Служба имен столкнется с различными проблемами технического и организационного характера. Имена НСИ, в отличие от стандартных доменных, никак не структурированы — не начнет ли это мешать эффективной работе с базой при возрастании ее объема? Ведь будет увеличиваться как число русскоязычных Web-страниц (по данным AltaVista, число страниц в домене .ru сейчас превышает 750 тыс.), так, надо полагать, и процент страниц, зарегистрированных в НСИ. Возможны осложнения и из-за способа, каким в НСИ обеспечивается уникальность имен (хотя вряд ли можно было бы предложить что-то более разумное). Если на какое-то имя имеют право несколько страниц (как в примере с запросом «машинный перевод»), оно присваивается той из них, для которой первой была подана заявка, так что она оказывается в никак содержательно не обусловленном привилегированном



положении. Некоторой компенсацией может служить то, что, пока происходит подключение, НСИ показывает страницу с приглашением поискать нужный ресурс также с помощью серверов «Яндекс» и «Следопыт». Однако оно довольно быстро исчезает, и новичок (а служба рассчитана в первую очередь на новичков) может не придать ему значения.

Не вполне ясно, насколько активно будет происходить регистрация. Перепись населения русскоязычного Internet на сервере «Созвездие Internet» (<http://www.stars.ru/people.htm>) проходила довольно вяло: в общей сложности было заполнено менее 35 тыс. анкет (сравните это с 750 тыс. Web-страниц). База адресов электронной почты «Электронная Россия» (<http://www.dubna.ru/eros>) содержит сейчас более 100 тыс. записей, но по большей части эти адреса не внесены в базу их владельцами, а получены автоматически путем сканирования телеконференций. Более чем прохладное отношение к регистрации, разумеется, связано со страхом перед попаданием анкет в руки спэммеров или еще каких-нибудь людей, преследующих неблагоприятные цели, причем определенные основания для опасений, конечно, есть: даже

если честные намерения анкетирующих не вызывают сомнений, невозможно гарантировать, что анкеты не будут похищены хакерами. Однако куда более важную роль, по-видимому, играет самая обычная лень.

Так или иначе, в обоих случаях совершенно бесплатная регистрация (среди участников переписи к тому же разыгрывались призы) не вызвала особого энтузиазма. Не упадет ли он просто до нуля, если за регистрацию станут брать деньги? Как ни странно, вполне вероятно, что все будет наоборот, — вспомним, как Том Сойер красил забор. Кроме того, «100%» планирует заключить с рядом провайдеров договоры, по которым они включают регистрацию в базе НСИ в комплекс своих услуг; разумеется, пользователь будет иметь право отказаться от регистрации, но можно рассчитывать, что он этого не сделает, — опять же в силу лени.

Проблемы неизбежны: НСИ — новаторский проект, и странно было бы, если бы они не возникали. Но главное уже сделано: служба убедительно продемонстрировала, что работать с Internet по-русски вполне можно и что это удобно.

## Оффлайн браузеры

Очень многие привыкли относиться к прогулкам по Internet, как к походам в бескрайний платный читальный зал. Такой подход свойственен не только людям экономным, но и той (не менее уважаемой) части нетизенов, которые используют добытые гипертексты в профессиональных целях. Жизнь этой категории населения, как правило, скрашена наличием в арсенале программных средств так называемых «оффлайн браузеров». То есть программ, во многом напоминающих привычные браузеры, но позволяющих накапливать ресурсы с тем, чтобы потом иметь возможность просматривать их в режиме оффлайн. Эти программы во многом напоминают так называемые менеджеры загрузки (download managers), но ориентированы не на перекачку одного или нескольких файлов одновременно, а на перекачку всего содержимого сайта с сохранением его внутренней структуры.

### Лучшие оффлайн браузеры

- BlackWidow
- DiscoPump

- Offline Explorer
- Web Stripper
- WebZip

## BlackWidow

BlackWidow — «чёрная вдова». Название изобличает разработчиков, как почитателей кинематографического жанра триллера (в частности, творчества Боба Рэфельсона) или, что скорее, как дотошных исследователей физиологии представителей класса членистоногих.

Инсталляционная строка: Next ⇨ Next ⇨ Next ⇨ Next ⇨ Yes ⇨ Yes ⇨ Finish.

**Внимание!** Желательно несколько усложнить инсталляционную процедуру отказом от хранения результатов деятельности программы на системном диске C.

При такой последовательности действий, на рабочем экране появляется черный квадрат с аккуратно вписанным кровавым паукообразным существом. Не пугайтесь. Это не кусается, но может пригодиться для работы с гипертекстами.

Подключитесь к Сети, запустите программу BlackWidow, укажите любой web-адрес в строке **Site URL** и нажмите на третью слева кнопку «**Explore/Scan a site**». И вы увидите, как зашевелиятся лапки вашего нового помощника, задержатся нити Паутины и на экране проявится истинное содержимое сайта, избранного на сей раз в качестве жертвы. Одно из замечательных свойств кибер-паука: порой он видит то, чего не разглядеть невооруженным взглядом.

Основные окна программы:

■ **Browser** — окно просмотра web-ресурса; используется браузер, установленный в системе по умолчанию.

■ **Structures** — схема организации внутренней структуры рассматриваемого сайта (папки и их содержание).

■ **@Emails** — адреса электронной почты, опубликованные как на данной странице, так и на страницах, стоящих выше на иерархической лестнице (например, для страницы [www.ukr.net/book](http://www.ukr.net/book) будут показаны все e-mails, «живущие» на ней и на <http://www.ukr.net/>).

■ **Ext Links** — если снять метку в **Setting ⇨ General Options ⇨ Stay Within Site(s)**, то в этом окне будут отражаться все внешние ссылки, опубликованные на страницах данного сайта.

■ **Link Errors** — сообщения о неработающих и/или ошибочных ссылках.

■ **Threads** — «нити»; отображаются те «нити паутины» (адреса внутренних ресурсов), за которые в данный момент «дергает Черная Вдова».

Итак, программа обнаружила все файлы данного сайта. Что дальше? Теперь вы можете выбрать интересующие вас файлы и (например, используя **правый клик мышью ⇨ Download...**) сохранить их на своем жестком диске. Если в **Setting ⇨ General Options** поставить галочку в разделе **Download While Scanning**, то файлы будут автоматически сваливаться на ваш диск. Еще раз (особенно при такой установке!) настоятельно рекомендуется изменить **Path: C:\PROGRA\FILES\BLACKWIDOW**, скажем на **D:\BLACKWIDOW\TRASH**.

## DiscoPump

Знаменитая «ДискоКачалка»! Самая популярная программа этого класса среди русскоязычных пользователей.

Инсталляционная строка: **Next (Далее)** ⇨ **Continue** ⇨ **Yes (Да)** ⇨ выбрать язык интерфейса (**English/Russian**), **OK** ⇨ при желании запустить программу немедленно, **Yes (Да)**.

Примечательно, что уже при первом старте программа просит нового пользователя позаботиться об основных настройках программы. Это действительно важно, поскольку последствия деятельности оффлайн браузера, при небрежном отношении к настройкам, могут быть самыми печальными.

В окне «**Миссия**», в строке «**Куда качать**» обязательно укажите папку, живущую вдали от системных дисков.

В окне «**Параметры**» не забудьте указать типы файлов, которые вы предполагаете скачать.

Чем эта программа лучше других программ, предназначенных для автоматического считывания страниц из Мировой Паутины? Для нас — это,

прежде всего, ориентированность на русский язык. Помимо возможности работать с русским интерфейсом, пользователь может также выбирать, в случае необходимости, нужную кодировку. «ДискоКачалка» предусматривает два режима просмотра результатов своей работы: во внутреннем окне и в отдельном окне браузера. Эту программу отличает богатство возможностей, предоставляемых пользователю по правому клику.

Программа следит за состоянием связи, вы можете попросить ее восстановить связь после обрыва или прервать связь по окончании перекачки.

Подробнейшую инструкцию на русском языке по работе с программой можно найти в разделе **Помощь (Help)**. На любые ваши вопросы о программе в течение 48 часов ответит служба технической поддержки.

## Offline Explorer

Самая легкая программа в своем классе.

Самая короткая инсталляционная строка: **Next** ⇨ **Next** ⇨ **Next** ⇨ **Finish**.

Интерфейсно сильно напоминает своего мейкрософтовского названного брата.

**Внимание!** По умолчанию считает основной директорией для хранения добытого материала **C:\TEMP**. Не забудьте переместить «хранилище» с системного диска: **View ⇨ Options ⇨ Advanced ⇨ Download Directory ⇨** выбрать **другое место** для перекачиваемых сайтов (например, **E:\TEMP**).

Производит довольно бледное впечатление. Тем не менее, программа достаточно популярна именно за счет своей легкости и подобности Internet Explorer.

### Web Stripper

Задача этой программки — раздеть сайт до нитки. В этом она не оригинальна.

Инсталляционная строка (**Next ⇨ Next ⇨ Next ⇨ Next ⇨ Finish**) отличается солидной стандартностью и неожиданным требованием (для программ этого класса) перезагрузить компьютер для «complete the installation».

Жизнь с программой начинается раздражающим сообщением о настоятельной необходимости («впрочем, как вам будет угодно...») подписаться на некий лист рассылки (обещают только новости от производителя и никакой рекламы; граница между этими двумя понятиями представляется весьма и весьма размытой). Есть и еще одно следствие «бесплатного характера распространения» программы — рекламный баннер, маячащий в верхней части рабочего экрана.

Следующий (необходимый) шаг — изменение адреса хранилища с **C:\Stripped**, скажем, на **E:\Musor (Options ⇨ Options ⇨ New Site ⇨ Default root folder)**.

В числе несомненных достоинств программы, прежде всего, удобство перехода для работы с перекачанным материалом в окне Windows Explorer (правый клик на **названии страницы ⇨ Explore**).

### WebZip

Как и в предыдущих случаях, последовательность действий при

инсталляции чрезвычайно проста: **Next** ⇨ **Accept** ⇨ **Next** ⇨ **Next** ⇨ **Next** ⇨ **Finish**.

Относительная легкость программы обеспечивается за счет отсутствия встроенного описания программы, за поддержкой или справкой пользователь вынужден идти на сервер компании-производителя. Это — серьезный недостаток, для того и существуют оффлайн браузеры, чтобы экономить время, проведенное онлайн.

Интерфейс самый развитый в ряду аналогичных программ, но весьма запутанный, по сравнению с другими аналогичными программами. Современный дизайн (в том числе возможность «переодеть» программу: **View** ⇨ **Options** ⇨ **Select Skin**) производит весьма симпатичное впечатление. Внешне WebZip напоминает пишущий плеер. Так и хочется поскорее приступить к работе.

Но будьте бдительны! По умолчанию WebZip использует для хранения «руды» раздел **C:\My Intranet**.

Как же быть? **View** ⇨ **Options** ⇨ **Preferences** ⇨ в **Personal Intranet** и **Zip Folders** установить, например, **D:\Temp**.

Научиться работе с программой совсем не сложно. Не забудьте, что основное назначение программы — работа в оффлайн, а не процедура накопления ресурсов. Очень удобно на ночь с помощью «расписания заданий» (**Shedule**) установить перекачку нужных сайтов, чтобы утром быстро просмотреть их содержимое. Для работы в оффлайн необходимо установить галочку в окошке **Work Offline**.

### Систематическая ошибка

Читатель, познакомившийся с предыдущей главой, вряд ли сможет не обратить внимание на рефрен: «Не забудьте изменить директорию для хранения результатов работы оффлайн браузера!» Массированное складирование файлов на системном диске, при работе под Windows, чревато сокращением свободного пространства ниже критической отметки и, как следствие, сбоями в работе Windows (рекомендуется всегда иметь не менее 50 Mb свободного места на диске **C**, аналогичные требования предъявляются к любому системному диску, если используется boot manager или иной мультисистемный режим).

Так что же, создателям оффлайн браузеров неведома эта простая истина? Поверьте, они в курсе. Они проектируют «свой склад» на С, поскольку нет никакой уверенности в том, что у пользователя есть диски с другими именами.

Не забывайте о недопустимости складирования мусора на системных дисках и о том, что процедура очистки от мусора порой оказывается более длительной и запутанной, нежели процесс «мусоронакопления».

### Бизнес в Internet

Когда-то давным-давно, в виртуальном сообществе рынка не было — если вы хотели купить книжку, надо было бежать в ближайший книжный магазин.

А все потому, что в те дни Сеть в основном состояла из сетей, финансируемых правительством, в которых открытая коммерческая деятельность была запрещена. Сейчас же сеть в основном поддерживается частными компаниями, и некоторые из них начали эксперимент с открытием электронных «магазинов» и тому подобных служб. Многие из таких магазинов содержатся книготорговцами, а службы варьируются

от доставки заказанных копий федеральных документов и до электронной рассылки хроники в надежде, что кто-то подпишется на какое-нибудь периодическое издание (напечатанное на бумаге). Некоторые компании также используют группы новостей Usenet ( из иерархии biz, относящейся к бизнесу) для рассылки пресс-релизов и информации о продукции.

Однако коммерческая деятельность остается на куда более низком уровне, чем на других сетях, таких, как CompuServe, с его Electronic Mall, или Prodigy, с рекламными объявлениями почти на каждом экране. Это частично связано с новизной и сложностью Internet как носителя коммерческой информации. Но частично это связано также с соображениями безопасности. Компании беспокоятся по поводу электронных взломщиков, которые могли бы проникнуть по сети в их систему, и вообще многим не нравится идея пересылать по Internet номер кредитной карты (сообщение электронной почты может пройти через несколько подсистем, пока достигнет места назначения). Эти страхи могут быть развеяны по мере того,

как пользователи Сети перейдут к таким средствам, как шифровка сообщений и «электронные подписи». Тем временем бизнесмены в Сети пока еще считаются чем-то вроде пионеров-освоителей Internet.

Несколько общедоступных систем и региональных сетей открыли «рыночные площади» для электронного занятия бизнесом.

Мир (World) в Бруклине, Массачусетс, в настоящее время арендует «место» для нескольких книжных магазинов и программистских фирм, а также для магазина «Игрушки для взрослых». Чтобы пролистать их предложения, свяжитесь через gopher с world.std.com. В главном меню выберите **Shops on the World** (Магазины Мира).

Msen в Энн Арборе предоставляет в распоряжение пользователя свой «Msen Marketplace» (рынок Msen), где можно найти туристское агентство и «Online Career Center» (Электронный транспортный центр), предлагающий рекламу по всей стране. Msen также предоставляет «Internet Business Pages» (Телефонный справочник организаций в

Internet) — электронный список компаний, которые ищут контактов с сообществом Internet. Доступ к Msen можно получить через программу gopher по адресу gopher.msen.com. В главном меню выберите **Рынок Msen**.

Сеть Nova Scotia Technology Network ведет службу gopher «Cybermarket» (Гиперрынок) по адресу nstn.ns.ca.

Там можно найти электронный книжный магазин, в котором можно заказать книги по электронной почте (придется ей доверить свой номер кредитной карты), и аналогичный «virtual record store» (магазин виртуальных записей). Оба эти магазина дают возможности поиска товара по ключевым словам и просмотра каталогов.

Есть и другие электронные бизнес-системы, в том числе:

■ **AnyWare Associates**

Эта компания в Бостоне ведет шлюз факс-Internet, позволяющий пересылать факсовые сообщения в любую точку мира по Internet (конечно, за деньги).  
Дополнительную информацию



запрашивайте по адресу sales@awa.com.

- **Bookstacks Unlimited**  
Этот книжный магазин в Кливленде предоставляет возможность поиска по ключевому слову в базе данных, содержащей тысячи выставленных на продажу книг. Адрес Telnet: books.com.
- **Counterpoint Publishing**  
Основным продуктом для Internet этой компании, базирующейся в Кембридже, штат Массачусетс, является индексированная версия федеральных журналов, в том числе Федерального регистра (ежедневный компендиум правительственных контрактов, законодательных предложений и т. п.). Пользователи Internet могут просматривать последние копии, но полный доступ стоит порядка нескольких тысяч долларов в год. С помощью gopher свяжитесь с enews.com и

выберите **Counterpoint Publishing** («Издательство «Контрапункт»»).

- **Dialog**  
Национальная база данных, к которой можно получить доступ через telnet по адресу dialog.com. Однако чтобы зарегистрироваться, необходимо предварительно установить себе раздел в **Dialog**.
- **Dow Jones News Retrieval**  
Телеграфное агентство, которое ведется информационной компанией — владельцем «Уолл стрит Джорнэл». Доступ через telnet по адресу djnr.dowjones.com. Как и для компании Dialog, чтобы зарегистрироваться, нужно иметь свой раздел.
- **Infinity Link**  
Просмотр каталогов книг, музыки, программного обеспечения, видеокассет и лазерных дисков. Система базируется в Малверне, Пенсильвания. Связь через

gopher по адресу columbia.ilc.com. Регистрация: cas.

- **The Internet Company**  
Эта компания, базирующаяся в Кембридже, Массачусетс, является некоторого рода сервисным бюро и работает с несколькими издательскими компаниями над продукцией, имеющей отношение к Internet. Ее электронный стенд новостей (Electronic Newsstand) предлагает краткие специальные цены для подписки на национальные журналы, от New Republic до New Yorker. Связь через gopher по адресу enews.com.
- **MarketBase**  
Можете попробовать эту систему классифицированных объявлений, разработанную этой компанией в Санта Барбаре, Калифорния. Связь через gopher по адресу mb.com.
- **O'Reilly and Associates О'Рэйли (O'Reilly)**, известный своими

книгами «Nutshell» (Ореховая скорлупа) по Unix, ведет три службы Internet. Сервер gopher по адресу ora.com. Даст информацию о компании и ее книгах. Аналогичная информация передается в группу новостей Usenet biz.oreilly.announce. Ее Навигатор Глобальной сети (Global Network Navigator), к которому можно получить доступ через Всемирную Сеть (World-Wide Web) представляет собой некоторый вид электронного журнала, который позволяет пользователям просматривать интересные их службы и каталоги.

## Вопросы и ответы

### Как перетащить bookmark из Netscape в Internet Explorer?

Это можно сделать несколькими способами:

1. Bookmark из Netscape представляет собою обычный файл (его стандартное имя — bookmark.htm), и его можно открыть в Internet Explorer, как вы открываете все прочие файлы (file-open; файл-открыть), а затем просто щёлкать по содержащимся в нём адресам (линкам). Как правило, bookmark.htm располагается в той папке, в которую вы установили Netscape, вернее, в одной из её подпапок. Если вы затрудняетесь его найти, воспользуйтесь поиском Windows (Start-Find-Files and Folders; Пуск-Поиск-Файлы и папки).

2. Если в вашем bookmark-е адресов немного, вы можете вручную превратить каждый из них в ярлык и поместить эти ярлыки в папку Favorites (Избранное), где, собственно, Internet Explorer и сохраняет свои адреса. Проще всего сделать это так. Откройте файл bookmark.htm в одном из ваших браузеров (в Internet Explorer или Netscape) и оставьте окно браузера открытым. С помощью программы My Computer (Мой компьютер) найдите на вашем диске и откройте папку Favorites (Избранное). Расположите окна браузера и My computer на экране так, чтобы видеть оба окна одновременно, а затем начните

просто перетаскивать мышью адреса (линки) из окна браузера в окно Favorites (Избранное). Попав в Favorites, ссылки Netscape превратятся в ярлыки Internet Explorer.

3. Если ваш bookmark большой и адресов в нём много, то перетаскивание линков вручную займёт много времени. В таком случае лучше воспользоваться специальной программой, которая выполнит эту работу за вас. Одна из таких программ называется Bookmark Converter. Она умеет превращать ссылки Netscape-а в ярлыки Internet Explorer-а и наоборот. Найти эту программу можно в Интернете по адресу:  
[http://www.abc.se/~m9761/bm\\_conv/](http://www.abc.se/~m9761/bm_conv/)

### **Я хочу поставить в Netscape-е иврит.**

Проблема, с которой вы столкнулись, известна давно. При попытке установить шрифты в программе Netscape Communicator (если он устанавливается в ивритской версии Windows Me) компьютер действительно намертво зависает. Есть несколько способов решения этой проблемы:

1. Прежде, чем устанавливать иврит в Netscape, откройте Control Panel (Luah ha-Bakara, Панель управления) — Fonts (Gufanim, Шрифты). Найдите шрифт под названием Times New Roman Backslanted и удалите его. После этого можете спокойно запускать Netscape и устанавливать шрифты. Компьютер больше не зависнет. Автор этого решения неизвестен, но метод многократно опробован и работает. Что такого страшного в вышеуказанном шрифте Times New Roman Backslanted, никто толком не знает, однако причина — именно в нём. Фирма Netscape никаких объяснений на эту тему не обнародовала.

2. Новое, совершенно неожиданное решение проблемы предложил Илья Дивинский (Иерусалим). Он обнаружил, что если при установке шрифта мышью использовать только в операциях Edit — Preferences — Fonts, а затем выбирать кодировку и фонты с помощью клавиатуры (рабочие клавиши: Tab, стрелки вверх/вниз, Enter), то компьютер не зависает.

3. Сотрудники компьютерного магазина «Big Bit» (Иерусалим) экспериментальным методом выяснили, что проблема решается также, если

устанавливать Netscape на «чистый» Windows, до инсталляции Microsoft Office.

Какой из этих методов вам больше подходит, решайте сами.

### **Как предотвратить зависание компьютера при установке шрифтов в Netscape Communicator?**

Установите в Windows необходимые вам шрифты (например, Web Hebrew AD и Web Hebrew Monospace). Используя мышью, откройте: edit — preferences — fonts

Начиная с этого момента, не трогайте больше мышью и пользуйтесь только клавиатурой. Порядок действий следующий:

- Tab — клавиша u — «стрелка вниз» (получили в верхней строке User Defined)
- Tab — клавиша w (высветился Web Hebrew AD)
- Tab — Tab — клавиша w (высветился Web Hebrew Monospace)
- Tab — Tab — Tab — Enter

Обратите внимание: после перехода в раздел **Variable Width Font** не следует

нажимать на «стрелку вниз» несколько раз, иначе компьютер зависнет.

Кстати, два слова и о другом методе решения проблемы (через удаление шрифта Times New Roman Backslanted). Некоторые опасаются применять этот способ, так как им жаль расставаться с соответствующим шрифтом. Эти опасения совершенно безосновательны. Если вам так дорог Times New Roman Backslanted (хотя на самом деле им почти никто не пользуется), то перед его удалением скопируйте его в какую-нибудь временную директорию (то есть сделайте его резервную копию), а после окончания установки шрифта в Netscape Communicator инсталлируйте его в Windows снова.

**Мне Outlook Express иногда предлагает послать message как Unicode или as is. Какая разница между ними? Что лучше?**

При выборе опции «as is» письмо посылается в той кодировке, которую назначили вы сами (например, Cyrillic Windows, koi8-r и др.). При выборе опции Unicode письмо (в какой бы кодировке вы его ни написали) «насильно» переводится

в так называемую универсальную кодировку.

К сожалению, не существует идеального решения вопроса, какой кодировкой лучше пользоваться при написании текстов электронных посланий. Формат Unicode хорош своей универсальностью по отношению к различным языкам общения, но при отправке письма из Outlook Express выбор Unicode может создать проблемы для тех корреспондентов, у которых не установлен Internet Explorer, MS Office или Netscape. Вместо текста они могут получить нечитабельную абракадабру. Советую вам решать вопрос о выборе кодировки письма с каждым корреспондентом индивидуально.

**Я очень признателен программе The Bat! Она позволила мне обнаружить весьма неприятный «баг» в Netscape Messenger Mailbox. После установки The Bat! я импортировал все письма из почтового ящика Netscape и, к своему ужасу, обнаружил письма, которые давно «сжѐг». Неужели невозможно очистить мусорную корзину в Netscape Mailbox?**

Действительно, такая проблема в Netscape (в отличие от Eudora или The Bat!) существует. После очистки мусорной корзины (trash) по стандартной схеме (правый клик — Delete Message) письма не удаляются полностью, а лишь становятся невидимыми. Для того, чтобы в этом убедиться, просмотрите файл Trash из директории Mail посредством текстового редактора Notepad (Блокнот). Полностью очистить мусорную корзину, не выходя из программы Netscape Messenger Mailbox, можно так: выбрать Local Mail в списке почтовых ящиков — в Netscape Message Center правый клик на корзину Trash — Folder Properties — Clean Up Wasted Space. Теперь Вы можете быть спокойны: элементарными методами Вашу уничтоженную переписку не восстановить.

**Меня интересует возможность чтения news (сейчас мне особенно интересны israel.jobs.\*) в режиме off-line (у меня доступ в Internet через телефоны 132/135). Возможно это реализовать через сайт (www.dejanews.com)? Какие на сегодняшний день существуют службы получения news посредством электронной почты?**

Режим offline позволяет вам читать сообщения, посланные в группу новостей (newsgroup), не будучи подключённым в этот момент к Интернету. После получения сообщений с сервера новостей на ваш компьютер, вы можете отключиться от сети и продолжить чтение сообщений в любое удобное для вас время, не заходя в Интернет. Работа в режиме offline позволяет сократить телефонные расходы и счета компании, предоставляющей вам услуги Интернета.

Рассмотрим, как организована работа в режиме offline в приложении Microsoft Outlook Express.

С помощью MS Outlook Express вы можете сгрузить с сервера новостей только заголовки сообщений, которые содержат тему, имя автора, размер сообщения и время его отправки. Вы также можете сгрузить полное сообщение, включая прикрепленные к нему файлы (attachments). Сгружая только заголовки сообщений, вы экономите и время, и деньги. К тому же информация, содержащаяся в заголовке, даст вам представление, о чем идёт речь в самом сообщении. Далее, работая в режиме offline, вы можете пометить заголовки

новостей, заинтересовавшие вас, для полного переноса самого сообщения на ваш компьютер с сервера новостей.

Вы можете сконфигурировать Outlook Express для переноса сообщений и/или заголовков на ваш компьютер из newsgroups, к которым вы создали постоянное присоединение (subscription). Для этого отметьте интересующую вас newsgroup в списке папок. В меню программы выберите: File-Properties. Откройте закладку Download и активизируйте контроль «When downloading this newsgroup, retrieve». После этого выберите одну из трех возможных опций:

- New headers — сгружать заголовки новых сообщений.
- New messages — сгружать заголовки новых сообщений и сами сообщения.
- All messages — сгружать все заголовки вместе с сообщениями.

Для переноса заголовков и/или сообщений на ваш компьютер выберите в меню программы Tools-Download All.

Если вы сконфигурировали Outlook Express для переноса только заголовков, вы можете отметить заголовки заинтересовавших вас сообщений, работая в режиме offline. Это поможет вам, подсоединившись к сети, полностью сгрузить желаемые сообщения. Для того, чтобы отметить заголовок, выделите его, пользуясь правой кнопкой мышки. В открывшемся меню выберите опцию «Mark message for download». Отметьте все необходимые заголовки. Подсоединитесь к Интернету. В меню программы выберите Tools-Download All.

Продолжая отвечать на вопрос, замечу, что узел Dejanews ([www.dejanews.com](http://www.dejanews.com)), созданный специально для оказания услуг пользователям Usenet не поддерживает режим offline на том уровне, как это реализовано в Outlook Express. Доступ к ресурсам Dejanews происходит посредством протокола HTTP, поэтому эффективность режима offline при работе с Dejanews зависит от того, насколько хорошо он реализован в вашем браузере. Outlook Express, который пользуется протоколом NNTP для доступа к ресурсам Usenet, создавался специально, чтобы

обеспечить максимум удобства при работе с news.

В настоящий момент мне неизвестно о существовании служб получения news посредством электронной почты. Технически такое возможно. В любом случае, оставляя сообщение в newsgroup вы можете попросить, чтобы вам ответили прямо на адрес e-mail.

### Поиск работы через Internet

Для начала вполне подойдут общеизвестные международные поисковые службы типа **Yahoo** ([www.yahoo.com](http://www.yahoo.com)) или **Excite** ([www.excite.com](http://www.excite.com)). Любая из них с радостью выдаст несколько сотен тысяч ссылок в ответ на ключевое слово «**job**». Тут уж следует быть осмотрительным и хотя бы немного владеть английским. Дело в том, что среди огромного количества ссылок, действительно подходящих — примерно процентов десять, остальное составляют описания контор, услугами которых вам вряд ли когда удастся воспользоваться.

Дело в том, что, как показывает практика (и подтверждается специально проведенными исследованиями), около

восемидесяти процентов рабочих мест распределяется посредством друзей и знакомых, и лишь двадцать процентов вакансий занимают через всякие объявления о найме. Таким образом, из сказанного вытекают два следствия. Во-первых, чем активнее человек, чем больше у него знакомых, тем выше его шансы найти наиболее высокооплачиваемую работу. Во-вторых, развитие систем всяких глобальных бюро, обросших прочными связями с нанимателями, в значительной степени влияет на перераспределение рабочих вакансий между конкурсным набором и рекомендациями друзей и знакомых.

Примером такой компании может служить **CFG Inc** ([www.cfg-inc.com/candidate\\_frames.htm](http://www.cfg-inc.com/candidate_frames.htm)). Неизвестно, сколько лет она зарабатывает себе комиссионные на предоставлении желающим всяких вариантов, но база у нее действительно впечатляет. Причем, вакансии самые разные, далеко выходящие за рамки обычных компьютерных специальностей. Хотя, если хочется найти не вообще работу, а в конкретной области, то и поиск лучше сразу конкретизировать. Благо, Internet такое позволяет.



Можно обратиться, например, в специальную поисковую службу по профессиям **JOBCENTER** ([www.jobcenter.com/search/default.asp](http://www.jobcenter.com/search/default.asp)). Система **JOBCENTER** дает возможность отыскать варианты, удовлетворяющие местами достаточно жестким требованиям (регион, специальность, условия, доход и так далее). Результатом работы этой поисковой службы является список заявок о найме сотрудников, с указанием контактных данных, позволяющих уже самостоятельно разослать свои заявления. Конечно, **JOBCENTER** не гарантирует, что вы обязательно получите эту работу, но одно дело искать варианты неделями, а другое — получить за три-четыре минуты несколько сотен конкретных, актуальных адресов. Тем более, что они принимают заявления и по электронной почте, что и недорого, и удобно, и достаточно быстро.

Хотя, если вы специалист по какой-нибудь области, перечень на слово «**job**» имеет смысл отсортировать по вашей профессии, так куда удобнее. В Сети есть достаточно специализированных мест, рассчитанных на конкретную профессию (если она широко универсальна, например, водители) или на четко

обозначенную группу профессий. К примеру, в Internet есть специализированная поисковая служба для желающих найти себе место в финансовой Америке. Она располагается на сервере «**ВСЯ АМЕРИКА**» ([www.allmerica.com/cs/jobs/function.htm](http://www.allmerica.com/cs/jobs/function.htm)) и позволяет найти новое место всяким финансовым консультантам, менеджерам по ценным бумагам, страховым агентам и так далее, и тому подобное. Это место сразу бросается в глаза своей обстоятельностью. Огромное множество профессий и направлений, объединяемых глобальным термином «финансы», четко подразделяются на группы и подгруппы, позволяющие максимально точно определить интересующую область приложения своих знаний и навыков.

Предоставляемая службой информация ([www.allmerica.com/cs/jobs/98A7098.htm](http://www.allmerica.com/cs/jobs/98A7098.htm)) достаточно кратка, но в то же время предельно конкретна. В переводе на «наши» понятия, она выглядит примерно так: требуется сотрудник для выполнения таких и таких операций, владеющий следующими навыками, например,

контактный, если речь идет о работе из области контактов с общественностью.

Если же у претендента есть некоторые оригинальные идеи, реализация которых может, если не перевернуть мир вообще, то принести большие дивиденды, — самое время обратиться на своеобразную ярмарку идей ([www. ideasjn.com](http://www.ideasjn.com)), где можно предложить свою в общих чертах и найти компанию, которая заинтересуется ею.

Правда, просить там себе пару сотен тысяч долларов на развитие системы утилизации бытового мусора в Богом забытой деревне бесполезно, как бесполезно баламутить мировую общественность очередной конструкцией вечного двигателя.

Честно говоря, «они» даже в передовых идеях достаточно консервативны и проявляют интерес лишь в довольно ограниченных направлениях. Однако все же простора для фантазии тут достаточно много.

Internet удобна еще и тем, что позволяет делать массу дел параллельно. В том числе и искать работу. Например, в одном окне обозревателя подыскивать

что-нибудь по имеющейся профессии, а в другом — просматривать предложения по трудоустройству, например, в США вообще ([www.psjobfair.com/ojf.html](http://www.psjobfair.com/ojf.html)).

Однако не стоит забывать, что поиск работы, это не прогулка по парку в погожий день ради собственного удовольствия. Поиск работы, простите за тавтологию, это тоже весьма не простая работа. Дело в том, что мы, испорченные единой тарифной сеткой, слабо себе представляем принцип оплаты труда в той же Америке. А тонкостей в этом деле ничуть не меньше, чем в хождении под парусом. То, сколько претенденту будут платить за его работу, зависит не только от его квалификации или среднего уровня оплаты труда в данной области. Тут играет роль масса факторов, в том числе схема местных налогов, уровень доходов на данной территории и так далее. Не редкостью в Америке является ситуация, когда вроде бы скромный доход оказывается на самом деле самым высоким, потому что и налоги меньше, и цены в местных магазинах ниже, и аренда жилья дешевле.

С другой стороны, приступая к такому совершенно естественному для

американцев действию, как торговля за величину зарплаты в контракте, вовсе не лишне знать «местные условия», иначе либо вы сами согласитесь трудиться почти даром, либо вам сходу откажут в месте по причине слишком высоких, по мнению нанимателя, исходных требований.

Чтобы представлять, в каких условиях вам предстоит жить и работать, полезно поинтересоваться в Сети о тех местах, где расположена та компания, место в которой вас интересует. Допустим, если это город Чикаго ([www.psijobfair.com/chicago.html](http://www.psijobfair.com/chicago.html)), то кое-какое представление можно получить, если узнать, что в среднем месячный семейный доход его жителей составляет \$34,682.00, что приблизительно шестьдесят шесть процентов жителей — белые, двадцать два — черные, двенадцать — испанцы. Кстати, Чикаго весьма перспективен с точки зрения поиска женихов и невест, так как четвертая часть населения (24,4 процента) в возрасте двадцати четырех лет и старше не состоят в браке. Для справки, этот показатель для США в целом составляет лишь 20,2 процента.

Используя Всемирную Паутину, можно вообще очень много узнать о

любом месте на свете. О том, как там живут люди, что они предпочитают, что не любят и насколько эта жизнь удобна и интересна (а также доходна). Пренебрегать такими данными, принимая действительно серьезное решение, попросту глупо. Тем более, что для средне эрудированного человека довольно многое может сказать даже статистическая таблица распределения жителей по профессиональным категориям. Например, если сказано, что в некотором населенном пункте более сорока процентов юристов и страховых агентов, то, скорее всего, это достаточно тихое и спокойное место, в котором живут хорошо обеспеченные люди. Стало быть, если вам предлагается место финансового агента или управляющего капиталом, то перспективы в подобном месте вполне радужны. А вот адвокатом или страховым агентом в подобном месте лучше не наниматься, если только вы не выдающийся гений, способный уболтать кого угодно на что угодно.

А если Труффальдино из Бергамо с вас писали, то можно попытаться счастья и напрямую, без всяких посредников и агентов. Ведь хотите вы того или нет, а в

Штатах комиссионные — дело святое, и если кто-то кого-то свел, то обязательно заплатит за это. Наверняка стоимость размещения объявления о найме будет под благовидным предлогом высчитано из вашего контракта. Таким образом, если есть стойкое ощущение, что вам сам черт не брат, то можно сразу обратиться к списку корпоративной Америки ([www.yahoo.com/Business\\_and\\_Economy/Employment/Jobs/Company\\_Job\\_Listings](http://www.yahoo.com/Business_and_Economy/Employment/Jobs/Company_Job_Listings)), так как там почти всегда кто-то нужен. Правда, сами американцы признают, что в Соединенных Штатах львиную долю рабочих мест дают не большие корпорации (менее десяти процентов занятых), а малые и средние предприятия, на которых работают около семидесяти процентов трудоспособного населения.

Тонкость этого дела заключается в том, что эти наниматели зачастую не могут себе позволить содержать так называемое виртуальное бюро по трудоустройству.

Многие работодатели и те, кто хочет найти себе новое место работы, прибегают к услугам сети Internet. Вообще говоря,

найти работу посредством Internet несложно. Вот получить ее — это уже другое дело, индивидуальное, зависящее только от вас и вашего везения. Поэтому далее будем рассматривать только непосредственно поиск нового рабочего места, причем не только в России, но и за рубежом.

Как показывает даже самый поверхностный взгляд, поиск или помощь в нахождении рабочего места давно превратились в своеобразную отрасль бизнеса. Сотни и сотни контор, бюро и служб занимаются поиском и сведением нанимаемых и нанимателей.

Рано или поздно проблема трудоустройства становится перед любым из нас. И каждый решает ее по-своему. Работу можно найти следующим образом:

- с помощью родственников и знакомых;
- по объявлениям в средствах массовой информации;
- обратившись в кадровое агентство.

Из вышеперечисленных пунктов многим знакомы первые два. К третьему из них прибегают, как правило, в последнюю очередь. Одна из причин этого явления — незнание того, как решить проблему выбора, в какое кадровое бюро (КБ) обратиться. Очень трудно определиться, какие из КБ предоставляют действительно качественные услуги.

С другой стороны, аналогичная проблема стоит и перед работодателем — как выбрать фирму, которая действительно обеспечит должный уровень сервиса, помогая найти необходимого специалиста.

Таким образом, возникает реальная проблема поиска критериев оценки качества работы КБ.

Мы предоставляем вашему вниманию свой первоначальный вариант, в надежде, что он поможет вам лучше ориентироваться на рынке кадровых услуг.

Итак:

- Опыт работы на рынке трудоустройства не менее 12 месяцев. Это означает, что у фирмы уже есть некоторый опыт работы на

рынке трудоустройства, что она не новичок.

- Средняя длительность удовлетворения заказа от работодателя — не более 8 рабочих дней. Это означает, что КБ может достаточно быстро найти нужного специалиста для фирмы-заказчика.
- Сумма месячных заработных плат трудоустраиваемых в течении месяца — не менее \$1000. Смысл этого критерия в том, что КБ ведет реальную работу и регулярно трудоустраивает или много неквалифицированных специалистов, или 1-2 профессионалов.
- Предоставление услуг по кадровому консалтингу. Выполнение этого требования означает, что фирма-заказчик может всегда надеяться на предоставление КБ, при желании, дополнительных услуг, на его помощь и

поддержку при решении разнообразных возникающих кадровых проблем, не решаемых методом простого увольнения/найма.

- Использование компьютерной обработки данных. В современном мире только компьютеризация обеспечивает необходимую скорость и полноту поиска нужной информации, и, следовательно, КБ без компьютеров — это КБ вчерашнего дня.
- Стоимость услуг лежит в пределах экономически обоснованного ценового коридора. Это означает честность КБ в его поведении на рынке — не заниматься демпингом и не «обдирать» клиентов.
- Предоставление услуг по профессиональному тестированию и оказание содействия по повышению квалификации. Такого рода

услуги позволяют сегодняшнему безработному ожидать скорого получения работы, даже если пока у него нет достаточной квалификации — ведь ему помогут приобрести необходимые знания. С другой стороны, фирма-заказчик, которая сотрудничает с таким КБ, может не волноваться, что ей пришлют «программиста» с 8 классами образования или «переводчика», умеющего лишь красиво произносить два десятка популярных фраз.

- Проведение психологического тестирования (в частности — с использованием соционических методик) при осуществлении индивидуального подбора специалистов. Это позволяет фирмам заказчикам быть уверенными, что подобранный КБ специалист не только имеет высокий профессиональный уровень, но и психологически склонен к выполнению работ

требуемого характера, без конфликтов впишется в уже существующий коллектив и т.п.

### **Анкор**

<http://www.ancor.ru/Center/>

На сайте представлены вакансии Московского и С.-Петербуржского отделений агентства, советы по составлению резюме. Имеется возможность отправить резюме.

### **Арсенал**

<http://www.arsenal-moscow.ru>

Информация о фирме, открытых вакансиях, советы по трудоустройству. Технология работы кадровых агентств под ключ.

Основные направления работы агентства: бухгалтеры, экономисты, банковские служащие, временный персонал, повара в рестораны \*\* и \*\*\*, инженерно-технические работники, макетисты-верстальщики, дизайнеры, маркетологи, специалисты в области рекламы и PR, менеджеры всех уровней и направлений, обслуживающий персонал, специалисты в области полиграфии и издательской деятельности, специалисты в

области финансов, юристы, администраторы локальных вычислительных сетей.

### **Триза**

<http://www.triza.ru>

Информация об агентстве, доска объявлений, резюме и открытые вакансии агентства.

Широкопрофильное агентство, занимающееся подбором персонала, а не трудоустройством.

### **Рикор-Электронникс**

<http://www.rabota.ricor.ru>

На сайте предоставлены следующие материалы: информация об открытых вакансиях агентства «Рикор-Электронникс» и других кадровых агентств, анализ рынка труда, информация о том, где получить новую профессию, как правильно заполнить резюме и куда обратиться за помощью. Работает бесплатная база данных «Анкета ищу работу». Постоянно присутствует около 500 актуальных вакансий кадровых агентств

### **Enter**

<http://www.arsenal-moscow.ru>

На сайте предоставлена общая информация об агентстве, тематические обзоры, открытые вакансии.

Основные направления работы: компьютеры и телекоммуникационные технологии, допечатные процессы.

### **Center Expert**

<http://www.inior.ru/centerexpert/>

Информация об агентстве, текущие вакансии и резюме специалистов, аналитические статьи. Возможность внесения в базу данных резюме для соискателя и работодателя.

Компания предоставляет базу данных агентств по подбору персонала, составление грамотного резюме и рассылка по электронной почте и факсам. Основной акцент работы Агентства делается на специалистов, в совершенстве владеющих иностранными языками.

### **Профиль**

<http://www.profil.ru>

На сайте предоставлены текущие вакансии.

Имеется возможность отправки резюме по электронной почте.

### **Биржа труда**

<http://www.adam.ru/Birza/>

Информация о текущих вакансиях и банк резюме. Возможность добавления заявки на специалиста и внесения резюме в базу данных

### **Евроменеджмент**

<http://home.ptt.ru/jobagency/>

Информация об агентстве, список вакансий.

### **Служба «Подбор персонала»**

<http://www.bax.labor.ru/>

Новости, услуги, добавление резюме и заявок, экспресс вакансии, советы.

### **Консалтинговая группа Альфа**

#### **Биржа труда от InfoArt**

<http://www.infoart.ru/it/>

Список резюме и предложений. Возможность размещения своего объявления.



### **Шанс**

<http://www.interset.ru/resume/>

Поиск в базе данных, заполнение резюме.

### **Комильфо**

<http://www.orc.ru/comilfo>

Сайт посвящен найму персонала: нянь, гувернеров, преподавателей, репетиторов, горничных, домработниц, поваров.

### **Russian Internet Job Agency**

<http://www.job.ru>

Крупнейший Российский сайт, посвященный работе. Здесь вы можете как добавить запись, так и осуществить поиск по базе вакансий и резюме. Приведена информация из телеконференций и предложения агентств. Большой список тематических ссылок.

### **Работа для всех**

<http://www.jobs.ru>

Есть возможность просмотреть, добавить и редактировать записи в базе данных по резюме и вакансиям. Также имеет место рейтинг вакансий и резюме. Вы можете ознакомиться со статистикой

(кто? когда? просматривал ваше объявление). Есть возможность на каждую вакансию отправить резюме (информация вносится в базу сайта).

### **Point Job Russia**

<http://www.pointjob.ru/>

Просмотр и публикация вакансий, поиск специалистов и заполнение резюме. Есть возможность online/offline публикации. Удобный интерфейс заполнения форм (вам не приходится каждый раз вносить общую информацию, т.к. она хранится на сервере). Советы по заполнению резюме.

### **Работа для Вас**

<http://www.rdw.ru>

Объявления о трудоустройстве, статьи и информационные материалы. Каждую неделю «Работа для вас» публикует в номере понедельника более 6000, а в номере четверга более 3000 вакансий по более чем 200 основным профессиям, а также сотни предложений по обучению. «Работа для вас» распространяется в Москве и Московской области.

### **Работа сегодня**

<http://www.job-today.ru/>

На сервере помещаются вакансии из объявлений, размещенных в последних выпусках газеты. На сервере объявления появляются в субботу. Еженедельная московская газета «Работа сегодня», публикующая объявления о вакансиях и мини-резюме людей, ищущих работу. Первый день распространения газеты — понедельник. Особенности аудитории издания: газета ориентирована на людей, имеющих конкретную специальность и квалификацию, пользующуюся спросом на рынке труда.

### **Работа за рубежом**

Не стоит надеяться «на авось» и выезжать за границу без железных гарантий получения работы. Попытка трудоустроиться нелегально тоже не самый лучший способ решения проблемы. Чудес не бывает, и просто так, «на блюдечке с голубой каемочкой» работу вам никто не предложит.

Помимо хорошего языка (кстати, проверить свой английский вы можете бесплатно на страницах <http://www.toefl.org/>), контактности и

неконфликтности при работе в коллективе, требуется как минимум двухгодичный подтвержденный опыт работы.

На данный момент в сети много объявлений о предложении работы за рубежом только для квалифицированных программистов и медиков. Не будем упоминать ученых, ибо если талантливый научный сотрудник еще здесь, то значит за граница ему не нужна. Представителям остальных профессий остается только ждать, когда в далекой Америке или не совсем далекой Европе вдруг неожиданно вырастет спрос, к примеру, на приезжих архитекторов или учителей.

Вы можете найти для себя компанию самостоятельно, либо через агентства по трудоустройству. Во втором случае с вас возьмут от 1000 до 2000 долларов за различные услуги и 5-15% от вашей зарплаты с работодателя.

Если же вы решили искать работу самостоятельно, можете воспользоваться услугами известных поисковых систем. Но это очень длительный процесс. Опыт показывает, что например поиск фразы «Н1В» в 90% выдает текст объявления

типа «Sorry, no H1B Visa candidates». Кстати, прежде чем начать поиски, советуем прочитать материалы, которые поступят в ваш адрес от рекрутера.

Для тех, кто не любит часами копаться в поисковиках, предлагаем список фирм (пока только США, Канады и Великобритании), набирающих к себе персонал из-за границы.

### **Памятка по составлению RESUME**

1. Форматируйте параметры страницы по американским стандартам: размер бумаги — Letter, поля:

- верхнее — 2.03 см
- нижнее — 2.29 см
- левое — 2.54 см
- правое — 2.03 см.

2. Нумеруйте страницы, не ставя номера на первой.

3. В Employment History указывайте данные в стандартном американском порядке: Дата (месяц, число, год) — Подразделение — Должность — Компания — Город — Страна. Адрес компании в России указывать не надо.

4. Свой домашний адрес, телефон, факс и E-mail в России не указывайте. (Ваши интересы перед работодателем будет представлять CompuPlus Resources USA, LLA, чьи реквизиты впоследствии и будут указаны в вашем резюме). Не указывайте и дату Вашего рождения.

5. Описательная часть Ваших служебных обязанностей и профессионального опыта должна быть достаточно подробной. Но пишите только существенное. Знание Вами, например, MS Office или Corel Draw не интересует работодателя.

6. Само резюме должно иметь объем от 1,5 до 3 страниц.

7. Согласно американской традиции, в любых документах сначала пишется имя и только потом фамилия (иначе американец не поймет, как же Вас реально называть). Имя и фамилию следует указывать непременно в варианте Вашего загранпаспорта.

8. В шапке не следует писать слов: Name, Phone, E-mail, CV, RESUME и т.д.

9. Ярко и выпукло обозначайте приоритетные области своего

профессионального опыта. Имейте в виду: американцы имеют привычку отбраковывать резюме по ключевым словам.

10. Даты старта и окончания работ ставьте в начале строки, а не в ее конце. Поменьше строчек выделяйте **Bold**, пользуйтесь чаще выделением Underline.

11. В разделе Education следует указывать Major, как основное направление учебы и Minor — как дополнительная специализация.

12. Резюме необходимо составлять сжато, писать только о том, что делал сам (не «мы», а «я»). Резюме — это текущий слепок Вашего опыта с акцентом на современные программные продукты. Перечислять языки программирования, операционные системы и прочее следует от более важных (для Вас) к менее важным.

13. Форматируйте, пожалуйста, текст с выравниванием по правому и по левому краю!

14. Называйте файлы с резюме первыми буквами Вашей фамилии (например, sidorov.doc, popova.doc). При составлении Резюме особое внимание

обратить (особо ярко и выпукло подать) на следующие разделы Вашего вероятного профессионального опыта.

## Приложения

### Глоссарий наиболее частых фраз и слов, на которые вы, вероятно, «наткнетесь» в Internet

#### **ANSI**

Компьютеры используют различные методы для представления информации на экране и для определения способа взаимодействия экрана и клавиатуры. ANSI — это один из методов подобной «эмуляции терминала». Он более всего популярен в системах досок объявлений (BBS) на основе PC, но используется и в некоторых подсистемах Сети. Для правильного его использования сначала надо его установить, или активизировать, в своем коммуникационном пакете.

#### **Application (Приложение)**

Программное обеспечение, выполняющее какую-либо конкретную функцию для пользователя, например, ftp. Сама эта полезная пользователю функция, например, пересылка файлов.

#### **Archie**

Система для определения местонахождения файлов доступных по анонимному ftp.

#### **ARPA**

The Advanced Research Projects Agency — бюро проектов передовых исследований. Подразделение DoD.

#### **ARPANet**

Предшественник Internet. Начата в 1969 году с финансированием от Defense Department's Advanced Projects Research Agency. (Агентство перспективных исследовательских проектов министерства обороны).

#### **ASCII**

Имеет два значения. ASCII — это универсальный компьютерный код для английских букв и символов. Компьютер хранит информацию в виде двоичных чисел. В коде ASCII буква «А» хранится как 01000001, независимо от вида компьютера — будь он сделан IBM, Apple или Commodore. ASCII также означает метод, или протокол для копирования файла из одного компьютера в другой по сети, в котором ни один компьютер не

ведет проверку на ошибки, могущие возникнуть из-за статического электричества или других причин.

**b**

bit — бит.

**B**

byte — байт.

### **Backbone (Хребет, хребтовая сеть)**

Высокоскоростная сеть, объединяющая несколько мощных компьютеров. В США хребтом Internet часто считается NSFNet, — финансируемая правительством сеть, объединяющая несколько суперкомпьютеров во всей стране.

### **Baud (бод)**

Скорость передачи данных модемом. Один бод — это, грубо говоря, один бит в секунду. Для передачи буквы или символа нужно восемь бит. Модемы редко передают данные с указанной в паспорте скоростью из-за статических помех или помех в компьютере. Более дорогие модемы используют системы типа Microcom Network Protocol (MNP), которые исправляют такие ошибки или

уплотняют данные для ускорения передачи.

### **BBS**

Broadcast Bulletin System — широковещательная система объявлений. Электронная доска объявлений, компьютерный аналог доски объявлений. Обычно — это компьютер с которым можно связаться по телефонной линии с использованием модема, где можно оставить свое сообщение, прочитать сообщения других пользователей, а также скопировать себе свободно распространяемое программное обеспечение, внести свой вклад в это хранилище и т.д. Изначально — некоммерческая система.

### **BITNet**

Другая система академической ориентации, международная компьютерная сеть, использующая для передачи данных другой набор компьютерных инструкций. К ней легко получить доступ из Internet по электронной почте, и в ней много конференция и баз данных. Ее название происходит от «Because It's Time» («Поскольку самое время»).

### **Bounce (Возврат)**

Это то, что делает ваша электронная почта, когда не может найти получателя — возвращает вам обратно — если не выбрасывает письмо в эфир, где оно пропадает навеки.

### **Bps**

bits per second — бит в секунду.

### **Bridge (Мост)**

Совокупность компьютера, программного обеспечения и разновидности модема (с большей пропускной способностью). Используется, чтобы соединять сети, использующие разные протоколы. Сети могут быть территориально удаленными. Мост является разновидностью шлюза, но обеспечивает более простое взаимодействие. Работает на канальном уровне модели OSI. Мост часто определяется как устройство уровня MAC — управления доступом к среде. Он никак не интерпретирует передаваемую им информацию. Две сети, связанные мостом, становятся в сущности единой сетью. Для борьбы с проблемами, возникающими из-за ограниченности пропускной способности и отказов различных

элементов сети, разработаны различные методики балансировки. Мосты все шире используются для управления потоком сообщений в сети и в целях безопасности. Они могут отфильтровывать из передаваемых данных сообщения местного масштаба, так чтобы внешние к этой сети пользователи не были допущены до них. Это не только повышает производительность сети, но и снижает опасность взлома.

### **CCITT**

Comitee; Consultatif International de Telegraphique et Telefonie — международный консультативный комитет по телеграфии и телефонии. Это комитет Международного Телекоммуникационного Союза (ITU), который, в свою очередь, является частью ООН. В задачи этого комитета входит разработка и утверждение стандартов международных сообщений.

### **CERN**

Centre Europeen des Recherches Nucleaire — Европейский центр ядерных исследований. Женева. Швейцария. Центр исследований по физике элементарных частиц. Организация с долевым участием ряда европейских стран.

**CERT**

The Computer Emergency Response Team — группа реагирования на компьютерную опасность. Государственная (бюджетная) организация США, занимающаяся проблемами компьютерной безопасности, как сетевых, так и отдельных компьютеров.

**Client**

Клиент.

**Command line (Командная строка)**

В местных системах Unix это место, где вы вводите команды, говорящие машине, чего вы хотите.

**Communications software  
(Коммуникационный пакет)**

Программа, говорящая модему, как ему работать.

**CSLIP**

Compressed SLIP — сжатый SLIP. Вариация SLIP. Обеспечивает большую пропускную способность на медленных устройствах последовательной передачи данных. Это достигается применением сжатых по Ван Якобсону (LBL (Lawrence)) TCP-заголовков.

**Daemon**

Вообще-то безвредная программа Unix, работа которой обычно незаметна для пользователя. В Internet вы, скорее всего, встретитесь с ней только когда ваша электронная почта не может быть доставлена адресату — вы получите свое исходное сообщение и к нему уродливое дополнение от «mailer daemon».

**data link**

Канал связи.

**Datagram (Дейтаграмма)**

Иногда, датаграмма. Пакет (кадр) информации, посылаемый компьютеру-получателю независимо от других пакетов — без установки логического соединения и квитирования, то есть, безо всякого предупреждения. В этом смысле дейтаграмма похожа на телеграмму: это самодостаточное сообщение, оно может прибыть в любое время без предупреждения. Дейтаграммы обычно используют в случаях, где информация пересылается нерегулярно, нечасто и небольшими порциями.



**DCE**

Data Circuit-terminating Equipment — оборудование, стоящее в конце линии передачи данных. Оно размещается между сетью и DTE. Его задачей является установление, поддержание и разрыв связи для передачи данных между машинами. Поддержание связи может включать средства кодирования и интерпретации передаваемых данных. Например, модем является DCE. Оборудование сетевого поставщика — тоже DCE.

**DDN**

Defense Data Network — оборонная информационная сеть. Часть Internet, которая соединяет военные базы США и их подрядчиков. Используется для сообщений, не требующих секретности. Принимает участие в финансовой поддержке NIC, где собран большой объем Internet информации.

**DEC**

Digital Equipment Corporation (часто просто digital). Одна из ведущих компаний по производству вычислительной техники и программных систем.

**DECnet**

DEC-версия локальной сети типа Ethernet: набор собственных сетевых протоколов, используемых операционными системами фирмы DEC вместо TCP/IP. С Internet эти протоколы несовместимы.

**Dedicated access (Непосредственный доступ)**

Доступ в Internet с максимальными пользовательскими правами и возможностями. В этом случае вы работаете прямо на сетевой рабочей машине (хосте).

**Dedicated line, leased line (Выделенная линия)**

Линия, постоянно включенная между двумя точками. Выделена в полное и безраздельное пользование. Обычно используется для подключения локальной сети средних размеров через компьютер поставщика в сеть.

**Dial-up (Вызов)**

Набор номера на телефоне и соединение с абонентом. Установление коммутируемой символьно-ориентированной асинхронной связи;

создание канала последовательной передачи битовых данных.

### **Distribution (рассылка)**

Способ ограничить места рассылки вашего послания на Usenet. Удобно для помещения объявлений типа «продается» или для обсуждения местных политических вопросов.

### **DNS**

Domain Name System — доменная система имен. Распределенная система баз данных для перевода компьютерных имен (типа club33.ora.bolizm) в числовые адреса Internet (типа 33.33.33.33). DNS позволяет пользоваться сетью, не забивая голову длинными цепочками чисел. Сетевой трафик, для этой работы занимает около 1% от всего в сети.

### **DoD**

Department of Defense —  
Министерство Обороны США.

### **Domain (Домен)**

Последняя часть адреса Internet, например, «news.com.». Домен, регион, зона, группа.

### **Dot file (файл с точкой перед именем)**

Файл в общедоступной системе Unix, который меняет способ взаимодействия вашего сообщения с системой. Например, файл .login содержит различные параметры для таких вещей, как текстовый редактор, который вызывается при посылке вами сообщения. При использовании команды ls эти файлы не появляются в списке директории, для их вывода воспользуйтесь командой ls -a.

### **Dot**

Когда разговариваете с ветераном Сети, если хотите произвести на него впечатление, называйте точку словом «dot», а не «period», например: «Мой адрес — john at site dot domain dot com.»

### **Down («упала»)**

Когда общедоступная система останавливается из-за технической неисправности и недоступна, она «упала».

### **Download**

Загрузка «вниз» — из центрального компьютера в периферийный. Копирование файла из местной системы в ваш компьютер. Есть несколько различных методов, или протоколов, для загрузки файлов, большинство из которых периодически проверяют файл, копирование которого они ведут, для гарантии неповрежденности и неискаженности информации в процессе передачи. Некоторые, такие, как XMODEM, позволяют передавать только один файл за один сеанс. Другие, такие, как пакетный YMODEM и ZMODEM, позволяют ввести имена нескольких файлов одновременно, для которых загрузка выполнится автоматически.

### **DTE**

Data Terminal Equipment — оконечное информационное оборудование. Оборудование пользователя, подключаемое к сети. Это может быть как просто терминал, так и большая ЭВМ. DTE и DCE могут объединяться в одном устройстве, как, например, персональный компьютер со встроенным модемом.

### **Duplex (Дуплексный)**

Позволяющий передавать информацию в двух встречных направлениях одновременно.

### **EMACS**

Стандартный текстовый редактор Unix, предпочитаемый Unix-овцами и ненавидимый начинающими.

### **E-mail**

Электронная почта — способ пересылки частных сообщений другому абоненту Сети. В английском языке используется и как существительное, и как действие.

Стандартный сервис Internet, реализующий аналог обычной почты; предоставляет полный спектр возможностей, доступных при пользовании обычной почтой, а также множество дополнительных удобств. Превосходит обычную почту по скорости на несколько порядков. Во всех отношениях удобнее и предпочтительнее традиционной. Единственный недостаток — слабая конфиденциальность, что, впрочем, также преодолимо. Также,

электронное письмо — сообщение (послание) электронной почты.

### **Ethernet**

«Эфирная (небесная) сеть». Тип локальной сети. Хороша разнообразием типов проводов для соединений, обеспечивающих пропускные способности от 2 до 10 миллионов bps (210 Mbps). Ethernet характеризуется способом принятия решения о порядке предоставления линии сети под управление подключенным рабочим станциям. Рабочие станции, получив управление, производят проделывают нужную им работу, например, пересылают электронную почту на узел сети. Довольно часто компьютеры, использующие протоколы TCP/IP, через Ethernet подсоединяются к Internet.

### **F2F**

Лицом к лицу. Фактическая встреча с кем-то, с кем вы обменивались письмами или наездами.

### **FAQ**

Frequently Asked Question — часто задаваемый вопрос. Или часто задаваемый

вопрос, или список таковых с приложением ответов.

(ЧаВО — Частые Вопросы и Ответы). Списки FAQ (FAQs) содержат многие группы новостей USENET и некоторые не USENET телеконференции (mailing lists) с целью избавить участников от угнетающего нудного занятия — отвечать на одни и те же вопросы, постоянно задаваемые новичками.

Многие группы новостей Usenet содержат такие файлы, которые раз в месяц или около того посылаются на сеть для новичков.

### **1 Film at 11**

Одна из реакций на слишком эмоциональный спор: «Предсказана неминуемая гибель Сети. Фильм начнется в 11.»

### **Filter (Фильтр)**

Средства, обеспечивающие связь двух однотипных локальных сетей. Фильтр выполняет прием пакетов из одной локальной сети, их буферизацию и пересылку в другую локальную сеть.

### **Finger**

Программа Internet, дающая возможность узнать кое-что о другом пользователе, если он предварительно создал файл .plan.

### **Flame (Наезд)**

Электронное рычание или ругань в чей-то адрес. Часто порождает войну наездов, которая иногда превращается в священную войну.

### **Followup (Отклик)**

Послание на Usenet в ответ на какое-то более раннее послание.

### **Foo/foobar**

Нечто вроде алгебраического символа, например: «Если вы интересуетесь, какая другая система ведется коммерческой компанией, ищите адрес в виде foo@foobar.com.»

### **Fortune cookie (Предсказание судьбы)**

Пустой, остроумный или глубокий комментарий, иногда находимый на сети.

### **Frame**

Кадр, пакет данных.

Конверт. При передаче данных — дополнительные управляющие биты, обрамляющие информационные биты кадра сообщения.

### **Freenet**

«Свободная (бесплатная) сеть». Способ организации предоставления свободного доступа в Internet на определенной территории, обычно через библиотеки. Представляет собой локальную сеть.

### **Freeware**

Бесплатное программное обеспечение.

### **FTP**

File Transfer Protocol — протокол передачи файлов. Система передачи файлов по Сети.

Протокол, определяющий правила пересылки файлов с одного компьютера на другой. Протокол уровня представления данных.

Прикладная программа, обеспечивающая пересылку файлов согласно этому протоколу.

**Gateway (Шлюз)**

Станция связи с внешней или другой сетью. Специализированный узел (станция) локальной (или специализированной) сети, обеспечивающий доступ узлов данной (локальной и т.д.) сети к другой (внешней) сети передачи данных. То есть совокупность аппаратных и программных средств, которая передает данные между несовместимыми сетями или приложениями. Типичный шлюз включает средства разборки/сборки пакетов и преобразования протоколов. Эта компьютерная система переформатирует данные при передаче так, чтобы они стали приемлемы для получателя — сети или приложения. Может обеспечивать, как связь двух несовместимых сетей, например DECnet и Internet, так и взаимодействие несовместимых приложений в рамках одной сети, например почтовых систем с различными форматами сообщений. Шлюзы работают на уровнях модели OSI: транспортном, сеансовом, уровне представления данных и прикладном. Термин часто (и неверно!) используется как синоним узла или маршрутизатора.

**Get a life (Поживи малость)**

Что сказать тому, кто, возможно, провел слишком много времени перед компьютером.

**GIF**

Графический формат обмена. Формат, разработанный в середине восьмидесятых годов в CompuServe для передачи графических образов с фотографическим качеством. Теперь часто используется при электронной передаче графики.

**GNU**

Это не Unix. Проект Организации бесплатного программного обеспечения (Free Software Foundation) для написания бесплатной версии операционной системы Unix.

**Gopher**

Интерактивная оболочка для поиска, присоединения и использования ресурсов и возможностей Internet. Интерфейс с пользователем осуществлен через сеть меню.

### **Hacker**

В Сети, в отличие от больших сообществ, это не плохой человек — это просто тот, кто любит напрягать аппаратное и программное обеспечение до предела, чтобы посмотреть, что может делать компьютер. Тех, кого многие называют хакерами, net.denizens определяет как взломщиков (crackers).

### **Handshake (Рукопожатие)**

Два модема при установке соединения сначала договариваются, как передавать данные.

### **Hang (Зависание)**

Модем не может повесить трубку.

### **HEPnet**

High Energy Physics net — сеть Физики Высоких Энергий. Широкомасштабная (глобальная) специализированная сеть. Предназначена для использования в научных исследованиях по физике элементарных частиц. Основана в CERN-е (Женева, Швейцария).

### **Holy war («Священная война»)**

Споры, которые затрагивают некоторые основные постулаты веры, несогласие с которыми и вызывает войну. Например: Машины IBM PC по жизни лучше машин Macintosh.

### **Host system (Местная система)**

Общедоступная подсистема, дает доступ к Сети тем, кто не входит в научное общество или правительственные службы.

### **Host (Хост)**

Сетевая рабочая машина; главная вычислительная машина; ведущая ЭВМ. Сетевой компьютер, который помимо сетевых функций (таких, как обслуживание сети и передача сообщений) выполняет пользовательские задания (программы, расчеты, вычисления и т.д.)

### **IAB**

Internet Architecture Board — архитектурный совет Internet. «Правящий Совет», принимающий решения о стандартах и других важных предметах обсуждения в Internet.

**ICMP**

Internet Control Message Protocol — Internet-протокол управляющих сообщений. Служебный протокол Internet. Посредством него узлы сети договариваются о выборе оптимальных решений. Например, с помощью ICMP передаются указания о модификации таблиц маршрутизации.

**IETF**

Internet Engineering Task Force — оперативный инженерный отряд Internet. Группа добровольцев, изучающих и разрешающих технические проблемы сети и дающих рекомендации IAB.

**IMHO**

In My Humble Opinion (По моему скромному мнению).

**Internet**

Всемирная система, объединяющая меньшие компьютерные сети. Сети, объединенные в Internet, используют специальный набор стандартов коммуникации, известный как TCP/IP.

С маленькой буквы — любое собрание отдельных сетей, работающих совместно.

С большой буквы — всемирная «Сеть сетей», которые соединены друг с другом согласно протоколам IP или другим аналогичным. Internet обеспечивает пересылку файлов, удаленный доступ и сеанс работы, пересылку сообщений электронной почты, службой новостей и много других услуг и служб. Объединяет более 12 тысяч локальных сетей, миллиона компьютеров и более 30 миллионов пользователей.

**IP**

Internet Protocol. Самый важный из протоколов, на которых основана Internet; протокол нижнего — сетевого уровня. Позволяет пакетам пересекать на пути к пункту назначения множество сетей.

**ISO**

The International Standards Organization — международная организация стандартов. Женева, Швейцария. Организация, которая определила отличный от Internet набор протоколов, называемых ISO-OSI-протоколы. Теоретически ISO/OSI протоколы, в конце концов, заменят протоколы Internet. Когда это произойдет



и произойдет ли это на самом деле — темы, ныне очень бурно обсуждаемые.

### **ISOC**

Internet Society — Internet-овское общество. Общественная организация, члены которой содержат международную информационную сеть. Международная профессиональная организация основанная с целью способствовать развитию, стандартизации и распространению методик и технологий, позволяющих связываться разнородным информационным системам. Это Общество издает журнал новостей, организует конференции и ведает распространением по электронной почте информационных бюллетеней в целях ознакомления мирового сообщества с глобальной сетью сетей, называемой Internet. ISOC содержит IAB и его различные IETF, поддерживает связи и взаимодействует с другими международными организациями и создателями стандартов, чтобы способствовать росту и развитию крайне важной инфраструктуры, коей является Internet. Также — правящий орган, пред которым отчитывается IAB.

### **К, к**

kilo — кило. Приставка, означающая умножение на 1000. В информатике и вычислительной технике — К (большое), в физике — к (маленькое).

### **Killfile (файл исключения)**

Файл, который позволяет фильтровать поступления на Usenet, исключая сообщения на определенную тему или от определенных лиц.

### **LAN (Локальная Вычислительная Сеть)**

Коммуникационная система, поддерживающая в пределах одного здания или некоторой ограниченной территории один или несколько высокоскоростных каналов передачи цифровой информации, предоставляемых подключаемым устройствам для кратковременного монопольного использования.

### **LBL**

Под этим сокращением известны две исследовательские организации. Это центры ядерных исследований, оба находятся в Беркли, США.

- Lawrence Berkeley Labs. Это научно-исследовательская, гражданская организация.
- Livermore Berkeley Labs. Военное секретное учреждение.

### **leased line**

То же, что dedicated line. Выделенная линия.

### **Log off**

Отсоединение от системы.

### **Log on/log in (Регистрация)**

Вход в местную систему или общедоступную подсистему.

### **Lurk (Затаиться)**

Чтение сообщений в группе новостей Usenet без каких бы то ни было сообщений в нее.

### **М**

Mega — мега. Приставка, означающая умножение на 1 000 000.

### **Mail reflector (Рефлектор почты)**

Специальный почтовый адрес. Электронная почта, посланная туда, автоматически рассылается оттуда по соответствующему списку, хранящемуся у

р.п. Обычно используется для организации почтовых дискуссионных клубов.

### **Mailing list (Список рассылки)**

В сущности, представляет собой конференцию, в которой сообщения рассылаются прямо в ваш почтовый ящик, а не в группу новостей Usenet. Включиться в список можно посылкой сообщения на некоторый электронный адрес, который часто представляет собой компьютер, автоматизирующий процесс.

### **MILNET**

Military net — военная сеть. Одна из сетей DDN, составляющая Internet; предназначена для несекретных сообщений военных США. Была построена с использованием той же информационной и структурной технологии, что и ARPAnet, но осталась работать, когда ARPAnet была демонтирована.

### **Modem (Модем)**

Происходит от модуляция-демодуляция. Устройство, преобразующее цифровые сигналы в аналоговую форму и обратно для передачи их по линиям связи аналогового типа, например, по телефону.

Служит для подключения компьютера к последовательной линии (обычно к телефонной или аналогичной).

### **MOTSS**

Members of the Same Sex — Участники того же пола. Электронная связь геев и лесбиянок. Изначально сокращение использовалось в федеральном цензе 1980 года.

### **Net provider, service provider (Поставщик сети, услуги)**

Организация, предоставляющая включение в Internet (ее местную подсеть). Если вы хотите получить Internet-доступ, вам следует договариваться именно с ними.

### **net server**

Сетевой сервер.

### **Net**

Сеть.

### **Net.god (Сетевой бог)**

Некто, кто в сети со дня основания, кто все знает и все перепробовал.

### **Net.personality (Сетевая личность)**

Некто, достаточно самоуверенный/болтливый/досужий, чтобы регулярно посылать сообщения на десятки групп новостей Usenet, и чье присутствие ощущают тысячи людей.

### **Net.police (Сетевая полиция)**

Презрительный термин для тех, кто хочет распространить свои стандарты поведения на всех пользователей сети. Часто используется в жестоких войнах наездов (и там иногда превращается в net.nazis — сетевые нацисты).

### **Netiquette (Сетевой этикет)**

Набор правил на основе здравого смысла, рассчитанных на то, чтобы не осложнять жизнь другим. Этикет.

### **Network (Сеть)**

Система коммуникации, связывающая два или несколько компьютеров. Это может быть просто нить кабеля между двумя компьютерами на расстоянии двух метров, и может быть оптоволоконная телефонная или спутниковая сеть, соединяющая тысячи компьютеров в разных концах света.

### **Newbie (Чайник)**

Новичок в сети. Иногда это прозвище используется презрительно сетевыми ветеранами, которые забыли, что и они когда-то были новичками, не знавшими ответа ни на один вопрос. «Clueless newbie» — всегда презрительное прозвище.

### **Newsgroup (Группа новостей)**

Конференция Usenet.

### **NIC**

Network Information Center — Сетевой Информационный Центр. Настолько близко к понятию центра сети, насколько это вообще возможно в сети типа Internet. Здесь обычно можно найти информацию о конкретной сети. Любая организация, ответственная за поставки информации о какой-либо сети.

### **NFS**

Network File System — сетевая файловая система. Системное программное обеспечение, обеспечивающее доступ к файлам на удаленных компьютерах так же, как если бы эти файлы находились на собственном

жестком диске. Вытесняется более совершенной AFS — Andrew File System.

### **NOC**

Network Operation Center — сетевой эксплуатационный центр. Группа отвечающая за постоянное сопровождение и поддержку сети. Каждый сетевой поставщик обычно имеет свой NOC, так что вам следовало бы выяснить, кого же именно звать на помощь, когда возникнут проблемы.

### **Node (Узел)**

Узел сети — компьютер связывающий две сети, использующие одинаковые протоколы. Узел обеспечивает только связь двух совместимых программ на двух таких сетях. При соединении двух сетей, использующих одинаковые протоколы с помощью мостов, мосты напрямую передают пакеты канального уровня (кадры) соседней сети, узлы и маршрутизаторы которой могут передать информацию по своей сети, истолковав служебную информацию, заключенную в конвертах, в которых передаются данные. Узлы сети (как, впрочем, и маршрутизаторы), являются, в отличие от мостов, устройствами сетевого уровня

работы. Также в отличие от мостов, узлы являются существенно протоколно-зависимыми, т.к. используют служебную информацию, сопровождающую данные.

### **NREN**

National Research and Educational Network — национальная исследовательская и образовательная сеть США. Попытка совместить сети, относящиеся к различным федеральным ведомствам США в одну сеть большой пропускной способности. Несмотря на то, что такая перестройка имеет важное историческое и научно-техническое значение, на пользователях это никак не сказывается.

### **NSA line eater (Строчкоглотатель NSA)**

Наиболее осторожные/параноидальные пользователи Сети верят, что агентство национальной безопасности (National Security Agency) обладает огромной мощности компьютером, который читает все, что идет через Сеть. Они шутливо (?) адресуются к этому строчкоглотателю в своих посланиях. Восходит к первым дням Сети, когда концы сообщений иногда пропадали без видимой причины.

### **NSF**

National Science Foundation — национальный научный фонд. В США это аналог нашего Министерства Науки. Содержит NSFNet — высокоскоростную сеть, которая когда-то была хребтом Internet в США.

### **Offline (Автономный режим)**

Когда вы не соединены с местной системой или Сетью, вы находитесь в автономном режиме.

### **Online (На связи)**

Ваш компьютер соединен с электронной службой, электронной доской объявлений или общедоступной системой.

### **OSI**

Open System Interconnection — связь открытых систем. Иной (не Internet) набор сетевых протоколов. Предложен ISO. Этот стандарт сетевого и межсетевого взаимодействия определяет семь уровней взаимодействия компонентов сети: физический, канальный, сетевой, транспортный, сеансовый, уровень представления данных и прикладной. Для каждого уровня разработан один или несколько протоколов, которые

обеспечивают сетевое взаимодействие широкого класса устройств.

### **OSPF**

Open Shortest Path First — открой кратчайший путь первым. Протокол маршрутизации — с его помощью составляются таблицы маршрутизации. Протокол нового поколения. Более эффективный, нежели ныне широко используемый RIP.

### **Packet (Пакет)**

Блок данных, имеющий строго определенную структуру, включающую заголовок и поле данных. В Internet данные разбиваются на маленькие части, которые заключаются в пакеты; каждый пакет пересекает сеть отдельно от других. Размеры пакетов могут меняться от примерно 40 до 32000 байт, в зависимости от оборудования, каналов связи в сети, но, обычно, пакеты не бывают длиннее 1500 байт.

### **Ping**

Программа, которая может отследить маршрут сообщения от вашей подсистемы к другой.

### **Plan file (файл .plan)**

Файл, в котором собрано все, что вы хотите сообщить о себе другим пользователям сети. Вы помещаете этот файл в свою домашнюю директорию на вашей местной общедоступной системе. После этого каждый, кто указывает на вас программой fingers, увидит этот файл.

### **Port (Порт)**

Число, определяющее конкретное приложение Internet. В пакетах Internet, которые компьютер отправляет другому, содержится информация об используемом протоколе (tcp или udr, например) и о приложении, которому предназначены эти данные. «Номер порта», как раз, и определяет это приложение.

Один из физических каналов ввода-вывода компьютера, обычно, разъем на задней панели.

- Последовательный (sequential) — как правило, подразумевает логическую упорядоченность и относится к процессам.
- serial — как правило, подразумевает временную или

физическую упорядоченность и относится к устройствам.

### **Post (Послать сообщение)**

Составить сообщение и послать его в группу новостей на прочтение остальным.

### **Postmaster**

Человек, к которому следует обратиться в конкретной подсистеме, чтобы попросить информацию о ней или пожаловаться на поведение пользователя.

### **PPP**

Point to Point Protocol — протокол «точка-точка». Протокол предназначенный для работы на двухточечной линии — линии, соединяющей два устройства). Протокол канального уровня. Дает компьютеру возможность использовать протокол TCP/IP (Internet) (и, таким образом, становится полноправным членом Internet) на связи по обычной телефонной линии, используя высокоскоростной модем.

PPP — новый стандарт для такого рода связи, он постепенно замещает своего прародителя — SLIP. Популярность PPP очень быстро растет.

### **Prompt (Приглашение)**

Местная система просит вас что-то сделать и ждет ответа. Например, если вы видите «login:», введите свое имя пользователя.

### **Protocol (Протокол)**

Метод, используемый для передачи файла между местной системой и компьютером. Их несколько типов, например, Kermit, YMODEM и ZMODEM.

Совокупность определений (соглашений, правил), регламентирующих формат и процедуры обмена информацией между двумя или несколькими независимыми устройствами или процессами. то есть описание того, как программы, компьютеры или иные устройства должны действовать, когда они взаимодействуют друг с другом. Протокольные определения охватывают диапазон от того, в каком порядке биты следуют по проводу, до формата сообщения электронной почты. Стандартные протоколы позволяют компьютерам различных производителей связываться друг с другом. Компьютеры могут использовать совершенно различное программное обеспечение, лишь бы они

поддерживали соглашение о том, как посылать и понимать принимаемые данные.

**Provider**

Поставщик; снабженец; продавец.

**Public (Общий)**

Экспортируемый. Например, общий файл — это файл, доступный любому пользователю. Существует несколько видов общих ресурсов с различными степенями свободы обращения с ними, во избежание проблем с законностью, их следует четко различать и знать границы дозволенного для каждого класса общности. Самый вольный класс — это «public domain». Такой файл, например, любой может читать, редактировать, использовать по своему усмотрению, в том числе продавать.

**Publicly available**

Публично доступный. Открытый для всеобщего доступа и использования.

**README files (файлы README)**

Файлы, имеющиеся на подсистемах FTP, объясняющие, что (файлы README) представляет собой данная директория

FTP или дающие другую полезную информацию (например, как пользоваться FTP).

**Real Soon Now (Как только, так сразу)**

Неопределенный термин, описывающий, когда нечто на самом деле произойдет.

**Resource (Ресурс)**

Логическая или физическая часть вычислительной системы, которая может быть выделена пользователю и/или процессу: время центрального процессора, область оперативной или внешней памяти, логическое или физическое внешнее устройство.

**RFC**

Request For Comment — запрос для комментария. Портфель документов, в которых опубликованы стандарты Internet, предлагаемые ей стандарты, а также общепринятые идеи, негласные стандарты и т.п. Эти документы фактически определяют Internet. Они описывают не только работу самой сети, но ее использование и даже планы на будущее. Их число приближается к полутора тысячам. Большинство же этих документов



имеет сугубо технический характер. Распространяются DDN NIC. На многих сетевых рабочих машинах хранятся наборы.

RFC-документации различной степени полноты и древности.

### **RIP**

Routing Internet Protocol — Internet-протокол маршрутизации. Протокол, с помощью которого происходит выбор оптимальных маршрутов, составление оптимальных таблиц маршрутизации.

### **ROT13**

Простой способ кодировать грубые шутки, обзор фильмов с продолжением, порнографию и пр. Суть в том, что каждая буква заменяется той, что отстоит от нее на 13 позиций в алфавите. Существуют декодеры, работающие в момент сеанса связи, в pp и gp они встроены.

### **ROTFL**

Rolling on the Floor Laughing — Катаюсь по полу от смеха. Ответ на особо смешной комментарий.

### **Rout (Маршрут)**

Последовательность узлов сети передачи данных, по которой данные пересылаются от источника к приемнику.

### **Router (Маршрутизатор)**

Узел, обеспечивающий маршрутизацию. Для этого он имеет таблицу маршрутизации. Данные устройства работают на сетевом уровне модели OSI. Маршрутизаторы, как и узлы, являются существенно протокольно-зависимыми, т.к. используют служебную информацию, сопровождающую передаваемые данные. Мосты напрямую передают пакеты канального уровня (кадры) соседней сети, узлы и маршрутизаторы которой могут передать информацию по своей сети, истолковав служебную информацию, заключенную в конвертах, в которых передаются данные. Маршрутизаторы должны обладать способностью обмениваться между собой информацией о текущем состоянии сети, чтобы осуществлять оперативную и максимально эффективную маршрутизацию сообщений.

### **Routing table (Таблица маршрутизации)**

Таблица, связанная с узлом сети с коммутацией пакетов или сообщений и указывающая для каждого адресата оптимальный выходной канал; может быть указано несколько каналов в порядке их предпочтительности.

### **Routing (Маршрутизация)**

Выбор маршрута.

### **RTFM**

Read the ....ing Manual. — Читайте свое .... руководство. Часто используется в наездах на тех, кто задает вопросы, ответы на которые можно за несколько минут найти в руководстве. Вежливее — RTM.

### **Screen capture (Перехват экрана)**

Часть коммуникационного пакета, открывающая файл в вашем компьютере и записывающая туда все, что проходит по экрану в сеансе связи с местной системой.

### **Server (Сервер)**

Компьютер, который может автоматически рассылать информацию и файлы в ответ на специально закодированное сообщение электронной почты.

Программа для сетевого компьютера. Дает возможность предоставить услуги одного компьютера другому компьютеру. Обслуживаемые компьютеры сообщаются с этой обслуживающей (сервер-) программой посредством соответствующей пользовательской (client-) программы, работающей на машине конечного Пользователя. Программа-клиент предназначена для работы в паре с программой-сервером.

### **Service**

Сервис; служба; услуги.

### **Session (Сеанс)**

Цикл работы клиента от момента входа в систему или запуска программы до выхода из нее. Или же цикл работы пользователя с диалоговой системой от входа в систему (вызова системы) до выхода из нее.

### **Shareware**

Программное обеспечение, доступное в Сети бесплатно. Если вам нравится это программное обеспечение и вы пользуетесь им, пошлите запрошенный гонорар автору, чье имя и адрес

содержатся в рассылаемом файле с программным обеспечением.

### **Sharing (Разделение)**

Совместное использование чего-либо, каждому достается соответствующая доля разделяемого. Разделение ресурса — совместное использование ресурса (дисковой памяти, процессорного времени, линий связи), когда либо право пользования предоставляется поочередно всем пользователям, либо каждому пользователю выделяется соответствующая часть этого ресурса.

### **Shell (Командный процессор)**

Командный язык и процессор командного языка ОС UNIX и ее разновидностей. Shell — один из наиболее развитых командных языков, являющийся полным языком программирования. Процессор командного языка — системная программа, принимающая и исполняющая команды и командные строки с терминала и из командного файла.

### **Sig file (Файл .sig)**

Иногда — файл подписи (.signature). Файл, который, будучи помещен в вашу домашнюю директорию на общедоступной

подсистеме, будет автоматически добавляться ко всем написанным вами сообщениям на Usenet.

### **Sig quote (Цитата .sig)**

Глубокомысленное/остроумное/еще какое-нибудь изречение, которое включается в ваш файл .sig.

### **Signal-to-noise ratio (Отношение сигнал-шум)**

Количество полезной информации, находимой в конкретной группе новостей Usenet. Часто с пренебрежительным оттенком, например: «В этой группе новостей очень низкое отношение сигнал — шум».

### **SIMTEL20**

White Sands Missile Range поддерживало гигантскую коллекцию бесплатного и дешевого программного обеспечения, которое «отражалось» на многие другие подсистемы ftp в Сети. Осенью 1993 года Военно-Воздушные Силы решили, что у них есть еще чем заниматься, кроме ведения библиотеки бесплатного программного обеспечения, и закрыли ее.

**SLIP**

Serial Line IP — IP для последовательных линий. Протокол, позволяющий компьютеру использовать протоколы Internet (и становится полноправным членом сети) на стандартных телефонных линиях и высокоскоростных модемах. Это протокол последовательной посимвольной передачи данных — протокол канального уровня. SLIP замещается PPP, но пока еще широко распространен.

**Smiley (Смайлик)**

Способ электронного выражения эмоций. Наклоните голову влево и посмотрите на эту картинку :-). Смайлики бывают разные — от хохочущего до унылого.

**SNA**

Systems Network Architecture — сетевая архитектура систем, протокол SNA. Разработанная фирмой IBM организация сети ЭВМ. В SNA определены три уровня взаимодействия компонентов сети — уровень управления передачей, уровень функционального управления и прикладной уровень.

**Snail mail (Письмо- улитка)**

Письмо, которое приходит на бумаге в почтовый ящик на вашей входной двери.

**Supercomputer (Суперкомпьютер)**

Супер-ЭВМ. Вычислительная система, относящаяся к классу самых мощных систем в данное время.

**Sysadmin**

Системный администратор; человек, который управляет местной или общедоступной подсистемой.

**Sysop**

Системный оператор. Тот, кто ведет электронную доску объявлений (BBS).

**TANSTAAFL (Танстаафл)**

There Ain't No Such Thing As A Free Lunch. Дармовой закуски не бывает.

**TCP**

Transmission Control Protocol — управляющий протокол передачи. Один из базовых протоколов Internet, а именно, транспортного уровня. Надежный (исправляющий ошибки) протокол связи с установлением логического соединения, то есть, для организации взаимодействия, при котором: между взаимодействующими

станциями сети устанавливается логическое соединение — виртуальный канал; по этому каналу передаются и принимаются кадры (пакеты) с регистрацией их последовательности; управляется поток кадров и организовывается повторная передача искаженных кадров; в конце сеанса канал разрывается. TSP базируется на IP, то есть, виртуальный канал эмулируется средствами дейтаграммного протокола.

### **Telnet**

Программа, которая позволяет связаться с другими компьютерами в Internet.

Протокол эмуляции терминала. Позволяет входить в другие компьютерные системы в Internet.

Прикладная программа, дающая возможность входить в удаленные системы вычислительных машин, используя протокол telnet.

### **Terminal emulation (Эмуляция терминала)**

Есть несколько способов определить, каким образом ваша клавиатура и экран должны взаимодействовать с операционной системой общедоступной

подсистемы. Большинство коммуникационных программ предлагают выбор «эмуляции», что позволяет копировать поведение клавиатуры, которая обычно присоединяется к компьютеру.

### **Token Ring**

Эстафетное кольцо. Технология организации Локальных Вычислительных Сетей, которые могут подсоединяться к Internet.

### **Traffic (Трафик)**

Поток сообщений в сети передачи данных (или автомобилей на дороге). Загрузка сети (дороги).

### **UDP**

User Datagram Protocol — протокол пользовательских дейтаграмм. Еще один протокол — транспортного уровня, на котором основана Internet. UDP — это ненадежный протокол связи без установления логического соединения, то есть, для организации взаимодействия, при котором транспортный уровень обеспечивает только посылку кадра одной или нескольким станциям сети. Проверка правильности и обеспечение целостности

передачи данных осуществляется на более высоком уровне.

### **UNIX**

Популярная операционная система, сыгравшая важную роль в развитии и популяризации Internet. И все же, вопреки молве, совсем не обязательно работать под UNIX для того, чтобы иметь возможность использования Internet. Существуют различные вариации UNIX. Самые распространенные — это BSD и System V.

### **Upload (Загрузка «Вверх»)**

Копирование файла из вашего компьютера в местную систему

### **USENET**

Неформальная, скорее даже анархическая, группа систем (сетей, хостов, компьютеров), которые обмениваются «новостями». Новости по сути своей аналогичны доскам объявлений других сетей.

### **User name (Имя пользователя)**

В большинстве местных систем при первом соединении вас попросят назвать имя пользователя, состоящее из одного

слова. Это может быть любая комбинация букв и цифр.

### **User node**

Узел пользователя.

Абонент сети.

Пользовательская станция. Рабочая станция (локальной сети).

### **UUCP**

Unix-to-Unix CoPy (Копирование из Unix в Unix). Метод передачи поступлений на Usenet и электронной почты, который требует гораздо меньше сетевых ресурсов, чем TCP/IP, но работает гораздо дольше. Инструмент для копирования файлов между двумя системами UNIX, на котором были построены службы новостей USENET и электронная почта. Также — соответствующий протокол. Хотя UUCP до сих пор применима, Internet имеет и лучше инструментарий для той же работы. Протокол канального уровня.

### **Virtual circuit (Виртуальный канал)**

В сети коммутации пакетов — средства, обеспечивающие передачу пакетов между двумя узлами с сохранением исходной

последовательности, даже если пакеты пересылаются по различным физическим маршрутам. Виртуальное соединение устанавливается при вызове и аннулируется после сеанса связи.

### **Virtual connection (Виртуальная связь)**

Связь двух модулей одинакового уровня на двух связывающихся компьютерах, в действительности осуществляемое сложной многоступенчатой передачей информации через множество подлежащих уровней; во время такой передачи происходит многократное преобразование передаваемой информации, — на каждом из проходимых уровней.

### **VT100**

Еще одна система эмуляции терминала. Она поддерживается множеством коммуникационных программ и общепринята в Сети. Более новая версия — VT102.

### **WAIS**

Wide Area Information Servers — широкомасштабные информационные серверы. Очень мощная система для просмотра в поисках нужной информации

в базах данных (или библиотеках) по всей Internet. Использует метод ключевых слов.

### **White Pages**

«Белые страницы». Сетевой аналог бесплатно распространяемых на Западе телефонных справочников. Списки пользователей Internet, доступных по сети. Имеется несколько различных видов служб, серверов, предоставляющих такого рода услуги.

### **World Wide Web**

Всемирная паутина. Гипертекстовая система для поиска и использования ресурсов Internet. Имеет смысл только на линиях с высокой пропускной способностью. Вдали от соответствующих серверов влетает в копеечку и вызывает раздражение при использовании.

## Использованные материалы

### **Сто процентов русского языка**

М. Суханова.

### **Поиск в стиле ДИСКo**

М. Суханова. (Из журнала «Мир ПК», N 3, 1999).

### **Поиск людей и организаций в Интернет: лидирующие поисковые системы и принципы их тестирования**

М. Талантов, Центр Информационных Технологий.

### **Как быстро найти нужную информацию на избранном Web-узле**

Михаил Талантов.

### **Поиск в Интернете — внутри и снаружи: Эффективная методика поиска информации в сети Интернет**

Артем Попов.

### **Поиск людей и организаций в интернете — попытка систематизации**

Левчик Василий Анатольевич. Для свободного распространения и цитирования при условии сохранения данного (levchik@icom.ru).

### **Internet: Наиболее интересные вопросы и ответы о программном обеспечении из рубрики «Есть вопрос — Есть ответ»**

iNTERface, Israel.

### **Как уцелеть среди акул?**

Х. Маккей

### **Как в Сети найти работу**

А. Запольскис

### **Юноше, начинающему делать карьеру...**

С. Юрлов

### **Поиск легальной работы в качестве программиста в США**

Д. Федорец

### **Работа в Internet**

Webmaster Roman (web\_design@mail.ru)

### **Работа по-русски**

<http://www.agava.ru/>

### **Как искать работу**

Консультационный кадровый центр «Тамань»

### **Как составить резюме**

БЭКАП (<http://hunter.bacup.ru/>)

### **В какое кадровое бюро обратиться?**

В.Л. Павлов, президент Днепропетровского клуба соционики и социальных технологий;  
Л.Б. Куницына, администратор кадрового бюро «Специалист» ЦБС «Баланс»



*Научно-популярное издание*

**Серия книг «Хитрости и тонкости»**

Кузнецов Сергей Дмитриевич

**Внутри Internet:  
Методы поиска информации**

Налоговая льгота  
«Общероссийский классификатор  
ОК 005-93-ТОМ2 953000 — Книги и брошюры»

Лицензия ЛР № 068246 от 12.03.1999 г. Подписано в печать  
10.07.2006. Формат 70x100/32. Бумага офсетная.  
Кол-во п.л. 7. Тираж 3000 экз.