

РАЗДЕЛ V

ИНТЕРНЕТ: ИНФОРМАЦИОННЫЕ РЕСУРСЫ И СЕРВИСЫ

Данный раздел пособия, не затрагивая теоретических аспектов работы сети Интернет или описания исторических интернет-ресурсов, ставит своей целью изложение основ работы в Интернете, а также дает основные рекомендации по поиску тематических информационных ресурсов в Интернете. В разделе также рассматриваются вопросы, связанные с поиском информации и оценкой веб-ресурсов в рамках требований, предъявляемых на коллоквиуме по соответствующей теме.

ГЛАВА 11. ОСНОВЫ РАБОТЫ В СЕТИ ИНТЕРНЕТ

11.1. ВЕБ-СТРАНИЦЫ И БРАУЗЕРЫ

Интернет – это глобальная компьютерная сеть. Технически это означает, что Интернет объединяет великое множество компьютеров по всему миру и позволяет передавать с одного компьютера на другой разного рода информацию. В содержательном же плане Интернет представляет собой огромный комплекс информационных ресурсов, своеобразную глобальную библиотеку файлов, расположенных на миллионах компьютеров и в принципе доступных любому пользователю, подключенному к этой сети¹.

Веб-страницы, на которые заходят пользователи Интернета, могут быть созданы с помощью различных программных средств, они могут представлять собой отображение заранее подготовленных файлов или формироваться динамически по запросу пользователя, однако все они реализуют представление *гипермедиа*. Это означает, что на экране пользователь видит текст, в который могут быть добавлены, во-первых, ссылки на другие страницы, фрагменты страниц или файлы и, во-вторых, рисунки и другие элементы мультимедиа. Это могут быть программы (скрипты и апплеты), выполняющие какие-либо действия, а также видео, звуковые файлы и т. п.

Для того, чтобы просматривать веб-страницы, используются специальные программы – *браузеры*. Браузер должен выполнять две функции –

¹ Значительная часть интернет-ресурсов являются закрытыми, то есть доступными только после авторизации (для чего нужно знать пароль) или доступными только для пользователей, заходящих в Интернет с определенных компьютеров.

обеспечивать соединение с удаленным компьютером и загрузку веб-страниц, расположенных по указанному адресу, а также корректно показывать веб-страницы на экране.

Разных браузеров немало, и их популярность со временем изменяется. На начало 2010-х гг. наиболее распространенными браузерами в мировом Интернете являются Google Chrome, Microsoft Internet Explorer, Mozilla Firefox, Safari и некоторые браузеры для мобильных устройств. Эта статистика имеет свою специфику в разных странах: в частности, в России довольно распространенными являются также браузеры Opera и Яндекс.Браузер, а вот Safari используется относительно редко. Принципиально все эти программы мало чем отличаются друг от друга, потому что они должны одинаково (в таком виде, как этого хотел разработчик) показывать одни и те же веб-страницы. Впрочем, абсолютное единство тут все равно недостижимо, прежде всего потому, что пользователи работают с разными экранами, и на разных компьютерах по-разному может смотреться одна и та же страница. Поэтому у браузеров есть довольно широкие возможности, позволяющие пользователям самим определять настройки, с учетом которых просматриваются веб-страницы.

На рис. 5.1 показано, как примерно выглядит панель управления в наиболее популярных веб-браузерах (хотя здесь многое зависит от версии программы и от конкретного пользователя: браузеры обычно предоставляют более или менее широкие возможности по настройке панели инструментов).

Раздел V

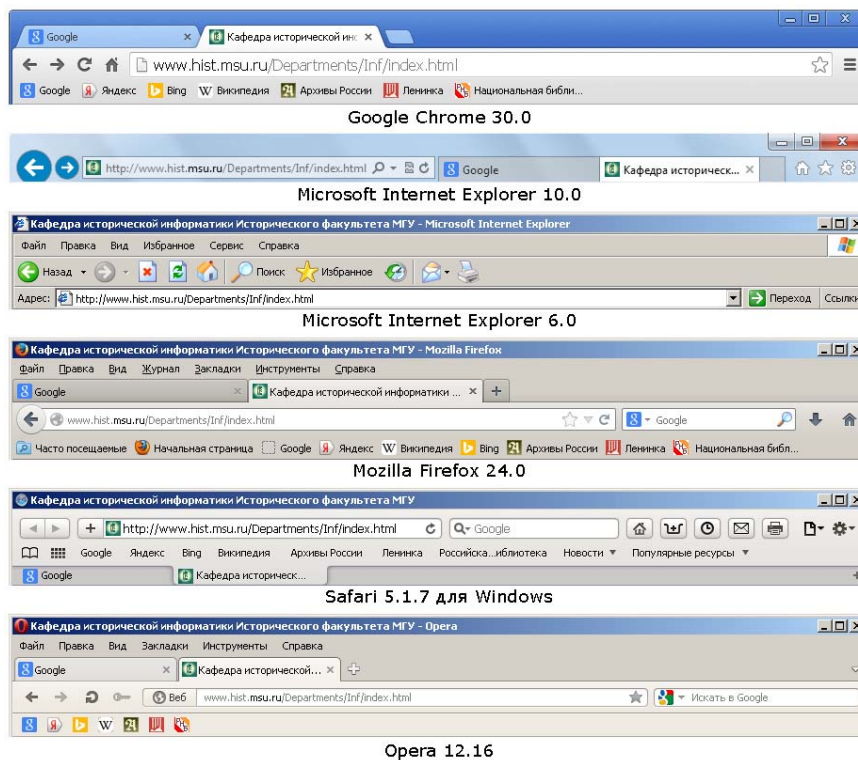


Рис. 5.1. Панель управления в разных браузерах.

Адрес страницы

Почти в каждом из представленных на рис. 5.1 браузеров (кроме старой версии Microsoft Internet Explorer) поддерживаются *вкладки*. Это значит, что в браузере может быть одновременно открыто несколько страниц. На рис. 5.1 в браузерах открыто по две вкладки: неактивной является поисковая система Google, активной (то есть той, которую пользователь читает в данный момент) – страница кафедры исторической информатики исторического факультета МГУ.

Основной элемент панели управления браузера – *адресная строка*. Здесь отображается адрес открытого или открывающегося в активной вкладке файла. Простой адрес обычно состоит из трех частей – используемого протокола, адреса веб-сайта и названия конкретного файла. Рассмот-

рим это на приведенном на рис.5.1 примере (http://www.hist.msu.ru/Departments/Inf/index.html).

Часть адреса, обозначающая **протокол**, представлена здесь буквами *http://*. Протокол – это совокупность команд, регулирующая метод передачи данных. Протокол *http* используется при связи с подавляющим большинством интернет-страниц, хотя в Интернете могут встретиться и другие варианты, с которыми работают браузеры, например *https://*, *ftp://* и др. Хотя протокол отображается в адресной строке и часто указывается в списках ссылок (как в Интернете, так и в печатных изданиях), писать его название в адресной строке необязательно – компьютер успешно определяет его сам.

Вторая часть – собственно адрес сайта, состоящий из идентификаторов, разделенных точками. В адресе *www.hist.msu.ru* первая часть – *www* – является указателем того, что это веб-сайт, последняя – *ru* – представляет собой имя одного из **доменов** высшего уровня и в данном случае указывает, что сайт зарегистрирован в зоне Интернета, отведенной для России, два оставшихся идентификатора соответственно принадлежат МГУ и историческому факультету. Этот адрес определяет собственно сайт исторического факультета МГУ.

Последняя часть адреса, *"/Departments/Inf/index.html"* – это указатель, позволяющий браузеру найти конкретный файл на сайте. В данном случае он показывает, что нужен файл *index.html*, находящийся в папке *Inf*, которая, в свою очередь, находится в папке *Departments* сайта. Расширения *HTML* и *HTM* являются основными для интернет-страниц, однако могут встретиться и другие варианты: *PHP*, *ASP*, *SHTML*, *XML* и прочие.

Переход к веб-странице

Существуют три возможности перехода к нужной веб-странице. Во-первых, если адрес искомой страницы известен, можно написать его непосредственно в адресной строке (после чего нажать **Enter**). При этом надо знать следующее:

1. В адресе имеет значение каждый символ, поэтому, в частности, нужно следить, чтобы там не появлялось лишних пробелов.
2. В адресах могут использоваться как прописные, так и строчные буквы. Воспринимаются ли прописная и строчная буквы компьютером как один и тот же или как два разных символа, – это зависит от настроек сервера, на котором установлен сайт, и вопрос может решаться по-разному.
3. Хотя адреса огромного числа веб-сайтов традиционно начинаются с указателя *www*, это не является обязательным. Есть сайты, на которых предусмотрено, что нет разницы в наборе адреса с указателем или без него;

так, адреса *www.yandex.ru* и просто *yandex.ru* равноправны. В других случаях правильное написание адреса важно. Например, для сайта Российской национальной библиотеки действует только адрес *www.nlr.ru*, а не *nlr.ru*, (правда, работает автоматический переход со второго на первый) а вот для сайта Ватиканских музеев, наоборот, работает только адрес *mv.vatican.va*, но не *www.mv.vatican.va*².

4. Для выхода на главную страницу сайта, как правило, не нужно указывать название файла, потому что для главного файла оно определяется автоматически (этот файл часто называется *index.html* или *index.htm*). Поэтому достаточно написать *www.hist.msu.ru*, чтобы получить адрес *http://www.hist.msu.ru/index.html*. Это может относиться и к отдельным разделам сайта. Так, если каким-то образом (скажем, по ссылке) получен сложный адрес вроде *http://www.lib.ru/INOOLD/ENGLAND/br_history.txt* (это текст "Истории бриттов" Гальфрида Монмутского в электронной библиотеке Максима Мошкова), а на самой странице нет выхода на верхний уровень сайта, то можно попробовать просто удалить часть адреса, записав *http://www.lib.ru/INOOLD/ENGLAND/* или *http://www.lib.ru/INOOLD/*, и вполне возможно, что при этом откроются главные файлы соответствующих папок, т. е. в приведенном примере мы получим доступ, соответственно, к разделам "Исторические тексты, имеющие отношение к Англии" и "Старинная европейская литература" той же электронной библиотеки.

Вторая возможность перейти на интересующую нас страницу – это воспользоваться ссылкой с какой-то другой страницы. Ссылки в том или ином виде содержатся почти на любой веб-странице, и это могут быть как внутренние ссылки, позволяющие переходить на другие страницы того же сайта, так и внешние ссылки, перебрасывающие пользователя к ресурсам, расположенным, может быть, на другом континенте. Ссылки стоят и на простых гипертекстовых страницах, и на страницах, представляющих собой специальные списки ссылок, и на динамических страницах со списками ссылок, создаваемых поисковыми системами по конкретным запросам пользователей, но технически все ссылки реализованы одинаково – имеется указатель, по которому нужно щелкнуть левой кнопкой мыши. Указателем может быть текст (обычно выделенный другим шрифтом, часто подчеркнутый), рисунок (в том числе сделанный в виде

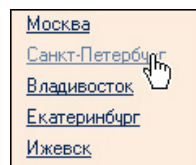


Рис. 5.2. Вид курсора мыши при наведении на гиперссылку

² Зачастую это связано с тем, что существует специальный сайт более высокого уровня, как и в данном случае – сайта *www.mv.vatican.va* нет, но есть общий главный сайт Ватикана – *www.vatican.va*.

кнопки с каким-нибудь указывающим текстом), даже часть рисунка. При наведении мыши на указатель ссылки курсор мыши изменяется и приобретает вид, представленный на рис. 5.2.

Третий способ перехода к веб-странице – использование какого-либо списка адресов, сохранившихся в памяти данного компьютера. Прежде всего, это относится к списку **закладок**. В любой момент браузеру можно дать задание запомнить адрес текущей страницы в специальном списке закладок. На рис. 5.1 для добавления адреса открытой вкладки в этот список служит знак пятиконечной звездочки (только в браузере Safari – знак «плюс» непосредственно перед адресом вкладки). Потом, в том же или следующем сеансе работы в Интернете, для перехода к этой отмеченной странице достаточно выбрать ее в списке закладок. В списке закладок отображаются не адреса, а заголовки, которые даны большинству страниц их разработчиками (когда страница открыта, они написаны в обозначении вкладки и иногда также в названии окна браузера – см. рис. 5.1, – однако в момент добавления закладки можно ввести и свое название ресурса). Кроме того, обычно в браузерах предусмотрена возможность отображения специальной **панели закладок** – специальной области, на которой постоянно располагаются некоторые наиболее полезные закладки (на рис. 5.1, где возможно, в панели закладок размещены ссылки на поисковики Google, Яндекс, Bing, сетевую энциклопедию Википедию, а также на сайты Архивной службы России, Российской Государственной библиотеки («Ленинки») и Национальной библиотеки России).

Закладки – это специальный инструмент для того, чтобы запоминать адреса посещенных данным браузером страниц; однако помимо этого компьютер хранит в памяти список адресов страниц, просмотренных последними. Насколько велик этот список – зависит от настроек браузера. Выбор нужной страницы из списка производится так же, как и в случае с закладками, только используется другой раздел. В браузерах Microsoft Internet Explorer и Mozilla Firefox этот раздел называется «Журнал», в других – «История» (см. рис. 5.3). Специальные кнопки существуют и для перемещения назад, на предыдущую просмотренную страницу, и обратно (стрелки «влево» и «вправо» на рис. 5.1).

Конечно, предусмотрена возможность открыть страницу, на которую ведет ссылка, в новой вкладке или даже в новом окне (в последнем случае открывается второе окно браузера с одной вкладкой). Для этого нужно щелкнуть по ссылке правой кнопкой мыши и в открывшемся меню выбрать пункт «Открыть ссылку в новой вкладке (в новом окне)» (см. рис. 5.4). Бывает, что страницы организованы именно таким образом, что ссылки с них открываются в новом окне, оставляя открытым и исходное окно.

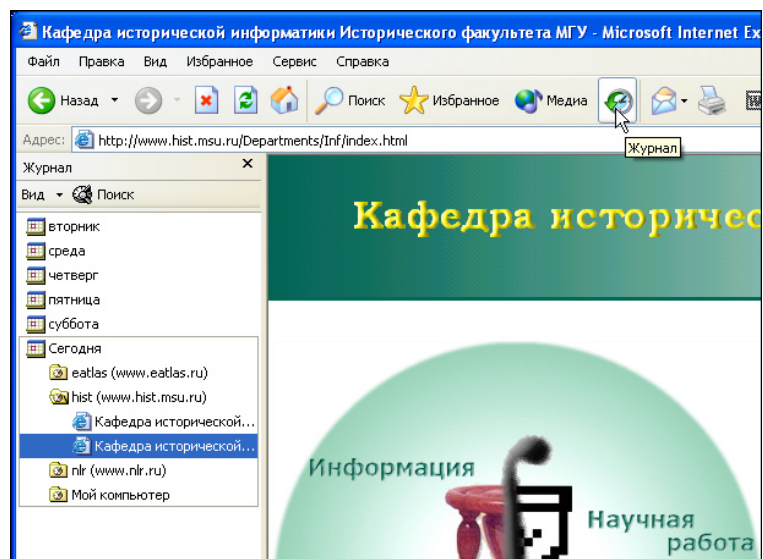


Рис. 5.3. Список истории посещений (Журнал в браузере Internet Explorer)

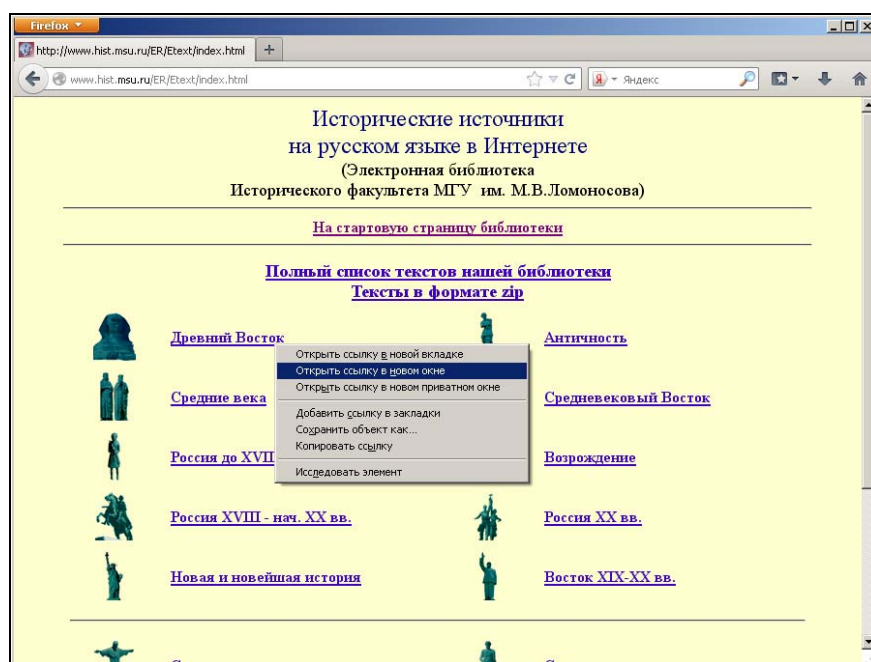


Рис. 5.4. Меню, позволяющее обработать ссылку на веб-странице. Вызывается нажатием по ссылке правой кнопки мыши.

Наконец, отдельно в памяти каждого браузера хранится адрес так называемой "Домашней страницы". На нее всегда можно переключиться посредством нажатия лишь одной кнопки (стилизованное изображение «домика» на рис. 5.1). В некоторых браузерах, например, в Internet Explorer, какая-нибудь страница должна быть открыта обязательно, так что сеанс начинается с загрузки "Домашней страницы".

Как уже говорилось, одна из особенностей файлов веб-страниц заключается в том, что их практически невозможно сделать такими, чтобы все пользователи увидели их абсолютно одинаково. Из-за этого возможно появление проблем при просмотре страниц. Наиболее серьезная из них связана с возможным рассогласованием в кодировках. В частности, только для простого кириллического шрифта (русского языка) существуют по крайней мере две распространенные в Интернете кодировки – Win1251 и KOI8. Если страница написана в одной из этих кодировок, а браузер по каким-либо

Раздел V

причинам считает, что нужно использовать другую кодировку, то страница будет отображена некорректно (см. рис. 5.5)³.

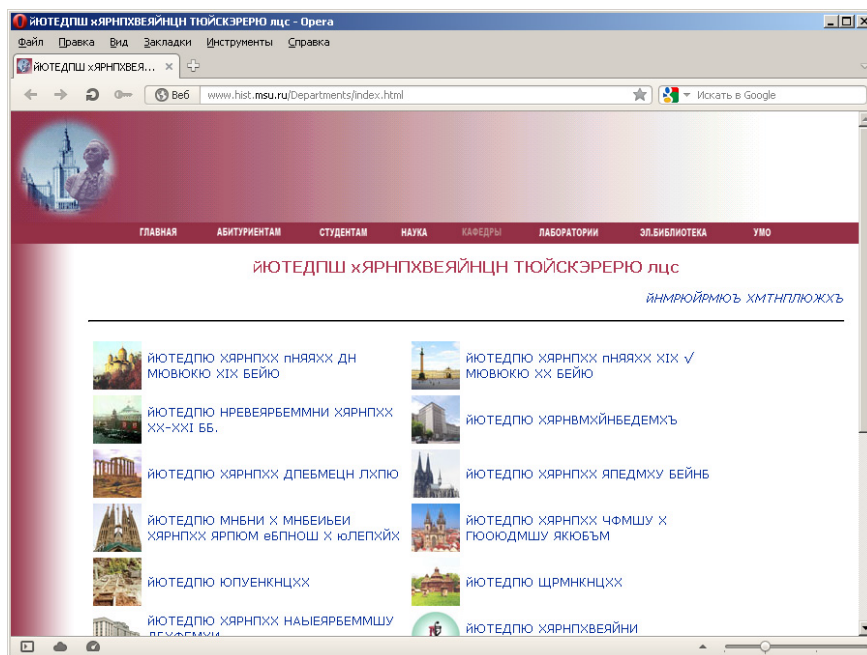


Рис. 5.5. Страница, показанная в неверной кодировке (здесь – KOI8 вместо Win1251)

³ В приведенном на рис. 5.5 примере указатели в полосе над основным текстом читаются правильно, – это происходит потому, что технически они представляют собой рисунки, и их отображение никак не зависит от выбранной кодировки текста.

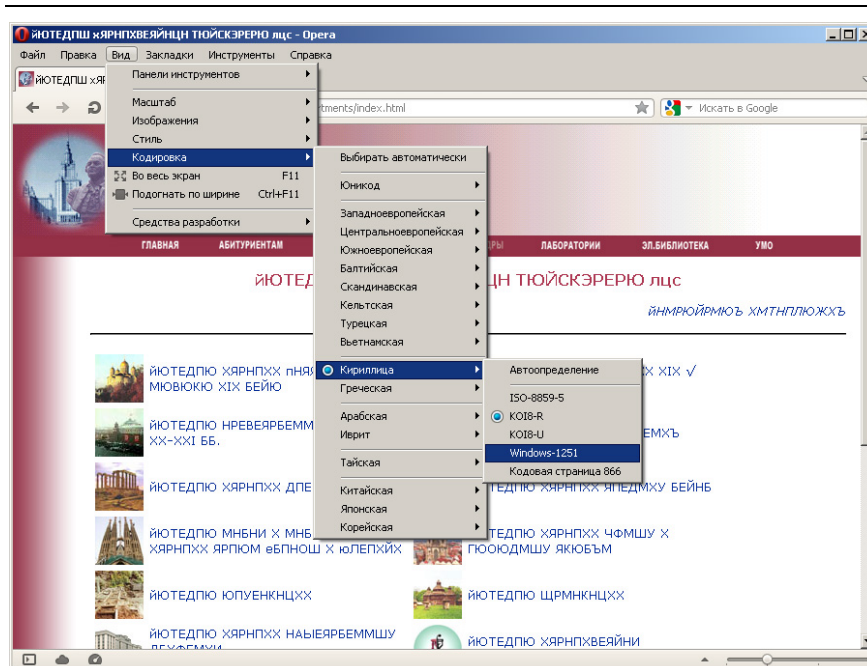


Рис. 5.6. Страница в неверной кодировке и выбор правильной кодировки

То же случится, если страницу, написанную на русском языке, браузер станет просматривать в западноевропейской кодировке или любой другой. В большинстве случаев пользователь может справиться с этой проблемой. Для этого нужно лишь выбрать правильную кодировку (рис. 5.6). В разных браузерах это делается по-разному, но такая возможность всегда есть. Справедливости ради нужно сказать, что программы все лучше справляются с автоматическим разрешением этой проблемы, и она постепенно уходит в прошлое.

Другая проблема связана с тем, что в зависимости от количества точек экрана шрифты и рисунки будут отображаться крупнее или мельче. Если шрифт страницы кажется слишком мелким или крупным, обычно можно изменить его размеры в тех же настройках просмотра страниц (рис. 5.7). В большинстве браузеров сейчас также очень просто изменяется масштаб просматриваемой страницы: для этого почти везде используется комбинация клавиш **Ctrl** и **+** или **Ctrl** и **-**. При изменении масштаба одновременно увеличиваются (уменьшаются) и шрифты, и рисунки. Правда, качество рисун-

ка при его увеличении не может улучшиться, и зачастую увеличенные рисунки выглядят неаккуратно ⁴.

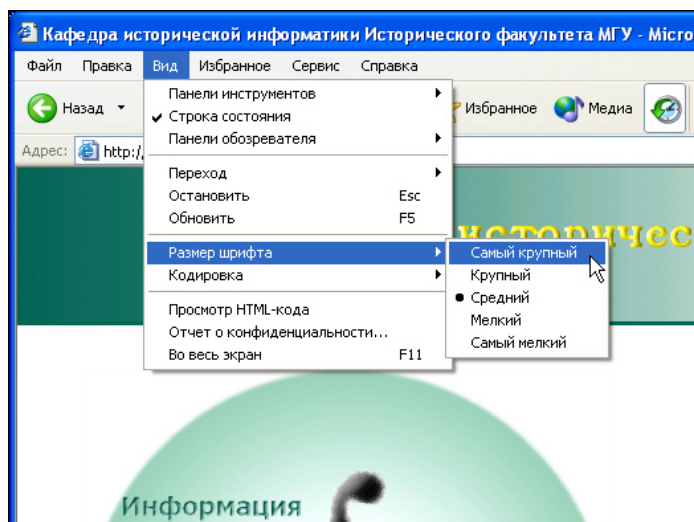


Рис. 5.7. Выбор размера шрифта в браузере Internet Explorer

Наконец, возможны проблемы с адекватным отображением страниц в связи с тем, что в браузере может быть отключен режим приема рисунков. Графические файлы вообще занимают существенно больше места, чем текст, и поскольку для некоторых пользователей Интернета даже сейчас еще принципиально важно уменьшить объем получаемой бесполезной информации, они могут отключать автоматическую загрузку изображений. В этом случае может оказаться, например, что текст будет написан белым цветом на белом фоне, потому что не будет загружен рисунок, служащий темным фоном. Впрочем, разработчик страницы обычно принимает меры к тому, чтобы этого не случилось.

Поскольку разработчик не может обеспечить жесткое форматирование страницы, в Интернете достаточно распространены файлы формата PDF. Это – псевдографический формат, то есть компоновка информации, сохраняющейся в нем, жестко задана – точно как на книжной странице. Хотя в большинстве случаев формат PDF позволяет проводить автоматический поиск заданного фрагмента текста, и теоретически возможен перевод

⁴ Это тот же эффект ухудшения рисунка, который проявляется, например, если смотреть через увеличительное стекло на фотографию, напечатанную в газете.

текста из этого формата в простой текстовый формат, PDF-документ сохраняет все форматирование страниц – положение и размеры абзацев, таблиц, схем, рисунков, начертание шрифтов и т. п. Естественный недостаток этого формата – относительно большой размер файлов. Для того, чтобы просмотреть файлы в формате PDF, нужна специальная программа, Adobe Acrobat Reader, установочные файлы которой можно бесплатно загрузить с веб-сайта компании "Adobe Systems" (ссылка «Acrobat» как на главной странице www.adobe.com, так и на региональной странице по России www.adobe.com/ru/) и во многих других местах ⁵. Если эта программа уже установлена, то браузер, получая команду открыть PDF-документ, копирует файл на жесткий диск, а потом автоматически запускает программу Acrobat Reader, в которой и открывает нужный файл (возможный вариант – PDF-файл открывается непосредственно в окне браузера).

Другой распространенный в Интернете формат для передачи многостраничных текстов – это специально разработанный для передачи сканированных текстов книг и статей формат DjVu ("ДежаВю"). Документ, сохраненный в формате DjVu, выглядит практически так же, как и в формате PDF, и к нему также возможно присоединять распознанный текст, хотя изначально формат DjVu не был ориентирован на это, и PDF предоставляет больше возможностей по работе с текстом и разметкой документа. Однако обработка графики в формате DjVu проводится весьма эффективно, и размеры файлов, составленных из сканированных страниц, получаются относительно небольшими, что и делает его по-своему привлекательным. Для просмотра файлов DjVu разработаны разные программы и *плагины* ⁶ к браузерам, расположенные, в частности, на сайте www.djvu.org.

Помимо указанных, есть и ряд других форматов файлов, которые браузер открывает с помощью соответствующих программ. Так, браузер Internet Explorer правильно воспринимает форматы текстового процессора Word (DOCX и RTF) и табличного процессора Excel (XLSX), и для отображения файлов с такими форматами, если на них указывают ссылки, открывает одну из этих программ, причем (в зависимости от версий) либо пытается совместить функции браузера с просмотром в нужной программе, заменяя командную строку программ Word или Excel командной строкой браузера

⁵ Например, часто с новым оборудованием (принтерами, сканерами, цифровыми камерами) поставляются компакт-диски, на которых, помимо прочего, записаны в формате PDF инструкции по работе с этим оборудованием. Чтобы с чтением инструкций не возникало проблем, на этих дисках записана также и программа чтения.

⁶ Плагины – это программы, которые работают не сами по себе, а в качестве приложений к каким-то другим программам, позволяя тем выполнять дополнительные функции. В данном случае, плагин DjVu делает возможным просмотр файлов формата DjVu непосредственно браузером.

Раздел V

(рис. 5.8), либо сразу открывает полученный файл в соответствующей программе.

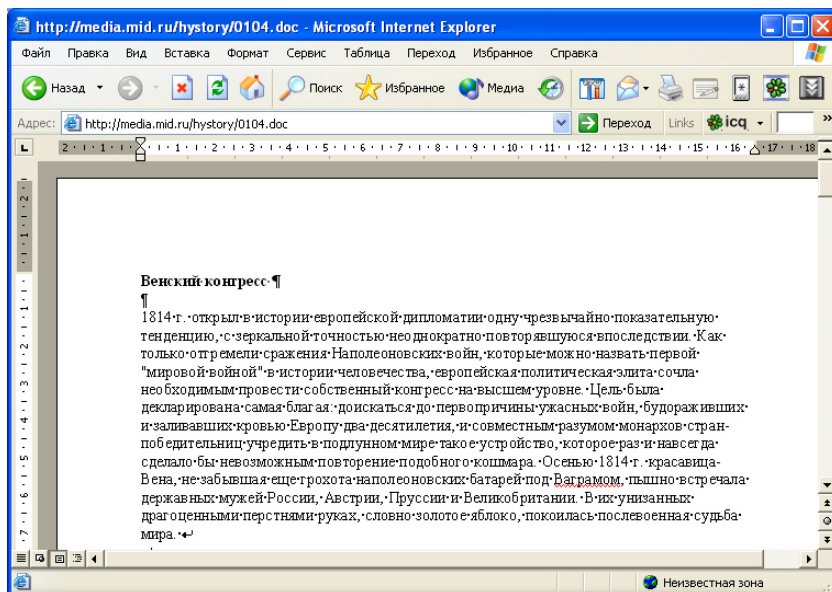


Рис. 5.8. Документ Word, открытый в браузере Internet Explorer

Когда же браузер встречает файлы неизвестных форматов (например, архивы вроде ZIP или RAR), то он предлагает просто сохранить их на диске (рис. 5.9)⁷. Ровно так же реагируют другие, помимо Internet Explorer, браузеры на файлы форматов DOCX, RTF или XLSX.

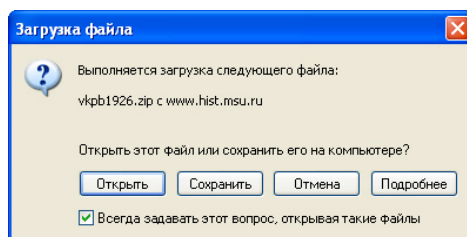


Рис. 5.9. Предложение сохранить файл формата ZIP

⁷ Строго говоря, как раз архивированные файлы "известны" браузеру, однако их все равно нельзя открыть непосредственно в окне браузера, а нужно сначала сохранить на диске и подвергнуть обработке (разархивировать).

В любом случае пользователь может сразу затребовать сохранение на диске файла, на который указывает ссылка. Для этого нужно выбрать соответствующий пункт («Сохранить объект как...») в контекстном меню, возникающем при нажатии правой кнопкой мыши на ссылку (рис. 5.4 на стр. 165).

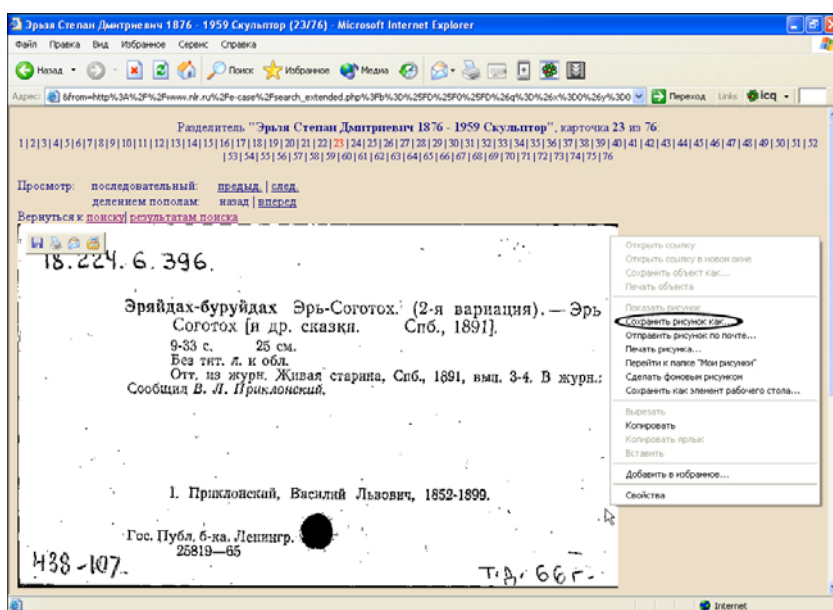


Рис. 5.10. Команда сохранения рисунка в контекстном меню (вызывается правой кнопкой мыши)

Точно так же – с помощью контекстного меню – можно сохранить на диске любой рисунок с уже открытой страницы (рис. 5.10).

Саму же открытую веб-страницу можно сохранить в разных видах, среди которых важнейшими являются "html-отдельная страница" ("только html") и "html-полная страница". Режим полной страницы сохраняет страницу вместе со всеми картинками и возможными вспомогательными файлами⁸, отдельная страница – это только html-код. Если же нужно сохранить лишь фрагмент текста веб-страницы, то иногда проще всего выделить этот фрагмент и скопировать его в какой-нибудь текстовый редактор стан-

⁸ Поэтому режимом сохранения полной страницы можно пользоваться для того, чтобы сохранять рисунки в случаях, когда меню правой кнопки мыши не действует из-за технических особенностей страницы.

дартным образом – через буфер обмена («Правка» | «Копировать» и «Правка» | «Вставить»).

Полезно знать, что и распечатать можно не всю веб-страницу, а только выделенный фрагмент (рис. 5.11).

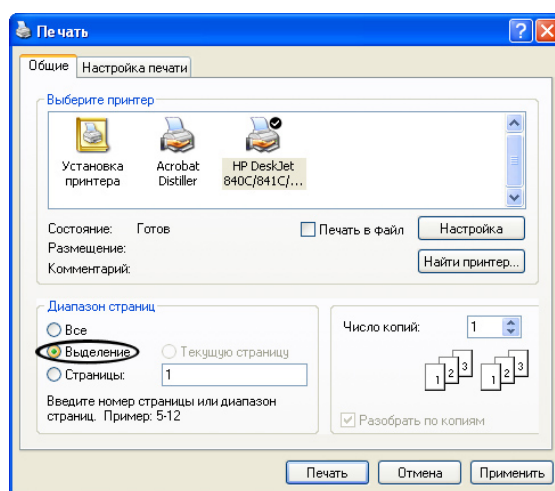




Рис. 5.11. Указание параметров печати

Оставшиеся нерассмотренными кнопки командной строки браузеров (рис. 5.1), как правило, используются редко. Можно лишь упомянуть кнопки перезагрузки и остановки загрузки страницы (для этого всегда используются значки  и ); также можно использовать клавиши **Esc** для остановки загрузки и **F5** для перезагрузки). Перезагрузка страницы полезна тогда, когда ее содержание должно было измениться в течение короткого времени. Дело в том, что любой браузер, принимая файлы, сохраняет их на жестком диске и обычно не загружает их снова в случае, если пользователь скоро вернется к той же странице. Количество сохраняемых таким образом страниц и срок, в течение которого они хранятся, зависят от настроек браузера. Обычно это очень удобно, однако если за прошедшее время страница изменилась, то браузеру нужно дать специальную команду на ее перезагрузку.

11.2. ЭЛЕКТРОННАЯ ПОЧТА

Основные понятия

Электронная почта (E-mail) представляет собой одно из средств общения между людьми, имеющими доступ к Сети. Это электронный аналог обычной почты, но электронная почта предоставляет и дополнительные возможности, например, возможность автоматической рассылки одного текста по целому списку адресов. Электронное сообщение передается в любую точку планеты практически мгновенно, и скорость получения письма зависит только от того, когда получатель заглянет в свой почтовый ящик⁹.

Адрес электронной почты состоит из двух частей и представляется следующим образом: *пользователь @ имя сервера*.

Во второй части адреса дается имя (адрес) *сервера* – того компьютера, на котором расположен электронный почтовый ящик. Адрес электронной почты не содержит пробелов, но может содержать точки, знаки подчеркивания ("_") и заглавные буквы, в том числе и в имени пользователя. Например, адрес *a_ivanov@hist.msu.ru* содержит имя пользователя – *a_ivanov* и адрес сервера – в данном случае это *hist.msu.ru* – сервер исторического факультета МГУ.

Электронный почтовый ящик – это специальная область памяти сервера, которая используется для хранения сообщений. Как правило, сервер обслуживает сразу множество адресов электронной почты, и все приходящие каждому пользователю письма накапливаются в соответствующем почтовом ящике и лежат там до востребования. Когда пользователь собирается проверить почту, он запускает специальную почтовую программу, которая соединяется с сервером, программа-обработчик сервера выбирает письма из почтового ящика именно этого пользователя и передает почтовой программе, которая уже может сохранять их в соответствующих рабочих папках, предоставляя пользователю все возможности работы с письмами.

Когда же пользователь отправляет письмо, почтовая программа снова соединяется с почтовым ящиком на сервере и передает письмо туда. В почтовом ящике это письмо ставится в очередь отправления, и сервер обеспечивает его правильную доставку до почтового ящика получателя. Как правило, это занимает секунды или доли секунд. Если сервер не может по каким-либо причинам сразу соединиться с указанным в адресе почтовым

⁹ Во многих почтовых программах предусмотрены вспомогательные возможности для привлечения внимания адресата: почтовая программа постоянно проверяет почтовый ящик и, например, подает звуковой сигнал в тот момент, когда приходит новое сообщение.

ящиком (потому, например, что адрес несуществующий – ошибочный или устаревший), он формирует письмо-отказ и помещает его в почтовый ящик отправителя.

Для организации почтового ящика нужно иметь почтовый сервер – как уже сказано, это компьютер, на котором выделено специальное место под письма пользователей и установлены соответствующие программы-обработчики. Доменное имя в почтовом адресе пользователя указывает именно на этот сервер. В частности, множество серверов в Интернете предлагает бесплатную услугу по предоставлению почтового ящика; часто это крупные веб-порталы (такие, как yandex.ru, mail.ru, gmail.com и др.), предоставляющие эту услугу наряду с другими (поисковая система, новости, словари и проч.).

Кроме самого почтового ящика, для работы с почтой нужно иметь почтовую программу, которая связывается с почтовым сервером, принимает и отправляет почту, а также осуществляет и другие процедуры по обработке почты. Так, почтовые программы позволяют использовать *фильтры*, с помощью которых сообщения автоматически сортируются и раскладываются по разным папкам. Все почтовые ящики на бесплатных серверах оснащены своими почтовыми программами, которые установлены на этих серверах и совмещены с программами-обработчиками почты, так что пользователь может работать со своими письмами просто посредством интернет-браузера, не устанавливая специальной почтовой программы. Впрочем, практически ко всем таким почтовым ящикам каждый может подключить и свою почтовую программу, и это бывает удобно, потому что специальные почтовые программы иногда позволяют осуществлять больше полезных функций по обработке и хранению писем.

Использование бесплатных почтовых серверов широко распространено: оно удобно тем, что для доступа к почте не нужно ничего, кроме подключенного к Интернету компьютера с интернет-браузером. Для того, чтобы завести почту на любом бесплатном сервере, нужно пройти процедуру регистрации (рис. 5.12). Главное, что нужно задать при регистрации – это *логин* и *пароль*.

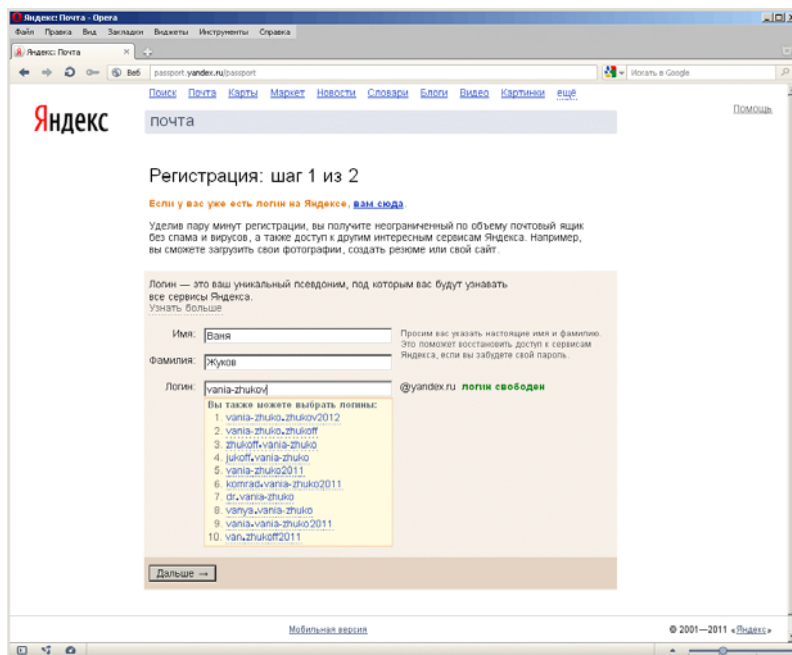


Рис. 5.12. Первый шаг регистрации пользователя в почтовой системе Яндекс

Логин – это уникальное имя пользователя, оно составляет первую часть адреса электронной почты. Как уже говорилось, для логина используются только латинские буквы, цифры и некоторые рабочие символы. При регистрации следует иметь в виду, что большинство простых логинов уже разобраны, поэтому просто имя или даже фамилия часто бывают “заняты”.

Пароль – это комбинация букв и цифр, которая обеспечивает доступ к ящику. Для пароля желательно (иногда обязательно) также использовать только латинские буквы и цифры; у некоторых почтовых серверов встречаются и другие требования к паролю (не короче определенного набора символов, не должно быть совпадения с логином, должен содержать цифры и т. п.). Конфиденциальность пароля – одно из важнейших требований к почтовым системам, поэтому при наборе пароля на экране отображается только последовательность звездочек. Для того чтобы пользователь был уверен, что он не сделал опечатки при первоначальном вводе пароля и ввел именно то, что хотел, при любой регистрации введено правило дублирования пароля. Забывать свой пароль не стоит, потому что в большинстве слу-

чаев никто не поможет восстановить забытый пароль, и все письма могут просто стать навсегда недоступными.

Бывает, что для разрешения подобной ситуации при регистрации вводится **контрольный вопрос** с ответом на него. Тогда в случае, если пользователь забудет свой пароль, можно сделать специальный запрос, и система задаст контрольный вопрос. Если ответ совпадет с тем, который был введен при регистрации, то система позволит задать новый пароль и, соответственно, даст доступ к почтовому ящику. С одной стороны, это может помочь тому, кто действительно забыл свой пароль, но, с другой стороны, это дает лазейку потенциальному взломщику почтового ящика, поэтому при регистрации лучше даже на простой контрольный вопрос давать неожиданный ответ, даже если он не соответствует действительности; ведь главное тут – не честность, а достижение поставленной цели при соблюдении базовых правил обеспечения безопасности информации.

Кроме того, при регистрации часто требуется ввести набор букв или цифр, криво или с помехами изображенных на приложенном тут же рисунке (так называемая **капча**). Это делается для защиты пространства, отведенного под почтовые ящики, от автоматической регистрации фиктивных адресов специальными программами.

Стандартные функции почтовых программ

Все почтовые программы, где бы они ни находились, обязательно выполняют следующий набор функций:

1. Прием почтового сообщения. Программа показывает список пришедших сообщений, в котором указано, от кого и когда они пришли, а также показана "Тема" каждого сообщения. Тема специально задается отправителем письма именно для того, чтобы уже при просмотре списка было видно, о чем говорится в письме.

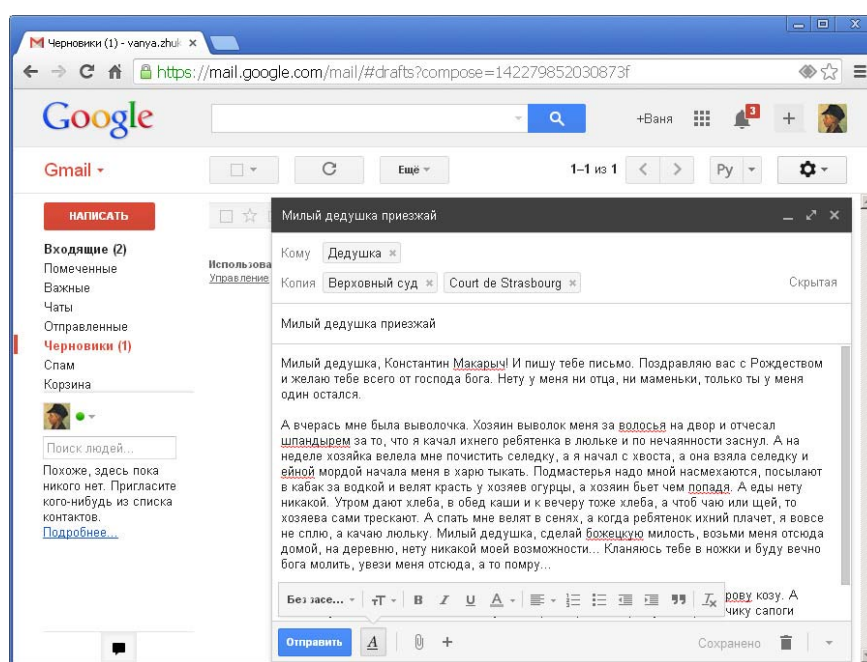


Рис. 5.13. Отправка письма в почтовой системе Gmail

2. Отправка сообщения. При отправке обязательно заполняется поле "Кому" с электронным адресом получателя. Через запятую¹⁰ можно записывать несколько адресов, по которым следует послать это же письмо, к тому же для записи дополнительных адресов дается и вторая строка, а также обычно есть еще третья строка "Скрытая копия"— то же самое, только так, что адресаты, указанные в первых строках, не увидят, что копию получил кто-то еще. В отдельной строке желательно указать тему сообщения (см. рис. 5.13 и 5.14).

¹⁰ Разделителем может служить и точка с запятой – ";".

Раздел V

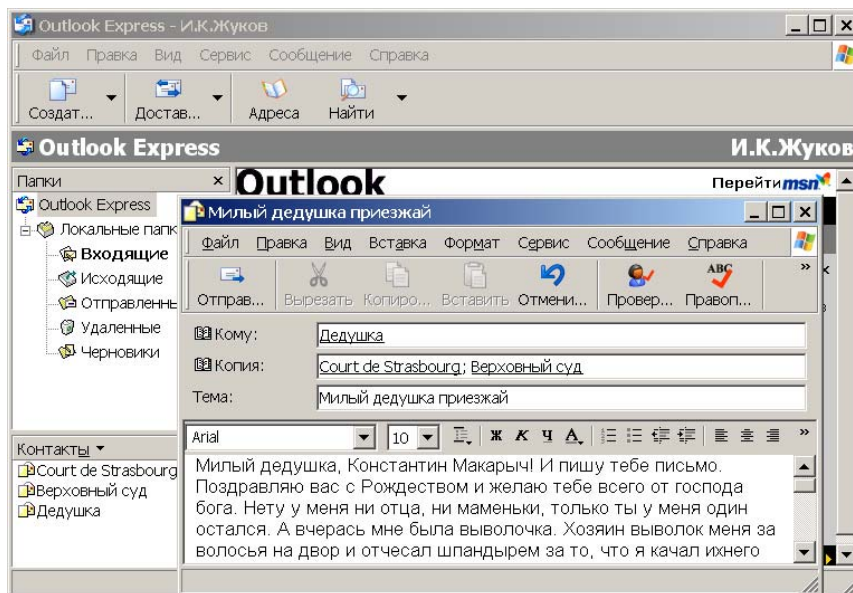


Рис. 5.14. Отправка того же письма через программу Microsoft Outlook Express (Собственно, здесь открыты два окна – на заднем плане окно программы и на переднем – окно самого письма)

3. Прочитав принятое письмо, пользователь может использовать команды «Ответить» или «Переслать». В первом случае организуется отправка письма, в которое автоматически будет подставлен адрес отправителя, а в тексте цитируется полученное письмо. Это удобно, например, для ответа на письмо по пунктам. При пересылке же текст полученного письма полностью подставляется в новое письмо, которое можно разослать по другим адресам.

4. Вместе с письмом-текстом можно отправлять файлы разных форматов. Такие файлы-посылки называются **прикрепленными файлами** или **вложениями**.

ГЛАВА 12. НАВИГАЦИЯ В СЕТИ ИНТЕРНЕТ. ПОИСКОВЫЕ СИСТЕМЫ

Простое перемещение по известным адресам Интернета или по ссылкам с них не позволяет достаточно эффективно и быстро находить нужную информацию в огромном количестве веб-ресурсов. Для решения этой проблемы были созданы информационно-поисковые системы. Поисковые системы собирают огромные, автоматически пополняемые базы данных, связывающие адреса веб-страниц со словами, словосочетаниями и изображениями, которые на них содержатся. Задавая запрос по ключевым словам, пользователь получает список адресов веб-страниц, на которых были встречены эти ключевые слова.

Сайтов, которые представляют возможность провести поиск по запросам в Интернете, довольно много, но сложность задачи постепенно привела к существенному сокращению поисковых систем, способных поддерживать собственную базу данных; при этом на многих сайтах, и в том числе весьма популярных, с небольшими изменениями используются поисковые базы данных одних и тех же крупных систем. Поэтому для пользователей Рунета (русскоязычного сегмента Интернета) сейчас* представляют наибольший интерес поисковые системы, поддерживающие свои оригинальные базы данных: Google (www.google.ru), Яндекс (www.yandex.ru), Bing (www.bing.com) и Mail.ru (go.mail.ru)¹¹. Как видно на рис. 5.15–5.17, разные системы представляют хотя и пересекающиеся, но вместе с тем и различные результаты на один и тот же запрос.

* Написано осенью 2013 г.

¹¹ Есть и другие поисковые системы, имеющие значительную популярность. Так, можно отметить еще два поисковика, в свое время внесших существенный вклад в развитие Интернета: американский Yahoo! (с 2009 г. использует поиск от Bing, однако по-прежнему превосходит Bing по популярности) и русский Рамблер (с 2011 г. использует поиск от Яндекса). Все эти поисковые системы уверенно работают с запросами на русском языке, но полезно знать и то, что сейчас третьей по обращаемости (т.е. по количеству ежедневных запросов) поисковой системой в мире (после Google и Yahoo!) является Baidu, ориентированная на китайский и японский сегменты Интернета.

Раздел V

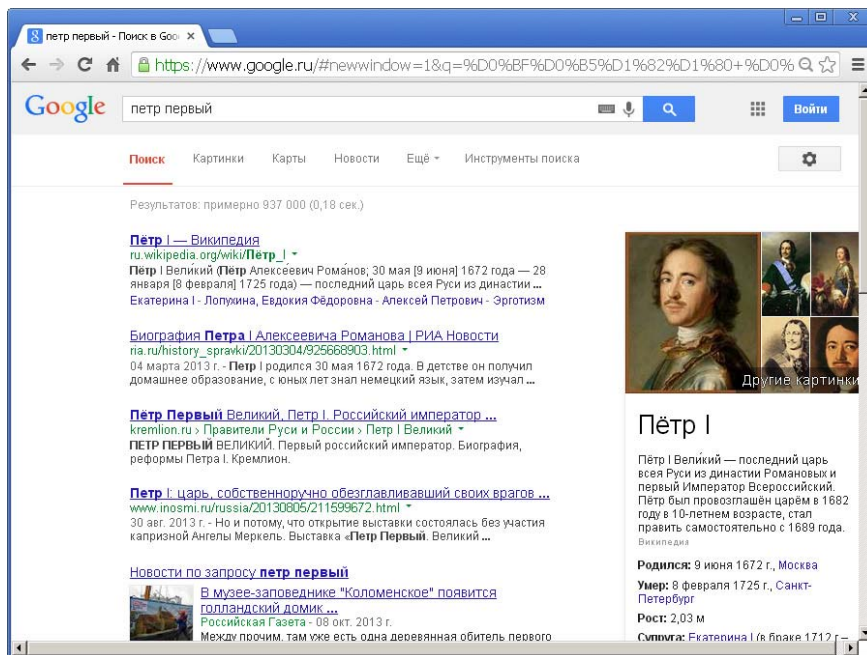


Рис. 5.15. Результаты поиска в системе Google

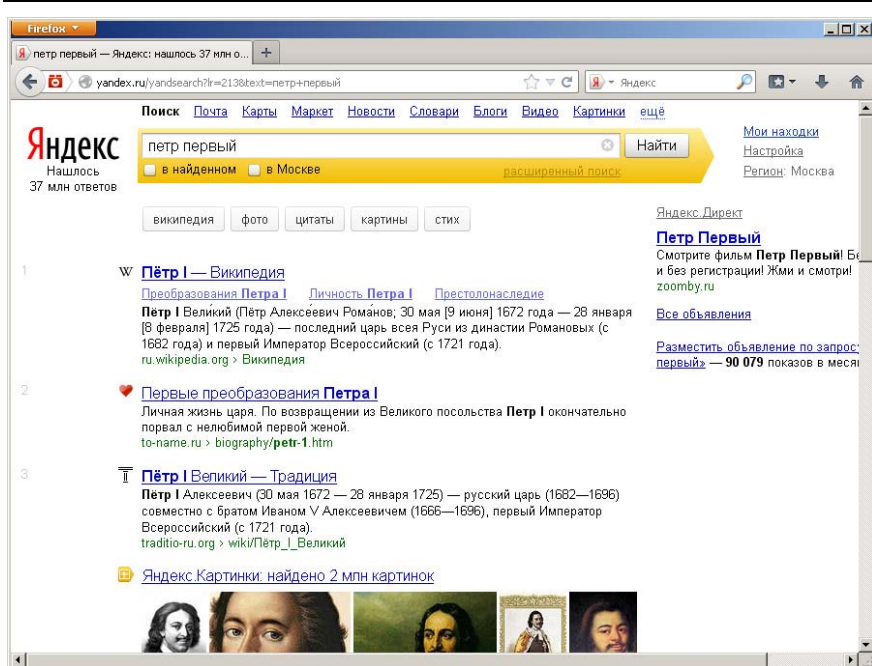


Рис. 5.16. Результаты поиска в системе Яндекс

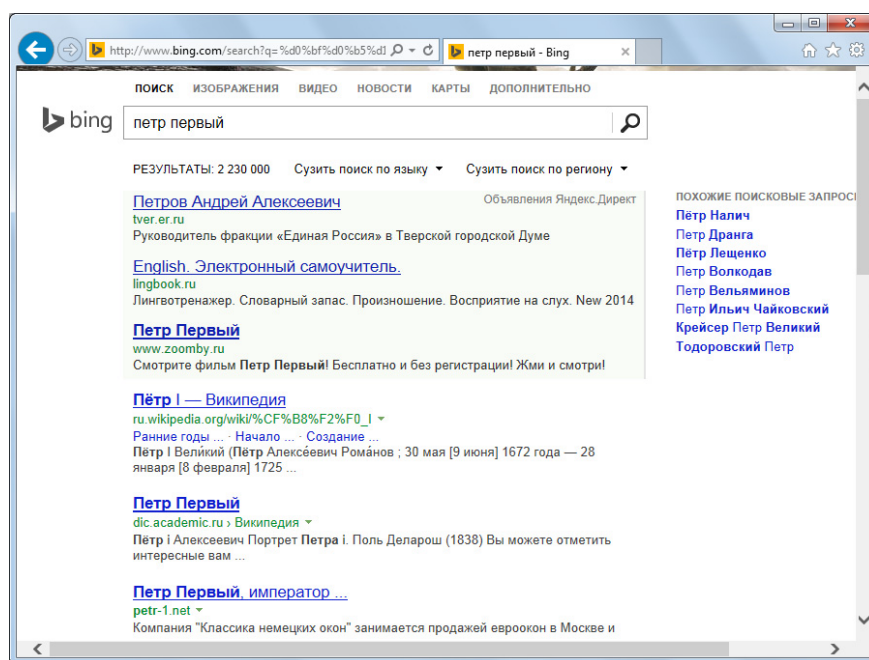


Рис. 5.17. Результаты поиска в системе Bing

12.1. ЗАПРОСЫ К ПОИСКОВЫМ СИСТЕМАМ

Общими для большинства поисковых систем являются правила формирования и выдачи результатов запросов. На главной странице сайта (рис. 5.15–5.17) находится окно для задания запроса. Пользователь набирает в нем текст запроса, нажимает на кнопку «Поиск» (или на клавишу **Enter**) и получает список найденных адресов. Список выдается страницами: по умолчанию обычно выдается по 10 адресов на одной странице, но эту настройку можно изменить). Переход к следующим страницам находится внизу списка. Для каждой найденной записи, входящей в список, представлены гиперссылка на нее и ее краткое описание. В описание в разных системах включается разная информация, характеризующая ресурс в целом или конкретный электронный документ.

Однако на самый простой запрос можно получить десятки тысяч ссылок, и важно из полученного списка вычленить *релевантные*, т. е. наиболее точно соответствующие запросу. Наиболее острой проблема релевантности является для сложных запросов, в которых фигурируют комбинации

слов и выражений. Скажем, если пользователь ищет информацию о Петре I, вводя в окне поиска *Петр Великий*, поисковый механизм может выдать ссылки на ресурсы, в которых слова *Петр* и *великий* разнесены и вообще не связаны между собой. В приведенном случае эти ссылки будут лишними, но можно придумать множество примеров, где такая работа поисковой машины поможет пользователю не упустить действительно полезные адреса. С другой стороны, ресурсы, в которых упомянут крейсер "Петр Великий", являются, конечно, тоже релевантными, так как механически они полностью соответствуют запросу. Поэтому если необходимо найти ресурсы, относящиеся непосредственно к личности императора Петра I, пользователю следует более аккуратно формулировать запрос.

Строго говоря, следует различать **релевантность** и **пертинентность** результатов запроса. Хотя и то, и другое говорит о качестве поиска, первый показатель говорит о том, насколько успешно поисковая система находит ресурсы по найденному запросу, а второй – насколько пользователя удовлетворяет та информация, которую ему удалось найти. Если пертинентность низкая, то это часто означает, что запрос был сформирован недостаточно удачно.

Часто поисковые системы позволяют провести поиск точного словосочетания, для чего в запросе это словосочетание нужно, например, записать в кавычках, хотя и это тоже может дать нерелевантный результат. Частичным решением проблемы нахождения релевантных запросу ресурсов является способность поисковых систем упорядочивать результаты поиска так, чтобы первыми были выданы наиболее подходящие запросу результаты. Поисковые системы постоянно совершенствуются и умеют выдавать все больше релевантных результатов даже на не очень толковые запросы.

В частности, нужно иметь в виду, что современные поисковые механизмы обязательно ведут статистику запросов и востребованных переходов по запросам, а также используют различные другие приемы для повышения релевантности выдаваемых ими результатов. Это приводит к тому, что поиск чего-то более востребованного делается более легким, а менее востребованного – более сложным. То есть, если говорить о приведенном выше примере, запрос *Петр Великий* почти во всех системах дает половину ссылок на информацию о крейсере и половину – об императоре, эти значения считаются примерно одинаково релевантными. Однако если написать *Петр первый*, все первые ссылки будут так или иначе связаны с императором, причем на некоторых страницах первого списка выдачи, возможно, даже нет непосредственно этих слов, заменяемых уже поисковой системой записями "Петр I" и "Петр 1", зато какие-то другие объекты с тем же названием, вроде московской гостиницы или российского круизного теплохода,

окажутся далеко внизу в списке результатов, хотя они, строго говоря, являются не менее релевантными. Поэтому если мы ищем что-то, что по имени/названию совпадает с более очевидными и популярными объектами, лучше сразу уточнить поиск дополнительными ключевыми словами. Очевидно, что запрос *Петр первый гостиница* приведет к искомой гостинице сразу.

В сложных случаях от грамотного пользователя зависит по-прежнему много, и весьма важно уметь задавать такие запросы, чтобы, во-первых, было найдено как можно меньше ссылок на нерелевантные ресурсы, и, во-вторых, как можно меньше лишних, формально соответствующих запросу, но ненужных по существу.

Поэтому нужно знать, что у каждой поисковой системы есть свой синтаксис запросов, и можно существенно повышать эффективность поиска, используя правила языка запросов и учитывая особенности работы конкретной поисковой системы. Кроме того, существуют и другие дополнительные возможности. Конечно, сами поисковые системы предоставляют справку о собственных поисковых инструментах:

- Google: <https://support.google.com/websearch/>
- Яндекс: <http://help.yandex.ru/search/>
- Bing: <http://onlinehelp.microsoft.com/ru-ru/bing/>
- Mail.ru: <http://help.mail.ru/go-help/>

12.2. ДОПОЛНИТЕЛЬНЫЕ УТОЧНЕНИЯ: РАСШИРЕННЫЙ ПОИСК

В настоящее время все поисковые системы поддерживают более или менее широкие возможности расширенного поиска, когда в запрос включаются дополнительные условия. Выход в расширенный поиск открывается, когда проведен какой-нибудь простой поиск. Частично возможности расширенного поиска можно получить и через использование команд языка запросов (об этом во всех случаях см. ниже, п. 12.3), но выход в расширенный поиск обычно открывает новое наглядное меню, в котором можно заполнить строки с уточнениями запроса (рис. 5.18–5.20).

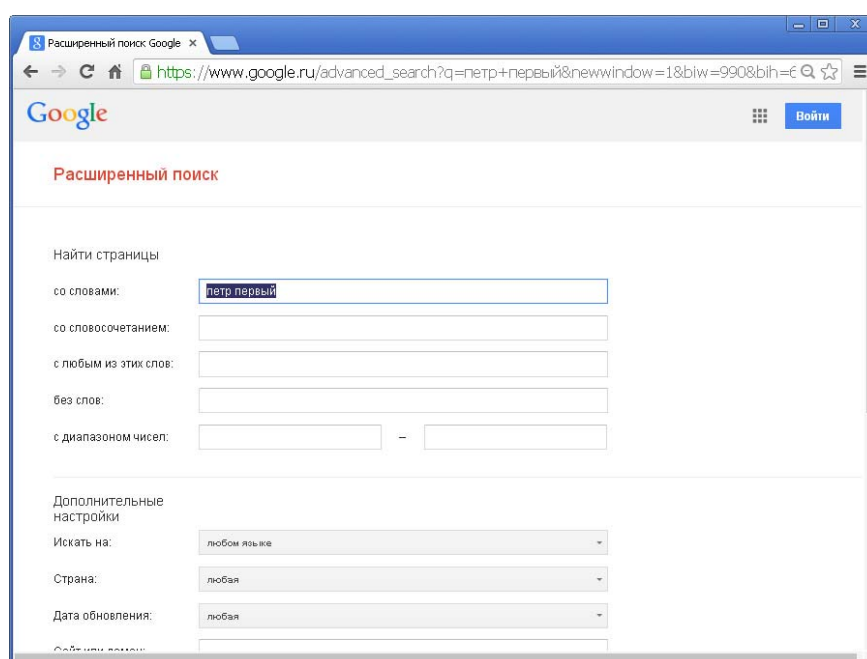


Рис. 5.18. Расширенный поиск в системе Google

Раздел V

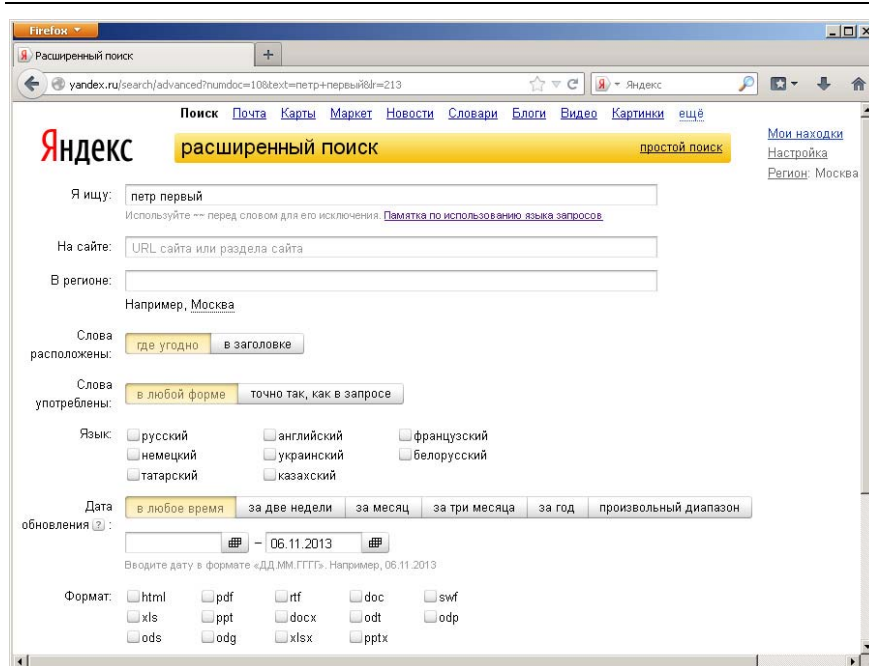


Рис. 5.19. Расширенный поиск в системе Яндекс

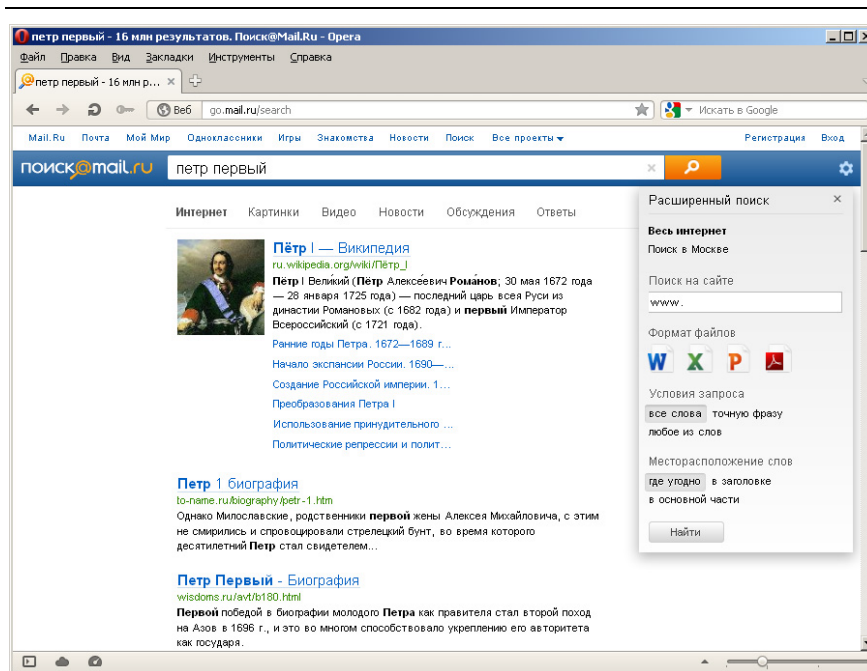


Рис. 5.20. Расширенный поиск в системе Mail.ru

Одним из наиболее полезных уточнений, без сомнения, является уточнение места расположения ресурса. Здесь встречаются два вида уточнений: по доменному имени или на конкретном сайте.

Поиск в определенной доменной зоне полезен, когда мы знаем, что некоторый набор сайтов связан с определенным доменом. Например, мы можем уточнить запрос так, чтобы искать информацию о подготовительных курсах только на сайтах факультетов МГУ, т. е. внутри доменов *msu.ru*. В расширенном поиске системы Google мы можем тогда записать *.msu.ru* в строке «Сайт или домен». В Яндексе можно сделать это только с помощью оператора запроса, но не через расширенный поиск. Конкретизировать поиск так можно и до доменного имени верхнего уровня (например только в доменной зоне "ru"), хотя вместо этого, может быть, порой полезней провести поиск страниц на избранном языке.

Другой вариант – когда мы знаем, что нужная нам информация находится на известном нам сайте. Тогда мы вписываем адрес этого сайта: в Google – в той же строке «Сайт или домен», в Яндексе – в строке «На сайте», в Mail.ru – «Поиск на сайте». В первых двух из этих систем можно

уточнить еще и раздел сайта (например провести поиск только в электронной библиотеке исторического факультета МГУ можно, если добавить к запросу запись *www.hist.msu.ru/ER/*). Но такое уточнение сработает только если структура искомого сайта устроена таким образом, что файлы разложены по отдельным папкам, а это верно далеко не для каждого сайта.

Пример 5.1. Кто автор фигуры, имеющей номер 002182786 в Государственном каталоге музейного фонда РФ?

Очевидно, что сведения об искомом объекте должны находиться именно на сайте этого Каталога. Легко узнать, что искомый сайт – *gokatalog.ru*, и можно провести поиск именно на нем: *002182786 site:gokatalog.ru*. Результат запроса оказывается очень точным, ссылка приводит непосредственно на описание объекта:

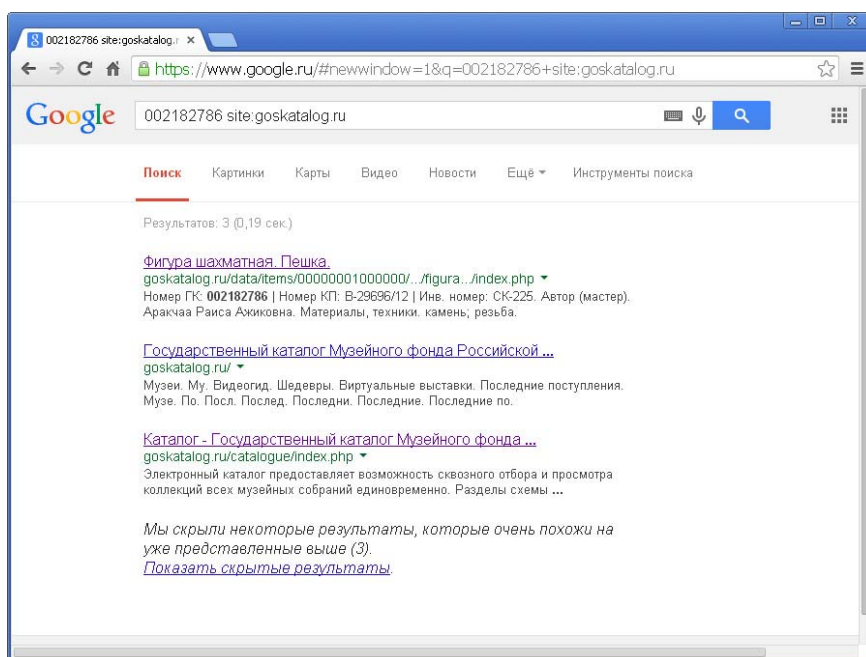


Рис. 5.21. Уточненный поиск по сайту *gokatalog.ru* в системе Google

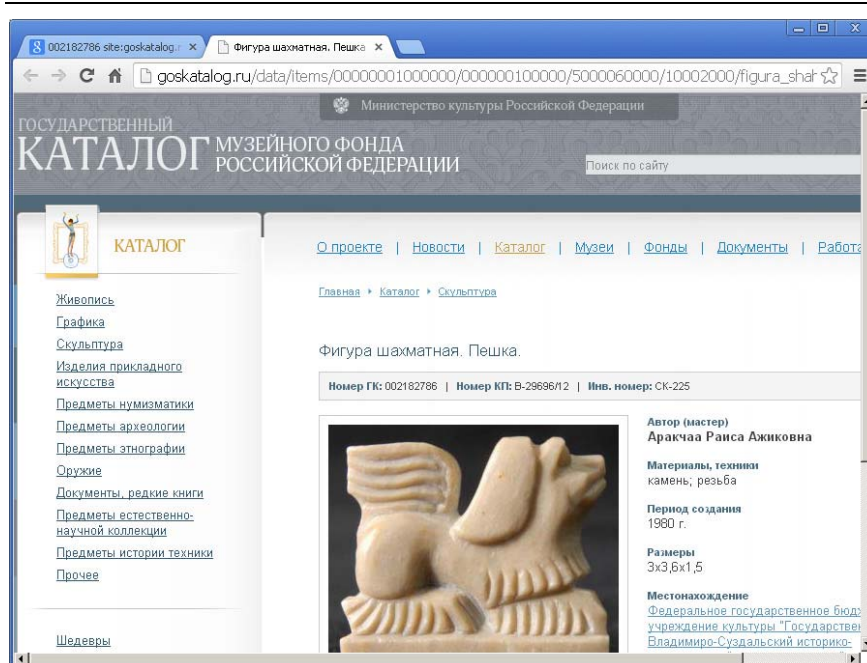


Рис. 5.22. Результат поиска на сайте goskatalog.ru

Заметим, кстати, что Яндекс и Mail.ru не справились с этим запросом, а Bing и вовсе не имеет такого уточнения.

Раздел V

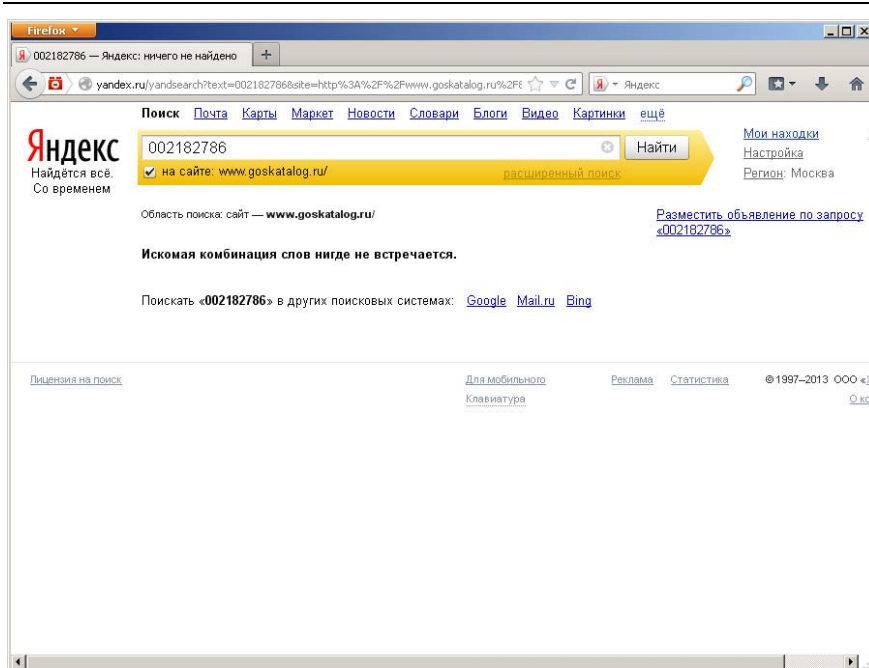


Рис. 5.23. Неудавшийся поиск по сайту goskatalog.ru в системе Яндекс

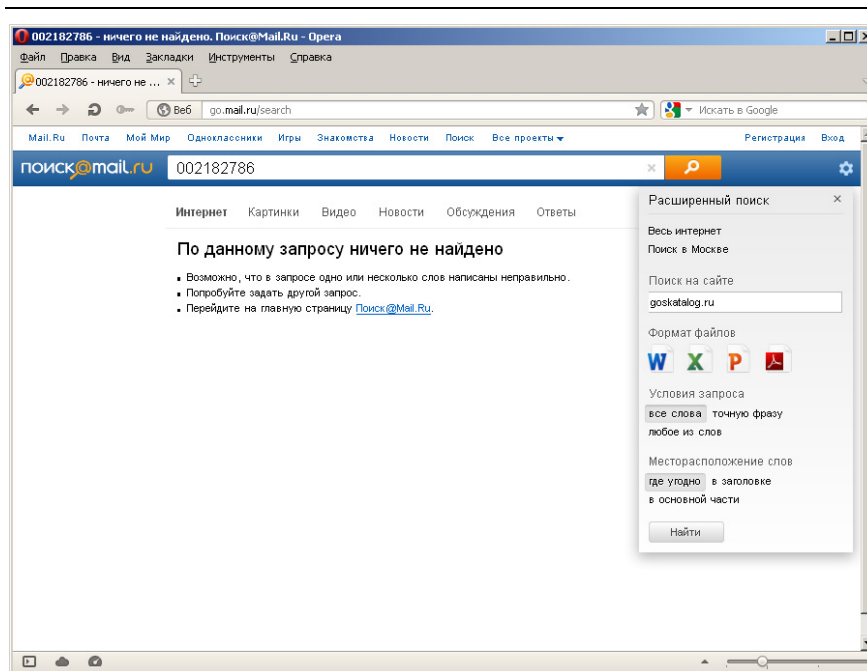


Рис. 5.24. Неудавшийся поиск по сайту gskatalog.ru в системе Mail.ru

Когда искомое ключевое слово или словосочетание очень широко распространено, и найденных результатов оказывается слишком много, может оказаться полезным требование поиска только в заголовке документа. Впрочем, современные поисковые системы и так считают более релевантными результаты, в которых слова из запроса находятся в заголовке, и обычно выводят такие результаты вперед.

Однако нужно иметь в виду, что поиск информации в Интернете не всегда нужно целиком исполнять с помощью только лишь средств поисковой системы. Именно потому, что наиболее ценными для поисковых систем являются слова, находящиеся в заголовках страниц, в первых абзацах страниц, повторяющиеся много раз по тексту и т.п., другие слова найти сложнее. Часто, если нам известно, в каком тексте должна быть искомая информация, проще найти вообще этот текст, а информацию искать в нем уже с помощью обычного контекстного поиска (**Ctrl** + **F**)

Пример 5.2. Кто назначался ответственным за охрану порядка на сельских сходах по Положению 19 февраля 1861 года?

В указанном примере трудно найти комбинацию ключевых слов, которая бы привела непосредственно к нужному ответу. Очевидно, однако, что ответ содержится в самом тексте Положения, поэтому проще найти сначала текст, а потом уже в нем без всякой специальной поисковой системы провести поиск какого-то не слишком часто встречающегося буквосочетания, например *поряд* (мы ведь не знаем, в каком падеже записано необходимое слово в источнике). Ответ находится довольно быстро: это сельский староста.

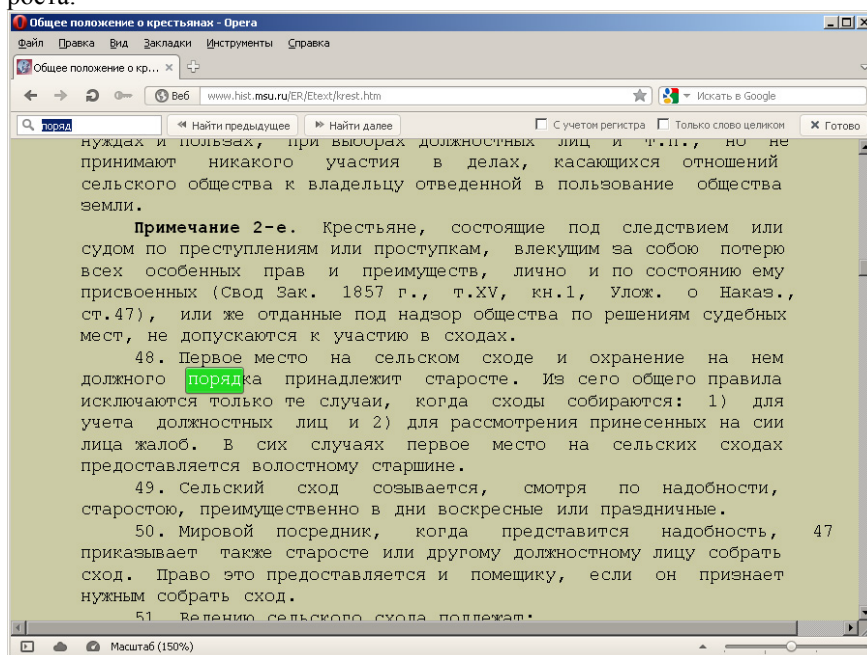


Рис. 5.25. Поиск информации в известном тексте

12.3. СИНТАКСИС ЯЗЫКА ЗАПРОСОВ

Поисковые системы вообще ориентированы на максимальное облегчение задачи пользователю, в них развиваются и все лучше действуют алгоритмы, позволяющие выбрать среди огромного количества найденных результатов более релевантные запросы и дать пользователю поиск именно к ним. Тем не менее, в сложных случаях оказывается очень полезным использовать специальный язык запросов.

Во всех системах действуют примерно одинаковые правила формирования запросов, хотя из рассматриваемых Яндекс и Google предлагают больше изощренных средств поиска. Детальное описание всех возможностей языка запросов можно увидеть в разделах помощи соответствующих систем; мы же здесь опишем только наиболее полезные из этих операторов.

В первую очередь, в синтаксисе запросов любой поисковой системы присутствуют средства логического объединения слов – операторы *И*, *ИЛИ*, *НЕ* (*И-НЕ*), группировка слов для работы с этими операторами и т. д. Оператор *И*, который означает, что все ключевые слова, указанные в запросе, должны присутствовать в найденных страницах, чаще всего задается просто пробелом: если пользователь написал два слова в запросе, поисковая система полагает, что они оба ценны, причем лучше, чтобы они находились рядом и в указанной последовательности; для пущей уверенности можно взять поисковую фразу в кавычки и усилить это требование в поисковой системе. Однако если результатов с обоими словами находится мало, поисковые системы обычно переходят к выдаче страниц, где встречено хотя бы одно из слов, то есть переходят к использованию оператора *ИЛИ*. По-настоящему полезным является оператор *И-НЕ*, который позволяет сократить поиск: найти главное слово, но убрать другие, которые мешают поиску.

Например, мы хотим найти информацию о разных мавзолеях, появившихся в истории. Простой запрос *мавзолей* даст массу страниц с наиболее популярными из них – мавзолеем Ленина и мавзолеем в Галикарнасе. Чтобы отсеять их, стоит использовать знак минуса (без пробела) перед словами, которые мы хотим выбросить из запроса: *мавзолей -Ленин -Галикарнас*. Результат будет совсем другим (рис. 5.26–5.27).

Раздел V

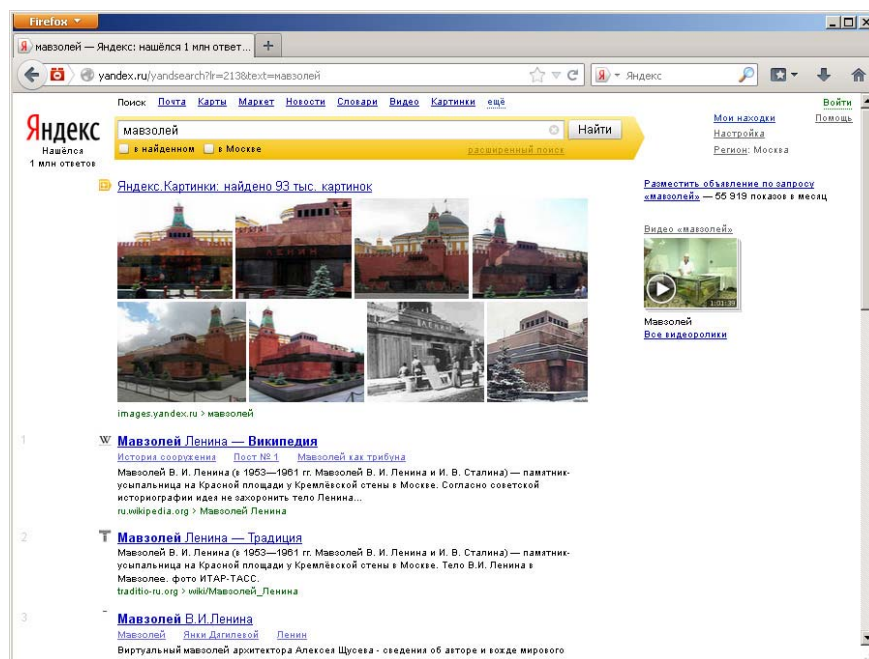


Рис. 5.26. Поиск без использования оператора «И-НЕ»

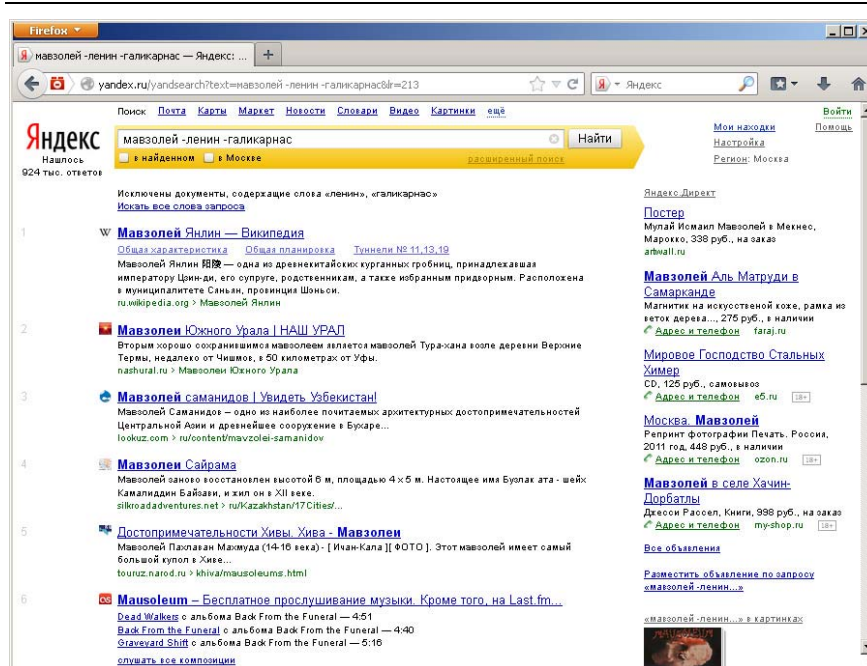


Рис. 5.27. Поиск с использованием оператора «И-НЕ»

Современные поисковые системы производят поиск ключевых слов во всех известных им словоформах, одновременно, как правило, они совершенно игнорируют часто встречающиеся и не имеющие самостоятельного значения служебные слова (*стоп-слова*) – предлоги, союзы, отдельные цифры и буквы и т. д. Однако в запросе можно специально обозначить требование, чтобы искомое слово обязательно было найдено. Чаще всего для этого нужно приписать знак "+" перед нужным словом (без пробела).

Бывает, что в системе предусмотрены способы точного поиска слова в той форме, в какой это слово задается пользователем. В Яндексе, например, для этого служит восклицательный знак. Запросы *!Ломоносова* или *!"имени Ломоносова"* укажут поисковой системе, что не нужно производить поиск этих слов в других падежах. Если же поставить два восклицательных знака, то этим указывается нормальная форма слова. Запрос *Кузнецов* найдет сайты, содержащие не только фамилию "Кузнецов", но и слово "кузнец" в самых разных формах. Уточненный запрос – *!!Кузнецов* – укажет программе, что нужно ограничиться лишь различными словофор-

мами этой фамилии, так что в результаты не попадут ссылки на сайты со словами "кузнецы", "кузнец" и т. д. В результатах, однако, может содержаться слово "кузнецов" в значении слова "кузнец", если оно встречается на странице именно во множественном числе и родительном падеже.

Заглавные буквы в запросе также иногда имеют значение. Так, если запрос в Яндексе записан строчными буквами, проводится поиск как строчных, так и прописных букв, если же записана прописная буква, то система ищет только такие же варианты. Например, по запросу *Орел* в Яндексе будут найдены сайты с упоминанием города Орел или же птицы, но только в тех случаях, когда слово "Орел" начинается с прописной буквы (например, в начале предложения).

В синтаксисе запроса порой предусмотрены и те возможности, которые предлагаются и в расширенном поиске. Так, для поиска на определенном сайте в системе Google используется оператор "site", а в Яндексе – целый набор операторов¹², причем оператор ставится перед названием сайта (домена, части сайта и т.п.) через двоеточие, обязательно без пробелов: *site:hist.msu.ru*.

Полезным оператором, пока имеющимся только в системе Google, является поиск по диапазону чисел. Если задать два числа через две точки (без пробелов: *1920..1929*), то система будет искать любые числа, входящие в этот диапазон. Так можно найти, в частности, событие, происшедшее в некотором хронологическом диапазоне, даже если пользователь не знает точной даты.

Пример 5.3. *В каком году XVII века был убит царь Александр II?*

Если мы не знаем, где правил этот царь, то какие ключевые слова мы ни выберем, на первых страницах выдачи результатов будет только информация про Александра II Романова, погибшего в 1881 г. Однако указание на то, что событие произошло в XVII веке, позволяет нам указать нужный нам диапазон лет, пусть и довольно большой: 1601–1700. Мы немедленно находим царя Кахетии Александра II, убитого в 1605 г.

¹² См. <http://help.yandex.ru/search/query-language/search-operators.xml>

Интернет: информационные ресурсы и сервисы

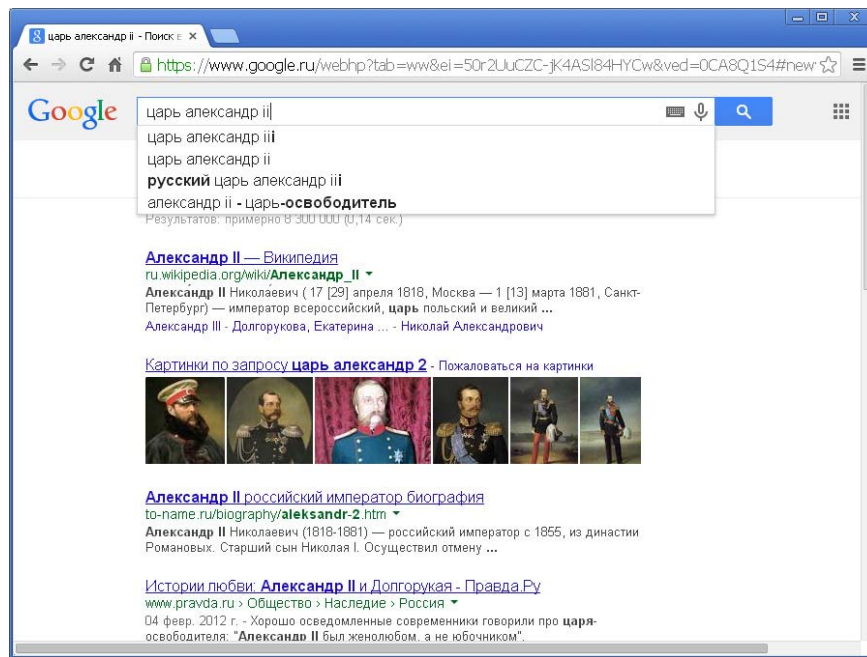


Рис. 5.28. Поиск без уточнения диапазона дат

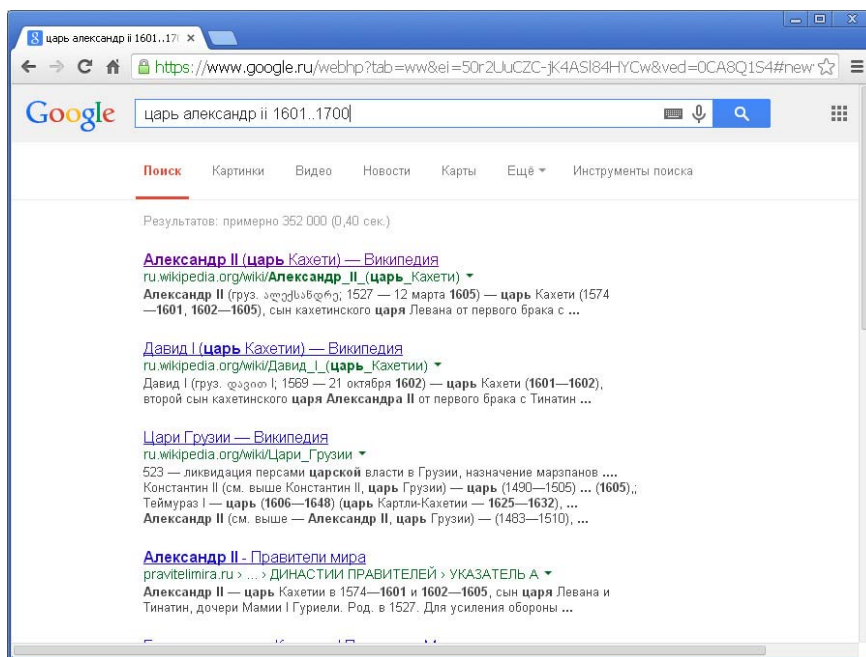


Рис. 5.29. Поиск с уточнением диапазона дат

12.4. КЛЮЧЕВЫЕ СЛОВА ЗАПРОСА

Знание синтаксиса языка запросов используемой поисковой системы помогает существенно улучшить результаты поиска, но не менее важно умение подобрать нужные ключевые слова – такие, которые позволят найти нужные ресурсы и отсеять ненужные.

Так, например, если тема, по которой нужно собрать сведения, звучит как *"Древняя Спарта: общественно-политический строй и внешняя политика"*, это ни в коем случае не означает, что всю тему можно вводить в строку запроса именно в таком виде. Из всех слов в этой теме только *"Спарта"* определенно является ключевым словом, полезным для первоначального поиска; с некоторой вероятностью будет полезно также словосочетание *"Древняя Спарта"*, но словосочетание *"внешняя политика"* уведет явно в сторону и даст огромное количество лишних документов. Дело в том, что для поисковой системы все записанные в запросе слова обладают равной ценностью, и если все слова данной темы будут включены в запрос, то сначала поисковая система попытается найти страницу, на которой бы упоми-

нались все слова сразу, но, скорее всего, не сможет ее найти и будет искать страницы, на которых есть хотя бы два-три слова. В итоге она покажет страницы как про Древнюю Спарту, так и про внешнюю политику вообще и про любой общественно-политический строй. Таким образом, по имеющейся теме можно и нужно задавать ключевые слова, которые по смыслу должны вести к нужной информации, но которых нет в названии. Так, для темы про Спарту это могут быть *"законы Ликурга"*, *"спартанцы"*, *"илоты"*, *"Геродот"* и т. п.

Приведем другой пример: если тема звучит как *"Князь и княжеская власть по Судебнику Ивана III"*, то не имеет смысла писать в запросе про княжескую власть, здесь именно Судебник – наиболее полезное ключевое понятие. Конечно, само по себе слово *"Судебник"* даст слишком много лишней информации, так что нужно использовать словосочетание *"Судебник Ивана III"*, но поскольку он же может где-нибудь проходить под названием *"Судебник 1497 года"*, желательно провести два поиска – по обоим словосочетаниям.

В связи с последним примером полезно отметить, во-первых, что использование в запросе фразы *"1497 год"* в большинстве случаев хуже, чем использование просто числа *1497*. В первом случае мы ограничиваем себя тем, что обязательно ищем в запросе введенное слово *"год"*, тем самым игнорируя все страницы, на которых написано только *"1497 г."*. Во втором случае мы, правда, можем получить и страницы, на которых число *1497* появляется в другом контексте, вне привязки к интересующему нас событию, но таких случаев будет, скорее всего, не очень много, особенно если конкретизировать запрос другим ключевым словом *"Судебник"*. Наконец, использование в запросе записи *"1497 г."* вообще некорректно, потому что без специальных указаний поисковые системы, скорее всего, проигнорируют и отдельную букву *"г"*, и тем более точку как незначимые символы.

12.5. ПОИСК ИЗОБРАЖЕНИЙ

У всех поисковых систем сейчас имеются возможности поиска изображений, так или иначе соответствующих ключевым словам. Поиск этот не идеален, потому что сбор данных для поисковой системы ведется автоматически, а программа не способна распознавать содержимое изображений (хотя разработчики поисковых систем работают над частичным решением этой проблемы). Сейчас есть два способа, по которым поисковая система помечает в своей базе данных те или иные изображения как соответствующие каким-то ключевым словам. Во-первых, такой разметкой могут озаботиться создатели сайта, на котором находится изображение: язык html позволяет приписывать рисункам ключевые слова в формате, понятном поис-

Раздел V

ковым системам. Такая разметка, однако, имеется далеко не у всех изображений, так что вторым способом является их атрибуция по словам, находящим в непосредственной близости – в ближайшем абзаце, в одной ячейке таблицы и т.п. Здесь, конечно, встречаются самые разные ошибки, и поэтому всегда в галерее изображений по конкретному запросу находится значительное количество вовсе не относящихся к делу картинок. Однако в целом такой поиск работает неплохо. Разумеется, и здесь мы сталкиваемся с той же проблемой релевантности: по запросу "Наполеон" найдутся портреты и Наполеона I, и Наполеона III, а с ними и изображения коньяка и (ниже, не вошло в снимок экрана) торта «Наполеон».

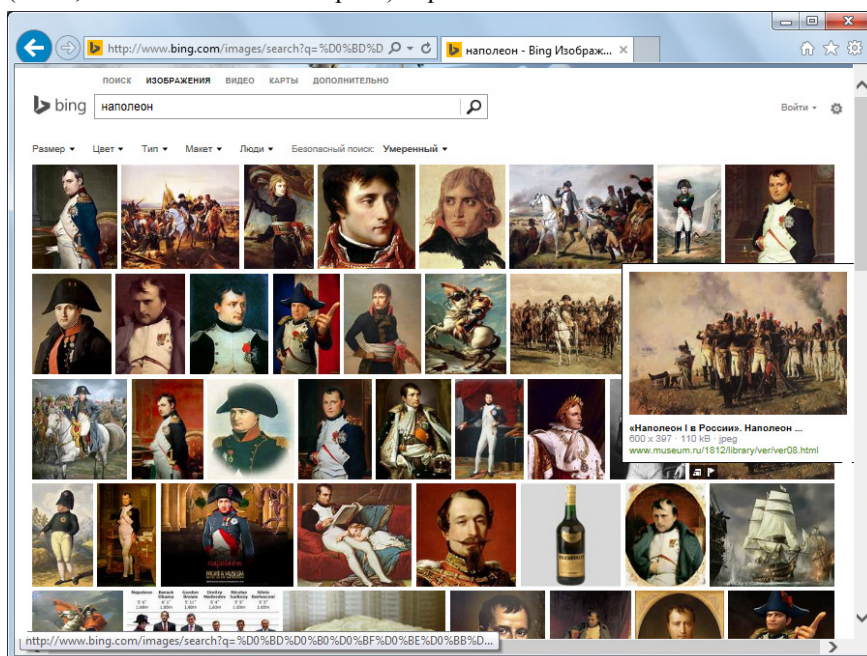


Рис. 5.30. Галерея картинок "Наполеон" в поисковой системе Bing

Существенно более сложной задачей для поисковых машин является обратная: по имеющемуся изображению найти информацию о нем. Впервые подобную задачу стала более или менее успешно решать специализированная именно на этом поисковая система TinEye (tineye.com), начавшая работу в 2008 г. На начало 2014 г. в ней, по официальной информации, учтено более 4,5 млрд. изображений. Чтобы задать системе поиск, можно ли-

бо загрузить нужное изображение в виде графического файла, либо указать точный интернет-адрес нужного файла, если он стоит на каком-либо сайте. Адреса найденных в Интернете изображений могут быть очень длинными и сложными, но обычно их можно скопировать в буфер памяти, используя меню по нажатию правой кнопкой мыши. Поисковая система сопоставляет полученный файл с другими имеющимися в базе, причем делает поправки на возможные отличия размера, цвета, полей, мелких деталей и надписей на рисунках.

Пример 5.4. *Найдите годы жизни представленного на портрете персонажа. (<http://www.hist.msu.ru/Departments/Inf/Quiz/person1.jpg>)*



Рис. 5.31. Исходный портрет для поиска по изображению

Чтобы ответить на поставленный вопрос, мы сообщаем адрес изображения поисковой системе (в правой части стартовой страницы),..

Раздел V

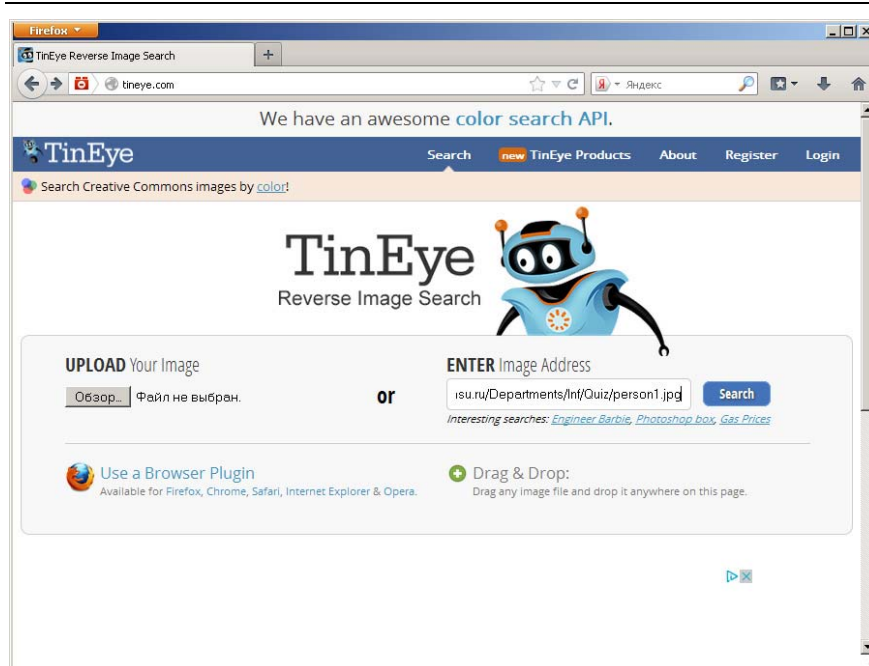


Рис. 5.32. Задание запроса в поисковой системе TinEye

...и получаем два результата, перейдя по которым мы узнаем, что это портрет работы А.Бронзино, и на нем изображен герцог Тосканы Козимо Медичи (1519–1574).

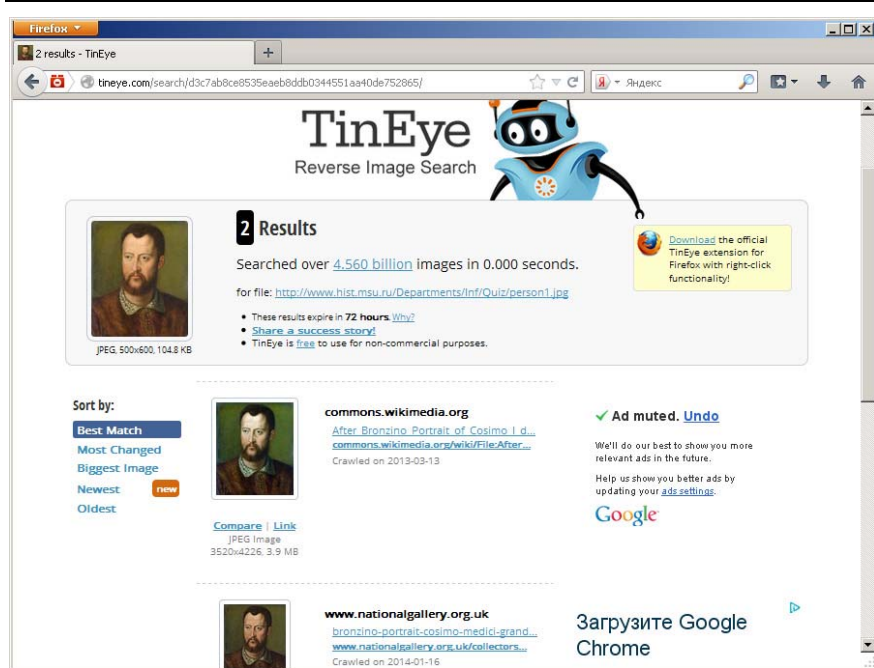


Рис. 5.33. Результат запроса в поисковой системе TinEye (2 сайта)

В 2011 г. ту же задачу поиска по изображениям стала решать поисковая система Google, и сейчас она выдает заметно больше результатов на тот же запрос. Для поиска нужно перейти в раздел «Картинки» и нажать на схематичное изображение фотоаппарата в правой части поисковой строки.

Раздел V

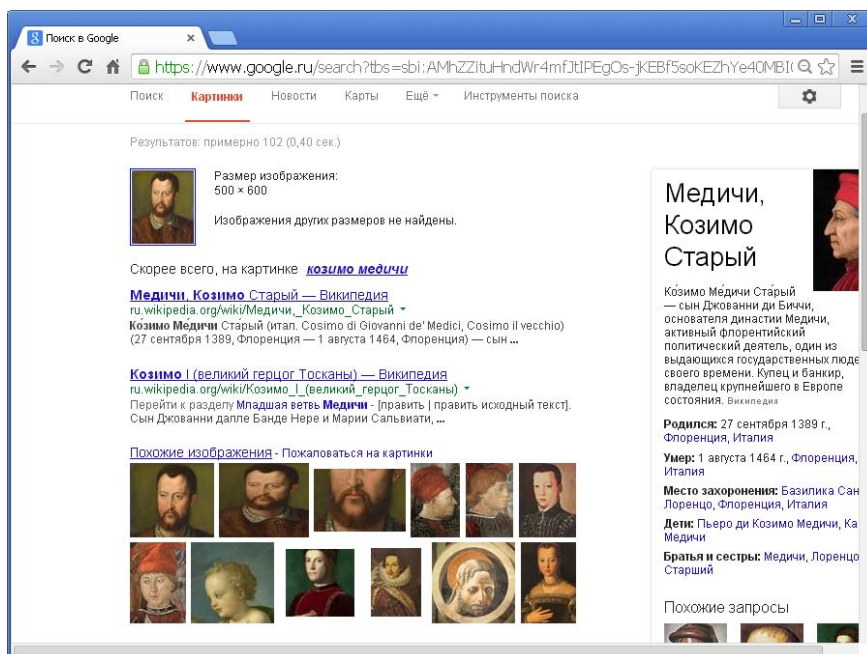


Рис. 5.34. Результат запроса в поисковой системе Google (102 сайта)

Ровно так же решается задача в поисковой системе Baidu (с 2012 г.; здесь нет не только русского, но даже и англоязычного меню, хотя разобраться несложно, и, кроме того, здесь довольно непривычная база данных, ориентированная на китайский Интернет, но иногда это как раз очень хорошо для поиска) и в Яндексе (с 2013 г.).



Рис. 5.35. Результат запроса в поисковой системе Baidu (5 сайтов)

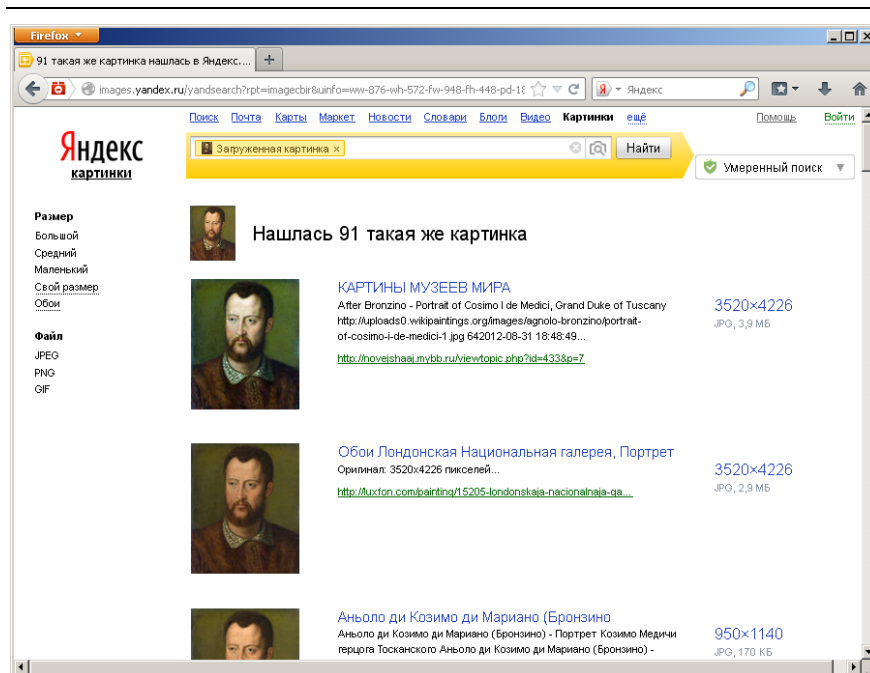


Рис. 5.36. Результат запроса в поисковой системе Яндекс (91 сайт)

Помимо поиска информации об изображении, описанный алгоритм нередко помогает решить и другую задачу: найти ту же картинку, которая у вас есть, но лучшего качества.

12.6. ПОИСК В НОВОСТЯХ И БЛОГАХ

Записи в новостных лентах и блогах занимают существенную часть Интернета. Практически все новостные агентства и средства массовой информации, глобальные, региональные и местные, выдают в Интернет новости, аналитику, интервью, в электронном наборе готовятся и также попадают в Интернет печатные выпуски любой периодики, и почти все это накапливается в Интернете в виде архивов новостей. Эти архивы существуют уже довольно продолжительное время. Приведем для примера список ряда крупных российских новостных сайтов, имеющих сейчас архивы:

- Издательский дом «Коммерсантъ»: архив разбит по изданиям, имеется архив газеты «Коммерсантъ» с первого выпуска 06.10.1992 и журнал «Коммерсантъ» – с 29.01.1990, хотя регулярно – с №11 от 19.03.1990 (<http://www.kommersant.ru/archive>).

- Lenta.ru: имеется архив с 31.08.1999 (удобной формы поиска по архиву нет, но можно менять дату в адресе – <http://lenta.ru/news/1999/08/31>).
- Newsru.com (начинавший как проект канала НТВ): имеется архив с 01.08.2000 с удобным поиском (<http://www.newsru.com/arch>).
- Утро.ру: имеется архив с 26.11.2001 (удобную форму поиска по архиву уже трудно найти, но тоже можно менять дату в адресе – <http://www.utro.ru/2001/11/26/>).
- «Российская газета»: самый старый из выпусков на сайте относится к 01.11.1990, но регулярно они выкладываются с №3263, от 28.07.2003 (<http://rg.ru/gazeta/>).
- РИА Новости: имеется архив с 01.01.2004 (есть архив <http://ria.ru/archive/>, но проще вводить дату в адрес: <http://ria.ru/archive/20040101/>).

Существуют архивы новостей на русском языке и у нескольких крупных зарубежных новостных агентств. Так, у британской корпорации BBC имеется на сайте архив новостей с 03.11.1997 с возможностью выхода на поиск по дате новости¹³. Архив есть и в новостях на русском языке (<http://www.bbc.co.uk/russian/search/>), но там нет возможности поиска по дате. У китайского агентства «Синьхуа» архив новостей на русском языке ведется с 29.06.2009 (<http://russian.news.cn/static/calendar/archives.htm>).

Таким образом, новостные сайты представляют собой гигантское, ежедневно прирастающее хранилище страниц, для которых характерна связь даты появления страницы и описываемого события¹⁴. Примерно то же можно сказать и про блоги: являясь по сути электронными дневниками множества пользователей, многие записи в блогах представляют собой большой интерес с точки зрения реакции на новости и происходившие в момент составления постов события.

¹³ Когда глава проверялась в последний раз (август 2014 г.), этот поиск был очень неудобным. Сначала нужно зайти в поиск <http://www.bbc.co.uk/search/>, затем ввести ключевые слова, и только потом, на странице результатов поиска, можно выбрать категорию "News", после чего появится поиск по диапазону дат.

¹⁴ При этом речь сейчас идет только об архивах веб-страниц, изначально составленных для размещения в Интернете. Отдельную тему составляет перевод в Интернет печатной периодики, который тоже постепенно ведется. Среди наиболее интересных бесплатных ресурсов такого рода стоит назвать цифровой архив газеты «The New York Times» с 1851 г. (<http://www.nytimes.com/content/help/search/archives/archives.html>) и раздел периодики электронной библиотеки «Gallica» Национальной библиотеки Франции (<http://gallica.bnf.fr>), где, например, выставлены все номера газеты «Le Figaro» за 1826–1840 и 1854–1929 гг.

Многие из записей в новостях и блогах находятся через обычный поиск, но полезно знать, что для них существуют отдельные разделы, позволяющие уточнить дату создания документа, местонахождение записи и др.

Наиболее серьезно и понятно эти функции представлены, на наш взгляд, в Яндексе: см. разделы «Расширенный поиск» по адресам <http://news.yandex.ru> и <http://blogs.yandex.ru>. Такой поиск есть и в Google: <https://news.google.ru> и <http://www.google.ru/blogsearch>.

Пример 5.5. В какие страны ездил с визитом министр иностранных дел России во второй половине июня 2005 г.?

Проблема с этим вопросом заключается в том, что хотя на самом официальном сайте МИД РФ сейчас, конечно, есть архив новостей, он начинается только с января 2008 г.:

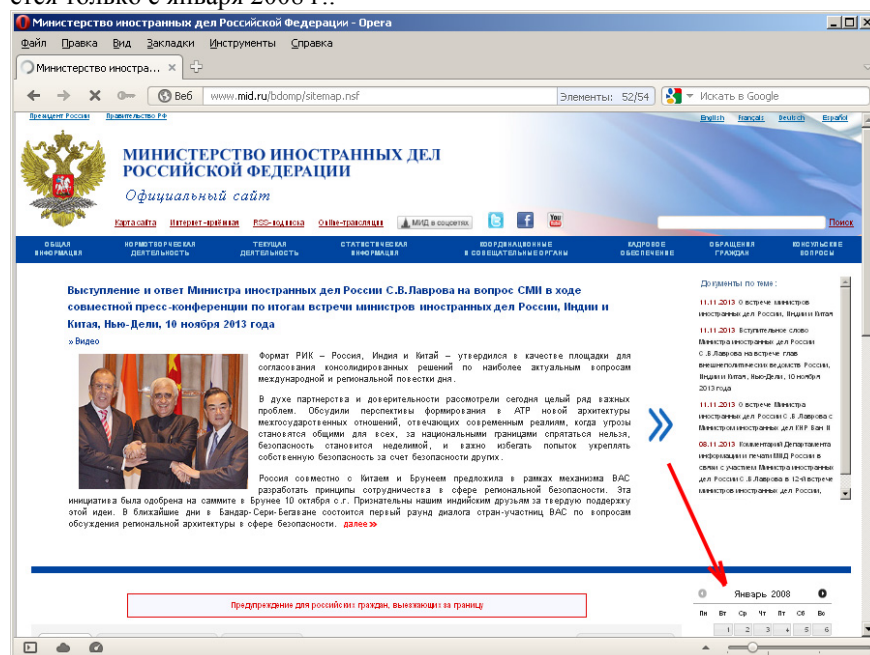


Рис. 5.37. Расположение архива новостей на сайте МИД РФ

Поэтому приходится проводить независимый поиск. Допустим, мы возьмем в качестве ключевых слов *Лавров*, *визит* и *июнь 2005*. Мы видим, что простой поиск, например, в Google, дает слишком большое количество

результатов, причем в первых же из них содержится ненужная нам информация:

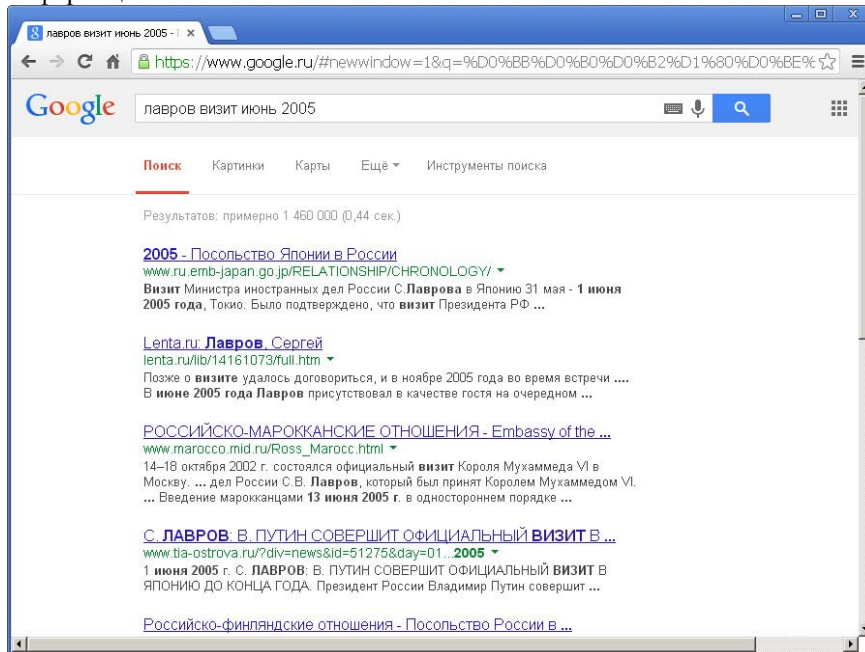


Рис. 5.38. Результат обычного запроса на поиск новости

Однако если мы ограничим поиск только новостями, оставим слова *Лавров визит* и конкретизируем даты интервалом 15–30 июня 2005 г., результат будет совсем иным. Google, правда, теперь не находит подходящих новостей:

Раздел V

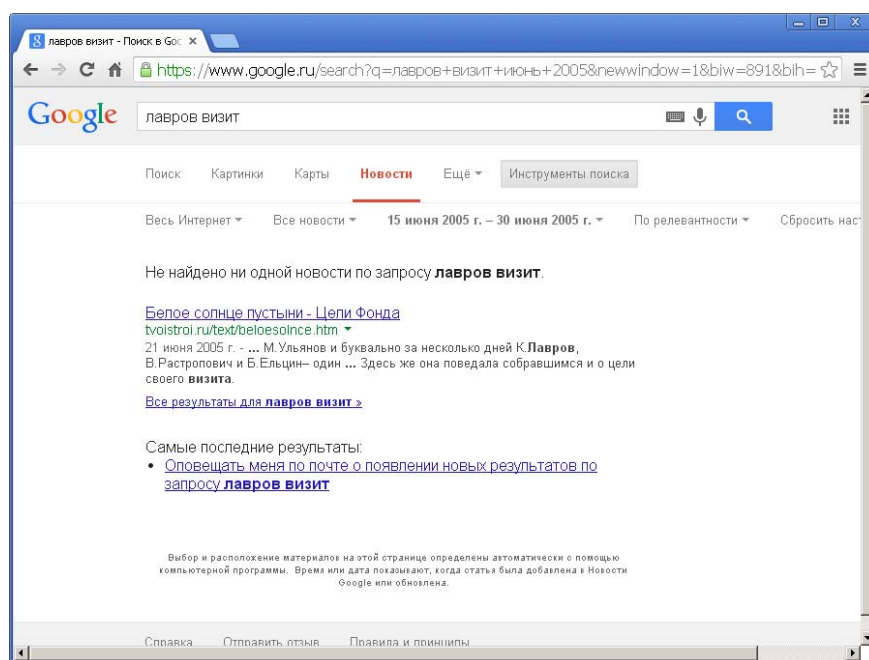


Рис. 5.39. Неудавшийся результат уточненного запроса на поиск новости в системе Google

Зато в Яндексе находятся именно те новости, которые нам нужны:

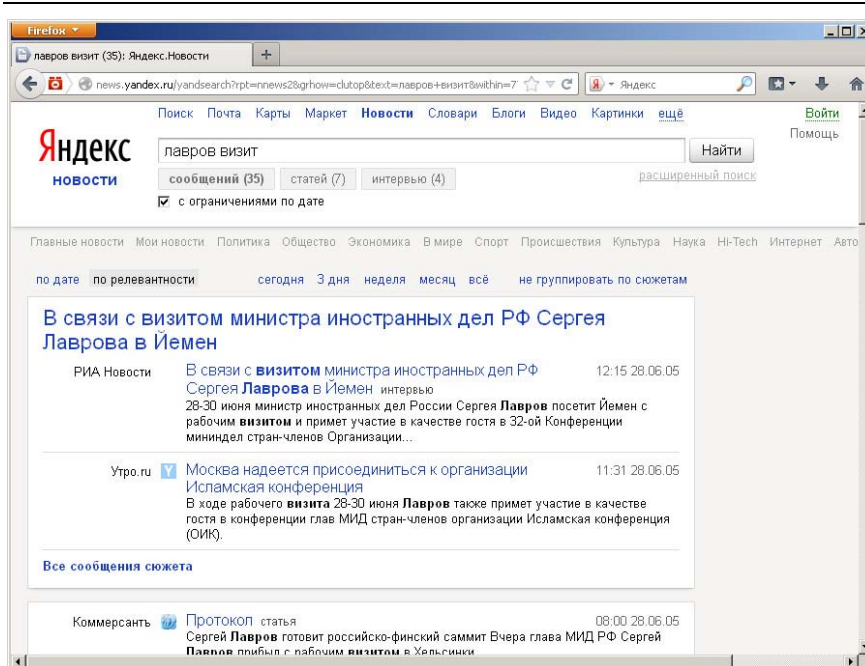


Рис. 5.39. Результат в системе Яндекс.Новости

Здесь найдено всего 35 новостей, их реально просмотреть все, и тогда мы узнаем, что искомыми странами являются Финляндия, Нидерланды и Йемен.

12.7. МЕТАПОИСК¹⁵

Каждая поисковая машина использует собственный алгоритм поиска и индексирования Интернет-ресурсов. В итоге, базы данных (т. н. индексы) поисковых систем далеко не одинаковы, информация, найденная ими в Интернете, совпадает лишь частично. Поэтому возникла идея для удобства поиска объединить результаты выдачи по тому или иному запросу сразу из нескольких систем. Так появились *метапоисковые* системы.

Метапоисковые системы не имеют собственных баз данных. В смысле содержания результатов запросов они полностью зависят от тех поисковых систем, на которые они опираются. Их задача состоит не в обследовании Сети в поисках обновлений, они призваны рационально объединять резуль-

¹⁵ Данный фрагмент раздела написан при участии к.и.н. А.Ю. Володина.

Раздел V

таты поиска: с одной стороны, чтобы результаты не повторялись, с другой стороны – чтобы самые важные результаты оказывались в начале списка выдачи.

Со времени выхода в свет первых метапоисковых систем метапоиск стали обогащать и другими интересными функциями. Все эти функции призваны сделать поиск более быстрым и эффективным. Прежде всего, следует назвать возможность специальной группировки результатов по тематическим кластерам.

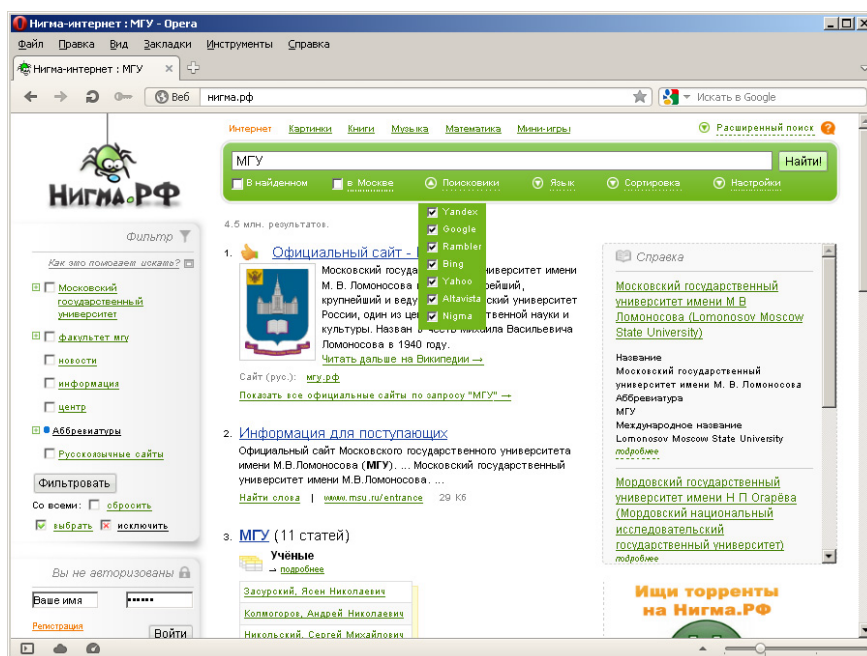


Рис. 5.40. Результаты запроса "МГУ" в метапоисковой системе Нигма

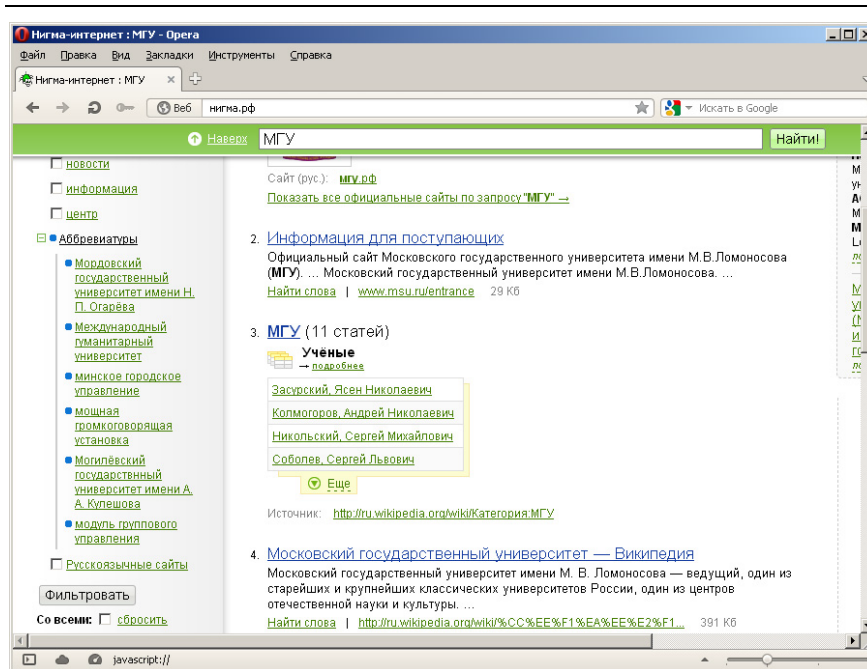


Рис. 5.41. Результаты запроса "МГУ" в метапоисковой системе Нигма

В качестве примеров современного метапоиска с возможностями кластеризации можно привести отечественную разработку "Нигма" (<http://nigma.ru> и <http://нигма.рф> – объединяет поисковые результаты всех рассмотренных нами поисковых систем, – Яндекс, Google, Bing и Mail.ru, – а также имеет свою собственную базу данных) и международную систему "Yippy" (<http://yipru.com> – однако для поиска на русском языке лучше использовать ее еще работающую систему-предшественник от того же разработчика, "Clusty" (<http://clusty.com>).

Например, введем в поле запроса системы "Нигма" запрос *МГУ*. Окно выдачи результата (рис. 5.40) содержит три раздела. Справа дается справка о том, что данное ключевое слово может относиться к разным объектам: здесь это, в частности, Московский государственный университет, но также Мордовский государственный университет, Марийский, Могилевский, Монгольский и др. Слева обозначены кластеры, т. е. группы, в которые автоматически объединяются главные результаты поиска. Например, кластер «Аббревиатуры» приводит представляет не только упомянутые аббревиатуры для МГУ, но и, как видим на рис. 5.41, «мощную громкоговорящую

установку» и «модель группового управления». В центральной части выдачи результатов запроса мы видим наиболее релевантные статьи, как это сделано в выдаче любой поисковой системы.

Подобным же образом действуют системы Yippy и Clusty (рис. 5.42–5.43).

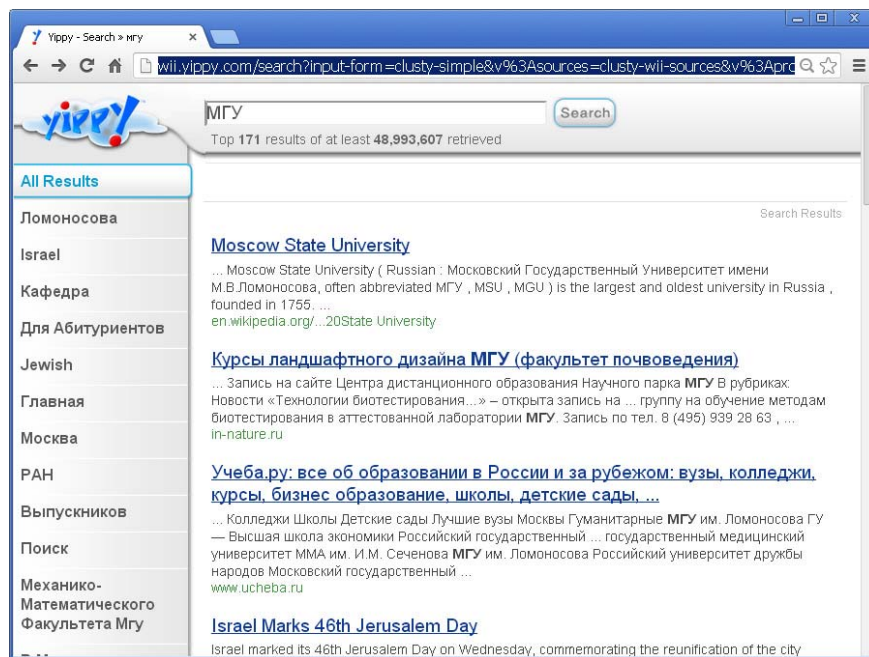


Рис. 5.42. Результаты запроса "МГУ" в метапоисковой системе Yippy

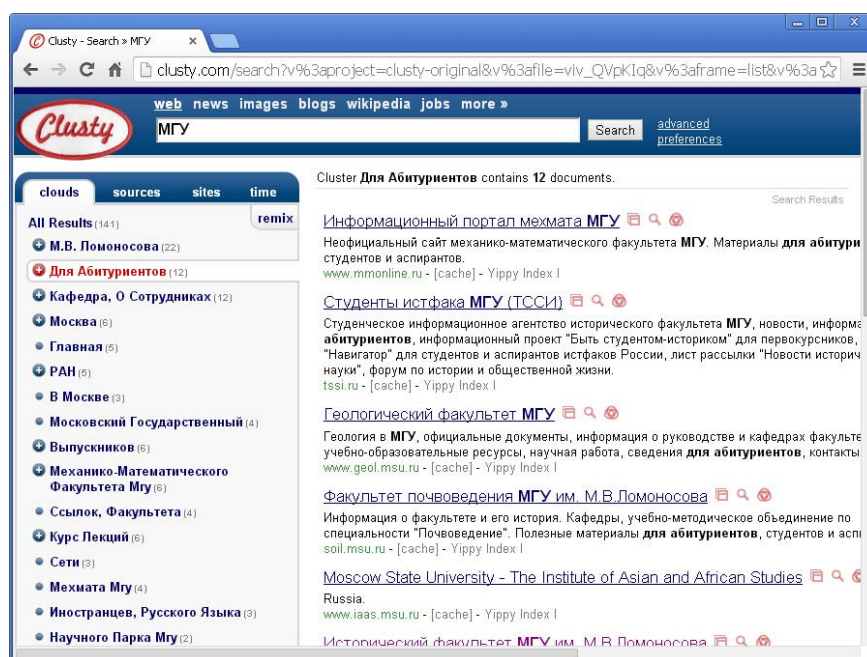


Рис. 5.43. Ссылки, входящие в кластер "Для абитуриентов", автоматически сформированный по запросу "МГУ" в метапоисковой системе Clusty

Среди профессионалов сегодня существуют два противоположных мнения о перспективах использования метапоисковых механизмов. Одни склоняются к тому, что метапоиск – это суррогатное изобретение, которое может давать не отражающие интернет-реальность результаты. Другие считают, что метапоиск – это будущее поисковых технологий и уверены, что одновременный поиск по нескольким системам удобнее того же поиска, но разнесенного по пяти-семи системам. Так или иначе, знать о метапоисковых системах полезно.

ГЛАВА 13. ПРОБЛЕМЫ ВОСПРИЯТИЯ И ОПИСАНИЯ ИНТЕРНЕТ-РЕСУРСОВ ТРЕБОВАНИЯ К КОЛЛОКВИУМУ ДЛЯ СТУДЕНТОВ

13.1. ПРОБЛЕМЫ ЭЛЕКТРОННЫХ ДОКУМЕНТОВ И РАБОТА С НИМИ

Интернет – информационная среда, заполненная текстами, – в том числе и по истории, – самого различного качества. Работа историка вообще в значительной степени связана с тем, чтобы изучать различные источники информации, часто текстовые, и при этом необходимо понимать, что это за тексты: кто их авторы, какова осведомленность авторов, в каких условиях и с какими целями написаны эти тексты и, в конечном итоге, насколько можно доверять изучаемой информации. Все это и многое другое входит в науку источниковедения, принципиально важную для историка. Когда мы изучаем историографию и знакомимся с мнениями других историков, поднимавших интересующую нас проблему, встают те же вопросы отношения к тексту.

Сказанное актуально для любого текста и любого исторического источника, независимо от метода передачи информации. Поэтому при оценке найденной в Интернете информации нужно прежде всего руководствоваться теми же подходами, что и при работе с любым другим, печатным текстом. Однако в дополнение к этому Интернет обладает своей спецификой, поскольку в ряде случаев представляет авторам более широкие возможности работы с их текстами и, соответственно, затрудняет эту пользователям. Перечислим эти особенности Интернета.

Во-первых, некоторые проблемы с оценкой размещенного в Интернете материала может создать *простота копирования*. Допустим, какой-то автор размещает свой текст, подписывает его и делает у себя на сайте вступление, помогающее понять, для какой аудитории и с какими целями написан этот текст. С таким материалом удобно работать. Но в Интернете мало что может помешать разработчикам других сайтов скопировать понравившийся им текст (или его фрагмент, в том числе, возможно, вырванный из контекста) и разместить у себя. При этом порой, – в случае недобросовестного копирования, но это не так уж и редко встречается, – теряются все указания на то, с какого исходного сайта взят текст, для чего он написан, кто его автор, есть ли у текста другие фрагменты, ссылки на источники и т.п. Поэтому нужно понимать, что если в Интернете найден текст, где отсутствует, прежде всего, указание на его авторство, то, скорее всего, мы имеем дело именно с таким недобросовестным копированием. При этом с большой долей уверенности можно полагать, что исходный текст, в кото-

ром содержится указание на автора и, возможно, прочая полезная информация, тоже существует где-то в Интернете, и имеет смысл сначала попробовать найти его. Для этого достаточно поставить в запрос поисковой системы сразу несколько слов, – хотя бы все первое предложение, – в кавычках, и тот же текст, если только он не был удален из Интернета, будет найден по другим адресам.

У простоты копирования информации в Интернете, впрочем, есть и положительная сторона. Если исходный ресурс пропадает (а такое может случиться, если сайтом перестают заниматься), то иногда тот же текст можно найти там, куда он был в свое время скопирован. Если копирование было добросовестным, то вся сопутствующая дополнительная информация также будет скопирована.

Во-вторых, проблему для историка составляет *возможность изменения материала*, размещенного в Интернете. В отличие от печатного текста, в Интернете в подавляющем большинстве случаев тот, кто имеет доступ к сайту, может в любой момент подкорректировать, заменить или вовсе убрать любой материал. В частности, особые неудобства возникают при работе с блогами как историческим источником: в ряде случаев блоги могли бы быть полезны для изучения как инструмент публичного высказывания их авторов, но имеет смысл еще дополнительно найти подтверждения тому, что точка зрения, высказанная несколько лет назад и имеющая соответствующую датировку, не изменена автором впоследствии. С той же осторожностью стоит подходить к любым ресурсам, находящимся под полным контролем автора, например, сайтам политических партий и общественных организаций, однако в этом случае работа упрощается тем, что политические партии, помимо интернет-архивов, оставляют массу других, надежно датированных материалов, в то время как с блогами такого обычно не происходит.

При этом, правда, существуют онлайн-архивы, которые в избирательном порядке сохраняют у себя копии веб-сайтов (и иногда блогов) на какие-то определенные моменты времени. Не углубляясь специально в этот вопрос, скажем, что наиболее известным из таких проектов является "Internet Archive" (<http://archive.org>). Поскольку к изменению этого ресурса авторы сайтов не имеют доступа, хранящиеся таким образом «слепки» сайтов можно считать достаточно надежным источником информации об их состоянии в прошлом.

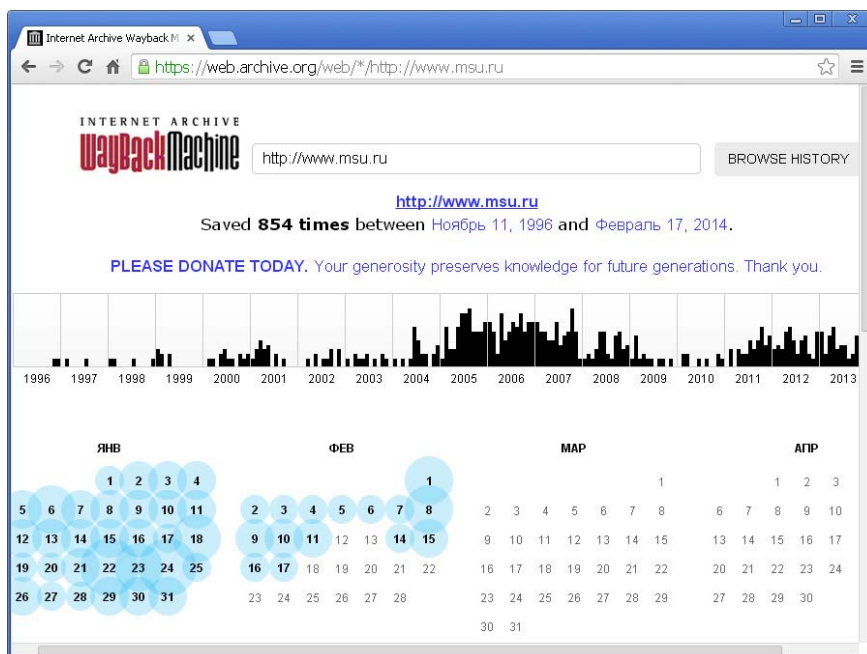


Рис. 5.43. Временная линия сохранений главного сайта МГУ в системе Archive.org

Положительным примером организации системы исправлений на сайте, изменять содержание которого имеют право многие пользователи, является электронная энциклопедия «Википедия» (<http://wikipedia.org>). К работе над нею привлекаются все заинтересованные пользователи, и одна и та же статья, особенно по предмету, вызывающему споры в обществе, может многократно изменяться, причем возможно и такое, что новые исправления ухудшают статью. Понимая это, организаторы проекта заложили функцию сохранения всех исправлений страницы: можно посмотреть, кто и когда вносил в статью какую правку или на странице «Обсуждение» подвергал сомнению ту или иную информацию в статье.

Правда, здесь мы сталкиваемся с еще одним проблемным свойством информации в Интернете, – нередкой **анонимностью авторов**. И в Википедии, и на множестве форумов, в блогах, вообще в статьях авторы нередко выступают под обезличенными псевдонимами («никами») или вовсе не указываются. Соответственно, почти невозможной становится оценка компетенции автора и затруднительным, иногда также невозможным, – сопос-

тавление высказываний одного и того же автора, сделанных на разных ресурсах.

Наконец, в совокупности с простотой копирования легкость изменения информации приводит еще и к широкому распространению разных форм плагиата как в самом Интернете, так и при традиционных формах передачи информации, заимствующих, тем не менее, информацию из найденного в Интернете. На рис. 5.44 приводится пример сайта (как на нем написано, «авторский сайт» некоего писателя), на котором, безо всяких ссылок, с собственным заголовком и с произвольными сокращениями приводится текст главы «Вавилонский плен» из книги «Иван Калита» профессора Н.С.Борисова.

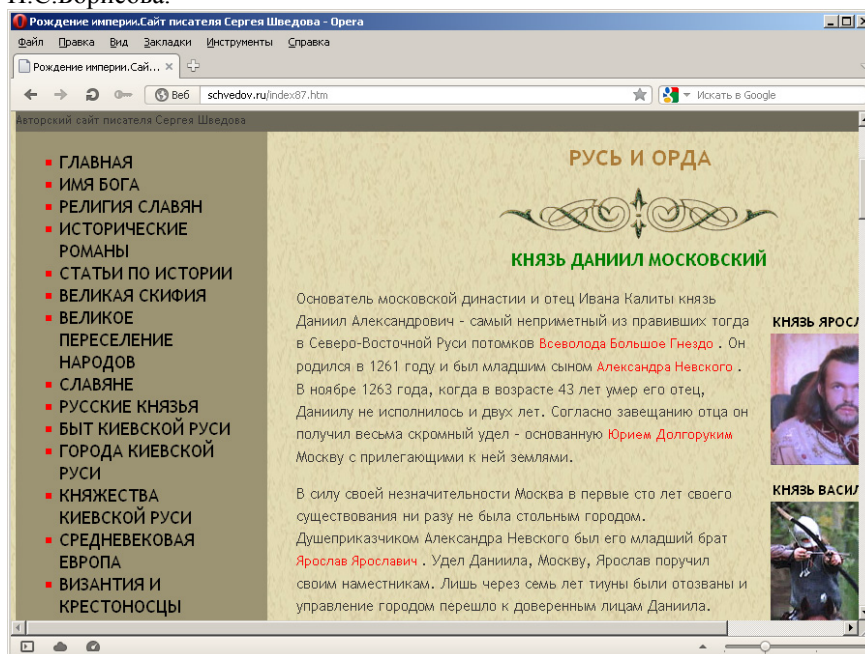


Рис. 5.44. Пример сайта с недобросовестным размещением чужого текста

Учитывая все сказанное, при использовании информации, найденной в Интернете, следует с особой тщательностью подходить к вопросу об ее исходном источнике.

В этой оценке можно руководствоваться некоторой схемой важных вопросов, а их следует обращать частично к самой анализируемой странице, а

частично – к тому электронному ресурсу, на котором она размещена. Назовем эти вопросы.

Вопросы к самому информационному материалу:

1) Является ли электронная публикация уникальной? Можно ли найти тот же материал на других сайтах?

2) Если существуют дубликаты публикации, можно ли предположить, какое из мест размещения электронной публикации является исходным? Если это возможно, то дальнейшие вопросы следует адресовать именно к исходному электронному ресурсу. Вместе с тем, сам ответ на этот вопрос редко бывает очевидным, и возможно, что сначала нужно рассмотреть другие вопросы, которые мы укажем ниже, – сравнить качество нескольких мест размещения электронной публикации и только потом обратить все внимание на вероятный оригинал.

3) Является ли электронная публикация копией какой-либо печатной статьи, опубликованного или оригинального текста исторического источника? Если да, то в каком формате выполнена электронная публикация? Воспроизводит ли она исходный текст визуально (то есть передает копию в графическом или псевдографическом формате) или является текстовой?

4) Если электронная публикация представлена в текстовом формате, заметны ли искажения источника? Нет ли заметных ошибок или лакунов в тексте? Вопрос о наличии лакунов, конечно, актуален и для графического представления источника. Снабжена ли публикация научно-справочным аппаратом? Имеется ли нумерация страниц (абзацев, строк, глав)?

5) Кто автор электронной публикации, когда она написана, с какой целью? Те же вопросы следует отнести к бумажной публикации, с которой сделана электронная, и к оригинальному источнику. Где хранится источник-оригинал? Что можно сказать о компетентности и информированности автора и публикаторов? Для ответа на последний вопрос, возможно, имеет смысл провести дополнительный поиск.

6) Если электронная публикация является переводом некоторого источника с иностранного или древнего языка, кем выполнен перевод? Какой принцип принят для представления текстов, в оригинале содержащих символы, отличные от стандартной компьютерной кодировки (например, ныне отсутствующие в алфавите древнерусские буквы)?

Вопросы к электронному ресурсу, на котором размещен документ:

1) Как называется электронный ресурс (сайт, веб-проект и т.п.)? Кто его авторы? Когда он создан? Что можно сказать о целевой аудитории ресурса?

2) Является ли электронный ресурс востребованным? Ссылается ли кто-то на него?¹⁶ Сколько таких ссылок? Какая тематика у сайтов, ссылающихся на рассматриваемый нами электронный ресурс?

По итогам ответа на все эти вопросы должно приниматься и общее решение о научной значимости электронной публикации. Оно состоит из двух частей: 1) отношение к самому материалу и возможности его использования в качестве источника для исторического исследования или достойного историографического материала и 2) возможность полноценного использования электронного материала в качестве замены соответствующей публикации. Последнее обстоятельство тоже важно: Интернет выполняет много функций, в том числе предоставляет удаленный доступ к материалам библиотек или архивов, и поскольку он является глобальным, через Интернет мы можем получить такие материалы, до которых не можем добраться иным способом. Это очень полезно, только нужно провести описанную выше процедуру анализа электронного документа и убедиться, что электронный доступ действительно заменяет нам поход в библиотеку.

13.2. ОТЧЕТНАЯ РАБОТА:

КОЛЛОКВИУМ ПО ПОИСКУ И ОЦЕНКЕ ИНФОРМАЦИИ В ИНТЕРНЕТЕ

В качестве отчетной работы по курсу «Информатика и математика» на историческом факультете МГУ проводится коллоквиум, состоящий из трех частей: поиск информации, поиск карточек в электронных каталогах библиотек и оценка предложенного электронного документа по некоторому ряду критериев. Поскольку коллоквиум проводится очно и время его выполнения ограничено, при оценке документа приходится отказываться от анализа некоторых важных вопросов из числа сформулированных выше, хотя в реальной научной работе их значение нельзя преуменьшать.

В приложении 5.1 приводится пример варианта коллоквиума.

¹⁶ Для ответа на данный вопрос можно использовать специальный элемент запроса "link" в поисковой системе Google (в настоящее время аналога в других рассматриваемых системах ему нет): по запросу *link:www.hist.msu.ru* будут найдены все проиндексированные поисковиком страницы, ссылающиеся на главную страницу сайта исторического факультета МГУ. Этот запрос не распространяется на весь сайт: если, скажем, нас интересуют ссылки не сайт исторического факультета, а на его электронную библиотеку, то этот адрес и нужно указывать в запросе: *link:www.hist.msu.ru/ER/*, а объединить результаты двух этих запросов через обращения к поисковой системе нельзя.

Приложение 5.1.

Пример задания на коллоквиуме по работе с Интернетом.

Коллоквиум III (Интернет)

Ответы на все вопросы следует сохранить в doc-файле. Его основой является этот самый текст с заданием, но файл переименовывается так, чтобы содержать фамилию студента: *Иванов_колл3.doc*

Также запишите свои имя, фамилию и группу: _____

Часть I. Проведите поиск информации в Интернете по предлагаемым запросам. Полный ответ на каждый вопрос должен состоять из двух частей: собственно содержательного ответа и адреса страницы (нескольких страниц, если нужно), на которой содержится этот содержательный ответ.

Задание 1 (1 балл).

Где находится библиотека Вольтера?

Ответ:

Адрес(а):

Задание 2 (2 балла).

Перечислите всех прадедов Петра I.

Ответ:

Адрес(а):

Задание 3 (2 балла).

О чем говорит статья 112 «сталинской» конституции?

Ответ:

Адрес(а):

Задание 4 (2 балла).

Кто автор текста, представленного здесь:

<http://www.ellada.spb.ru/?p=3010117>

Ответ:

Адрес(а):

Задание 5 (2 балла).

Первый съезд работников какой профессии прошел в Москве в 1870-х гг.?

Ответ:

Адрес(а):

Задание 6 (1 балл).

Найдите годы жизни персонажа, представленного на портрете:

<http://www.hist.msu.ru/Temp/who.jpg>

Ответ:

Адрес(а):

Часть II. Представьте, что вы пишете доклад на тему «Фабричная инспекция в России в XIX в.» Проведите поиск в электронных каталогах библиотек. Найдите 4 каталожные карточки книг (статей), полезных для работы над этим докладом, в 4-х разных библиотечных системах. Скопируйте каталожные карточки на сайте и вставьте их сюда. Ответом является либо копия текста карточки, перенесенная с помощью копирования через буфер, либо изображение соответствующей карточки (если выбранный вами каталог представлен набором сканированных изображений каталожных карточек). Если по названию книги/статьи или по ключевым словам в карточке нельзя понять, чем карточка может быть полезна для доклада, следует дать ваш краткий комментарий (2 балла за одну карточку).

Карточка 1:

Библиотека, откуда взята карточка 1:

Карточка 2:

Библиотека, откуда взята карточка 2:

Карточка 3:

Библиотека, откуда взята карточка 3:

Карточка 4:

Библиотека, откуда взята карточка 4:

Часть III. Оцените электронный документ <http://rusarchives.ru/smuta/01-001-venchanie-na-carstvo-fedora-ivanovicha.shtml> по прилагаемому протоколу.

**ПРОТОКОЛ ОЦЕНКИ ЭЛЕКТРОННОГО ДОКУМЕНТА
I. АТТРИБУЦИЯ ЭЛЕКТРОННОГО ДОКУМЕНТА (2 балла)**

	Оригинал (заполняется только при описании источников)	Бумажная публикация	Электронная публикация
Автор или создатель			
Название			
Дополнительная информация	<i>Датировка, место хранения:</i>	<i>Выходные данные:</i>	<i>Формат публикации:</i>

Раздел V

Является ли публикация уникальной? Есть ли дубликаты или другие варианты публикации?

II. ХАРАКТЕРИСТИКА ЭЛЕКТРОННОГО РЕСУРСА (ЭР) (2 балла)

Адрес ЭР	
Название ЭР	
Авторы ЭР	
Тематика ЭР	
Сколько ссылок на данный ресурс? Какая тематика у этих сайтов? Приведите примеры	

III. КРИТИЧЕСКИЙ КОММЕНТАРИЙ ЭЛЕКТРОННОГО ДОКУМЕНТА (2 балла)

Можно ли пользоваться данным документом для научных целей? Есть ли научно-справочный аппарат? Есть ли нумерация страниц (абзацев, строк, глав)? Есть ли в публикации заметные ошибки, лакуны? Есть ли возможности для удобной навигации? Имеет ли электронная публикация самостоятельное значение?

--

IV. БИБЛИОГРАФИЧЕСКАЯ ССЫЛКА (1 балл)

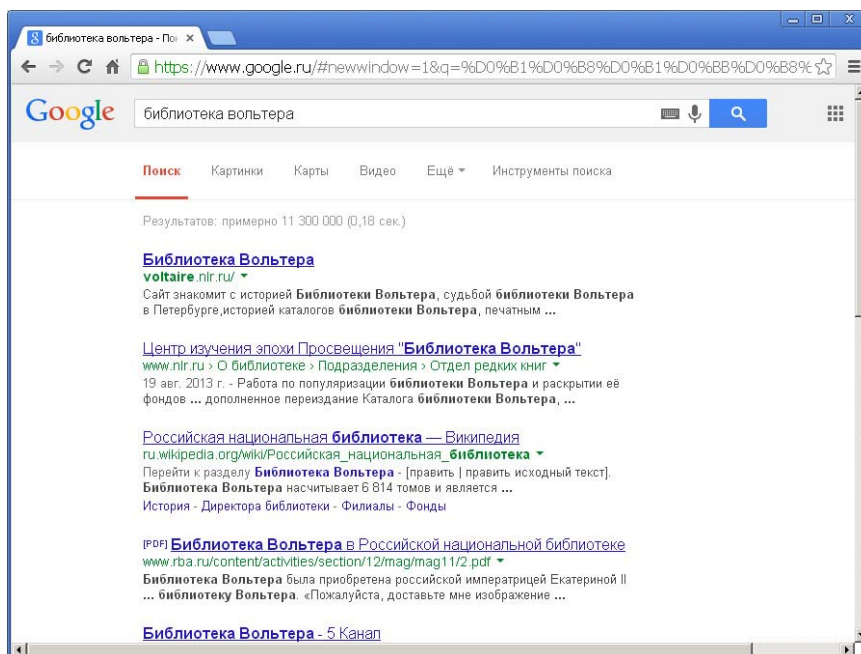
Оформите библиографическую ссылку на этот электронный документ по ГОСТу Р 7.0.5-2008:

Рассмотрим этот вариант.

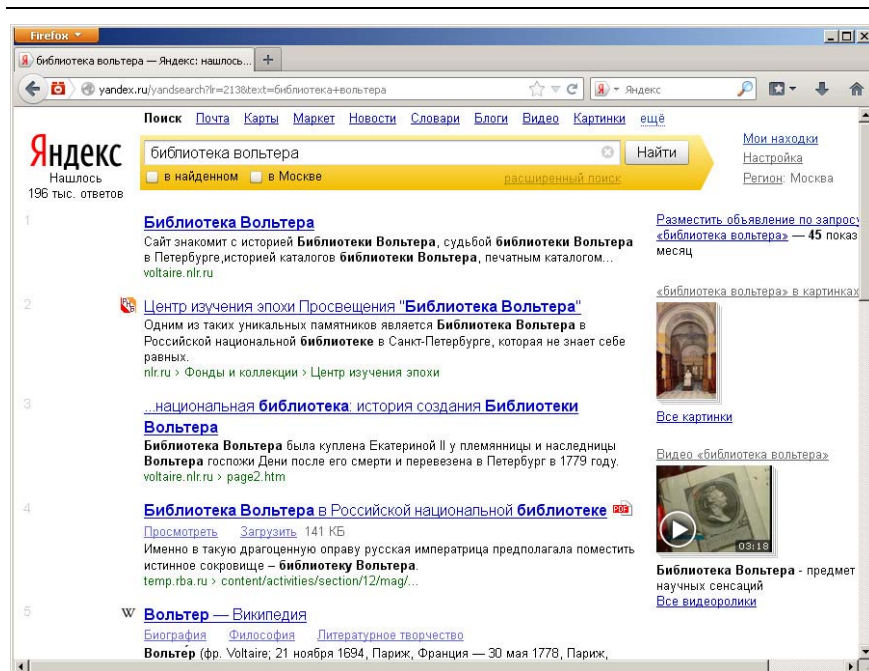
Часть 1. Поисковые запросы.

В каждом варианте коллоквиума содержится 6 вопросов. Правильный ответ на вопрос предполагает не только собственно содержательный ответ, но и указание адреса страницы, на которой он найден. Дело в том, что задание коллоквиума дается не на общую эрудицию, а проверяет способности по поиску в Интернете, поэтому, с одной стороны, студенту не должно сильно помочь прямое знание ответа на поставленный вопрос, а с другой – оценка не должна быть снижена, если в Интернете вдруг найдется неправильный ответ: в условиях ограниченного времени нет возможности искать подтверждения правильности найденной информации. В научной работе, конечно, критическое отношение к найденным в Интернете ответам очень важно.

Вопрос 1 требует провести простой поиск, он составлен так, что ответ будет быстро найден в любой поисковой системе, вероятно, даже без использования такого элементарного инструмента, как кавычки. На приведенный в варианте запрос ответ находится легко:



Раздел V



Библиотека Вольтера была куплена Екатериной II после его смерти и сейчас составляет отдельный фонд в собрании Российской национальной библиотеки в Санкт-Петербурге. В качестве второй части ответа можно поставить любой из представленных на рисунках адресов, но самым авторитетным по данному вопросу, очевидно, является веб-страница непосредственно этого фонда: <http://voltaire.nlr.ru>.

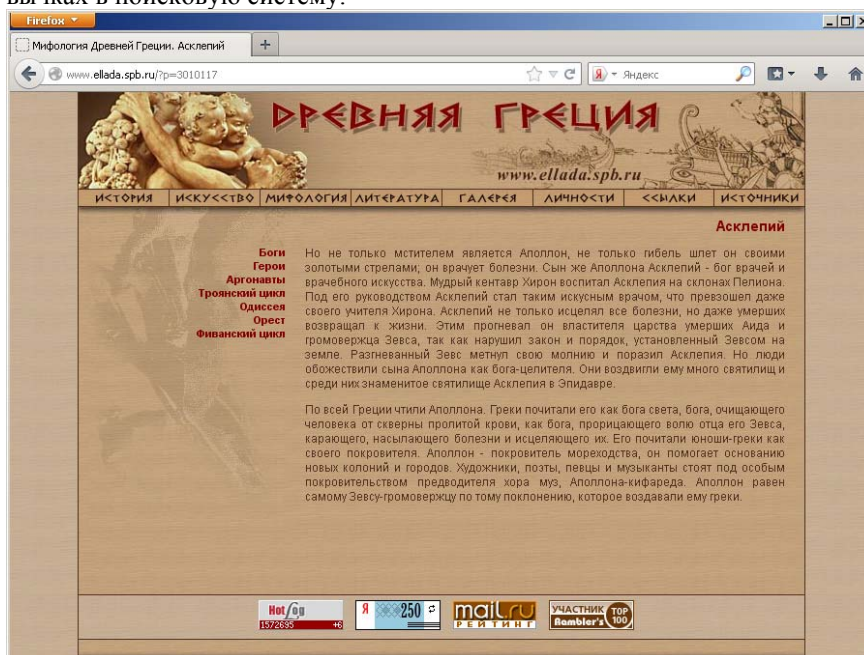
Второй вопрос ориентирован тоже на простой поиск, но этот поиск требует просмотра нескольких страниц. Так, искать всех прадедов Петра I проще всего не вместе, а через выяснение биографий каждого из его предков в первых двух поколениях. Впрочем, при особенном желании можно найти и страницы, на которых вся нужная информация находится сразу, например, здесь: http://russia-today.narod.ru/past/gen/rom_petrl.html:

Федор Никитич Романов (патриарх Филарет);
Лукьян Степанович Стрешнев;
Полуэкт Иванович Нарышкин;
Леонтий Дмитриевич Леонтьев.

Третий вопрос требует провести поиск в конкретном тексте, который и нужно сначала найти (см. раздел 12.2, пример 5.2). Здесь это Конституция

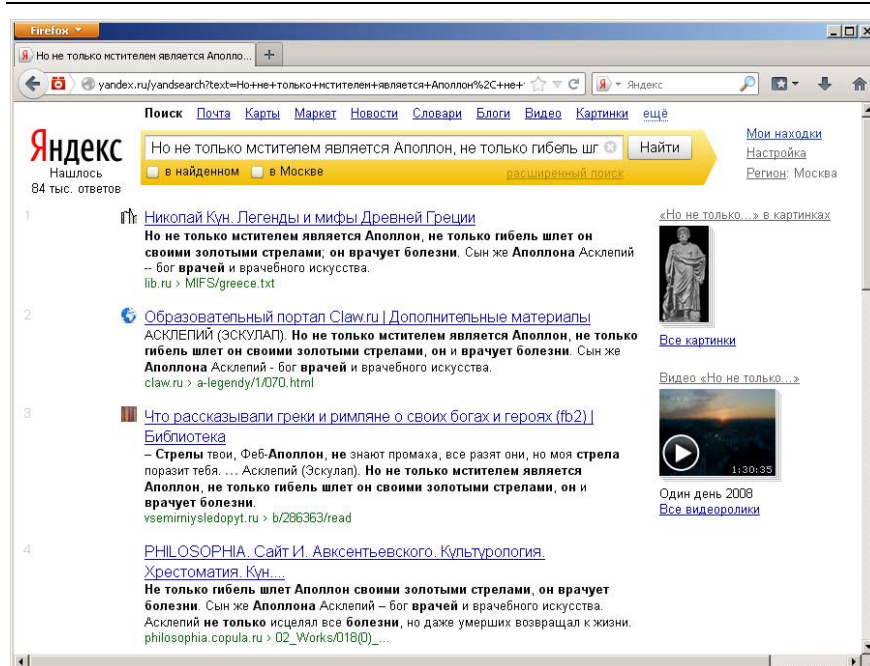
СССР 1936 г., 112-я статья которой гласит: «Судьи независимы и подчиняются только закону» (<http://www.hist.msu.ru/ER/Etext/cnst1936.htm>).

Ответом на **вопрос 4** будет автор текста на странице, полученной недобросовестным копированием. Как уже говорилось, самый простой способ поиска оригинала – ввести целиком текст первого предложения в кавычках в поисковую систему.



Уже первая ссылка прямо выводит на полный текст популярной книги Н.А.Куна «Легенды и мифы Древней Греции»:

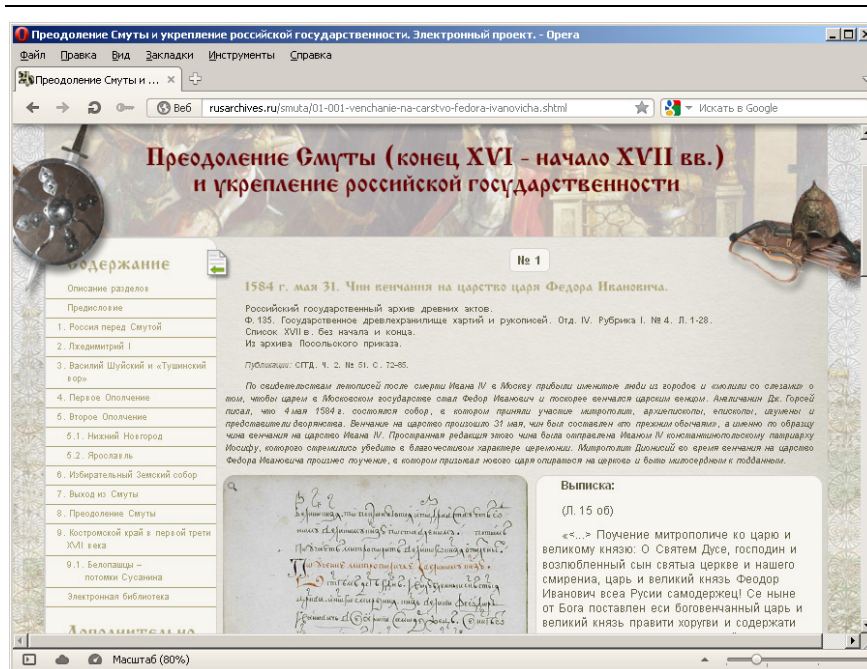
Раздел V



В пятом вопросе лучше использовать поиск по диапазону чисел (см. раздел 12.3, пример 5.3). Ответ на приведенный здесь вопрос – съезд юристов 1875 г. (http://ru.wikipedia.org/wiki/Первый_съезд_русских_юристов).

Шестой вопрос – поиск по изображению (см. раздел 12.5, пример 5.4). По приведенному адресу находится портрет генерал-фельдмаршала И.В. Гурко (1828–1901).

В части оценки документа приводятся выдержки из Чина венчания на царство царя Федора Иоанновича, выставленные в рамках проекта Росархива «Преодоление Смуты...»



Рассмотрим, как примерно должны выглядеть ответы на поставленные в коллоквиуме вопросы.

Электронный документ представляет собой публикацию исторического источника XVI века, дошедшего в некоторой рукописи и опубликованного ранее. Поэтому в описании его авторства имеет смысл рассмотреть все три элемента.

Автор собственно текста – митрополит Московский и всея Руси Дионисий, произнесший данную речь. Текст записан, естественно, каким-то неизвестным чиновником; более того, передан (как указано и в комментарии на той же веб-странице) в списке (то есть копии) XVII века, изображение которого и приводится. Эта копия, как написано здесь же, хранится в Российском государственном архиве древних актов (РГАДА), приведены архивные реквизиты дела. Эту информацию следует считать абсолютно достоверной, учитывая, что она выставлена на проекте Росархива, аккуратно публикующем архивные документы. То, что исходный источник представляет собой не оригинал записи речи митрополита, а позднюю копию, может не иметь значения, а может иметь. Во всяком случае, на это обстоя-

тельство нужно обратить внимание, если привлекать его для полноценного исследования.

Источник издан, на сайте приводятся реквизиты публикации: СГГД, Ч.2. №51. С.72–85. Что это значит, необходимо посмотреть отдельно: СГГД – это «Собрание государственных грамот и договоров, хранящихся в государственной коллегии иностранных дел». Часть 2 издана в Москве в 1819 г.

Электронная публикация является частью проекта «Преодоление Смуты...», и ее авторами можно считать участников проекта, указанных на отдельной странице (<http://rusarchives.ru/smuta/organizationaly-i-uchastniki.shtml>). Текст выставлен параллельно в двух форматах: html-текст и графические изображения источника. Источник, естественно, написан скорописью XVII в., а текст приводится на языке XVII в., но в передаче современным алфавитом, с заменой отмененных сейчас букв их аналогами и раскрытием титл. По тексту видно, что в нем имеются лакуны; их можно заметить и по сопоставлению текста с изображением источника (где лакун нет), причем по изображению вполне можно увидеть, какие фрагменты исключены, но надо заметить, что и на страницах, представленных изображениями, дан не весь текст речи митрополита. Изображение достаточно качественное для чтения текста, предусмотрена функция увеличения масштаба рисунка.

Текст продублирован в нескольких блогах и на форуме, например: <http://rublev-museum.livejournal.com/388863.html> и <http://karpets.livejournal.com/975019.html>.

Очевидно, именно рассматриваемый текст является оригиналом для этих дубликатов.

Электронный ресурс называется полностью «Преодоление Смуты (конец XVI – начало XVII вв.) и укрепление российской государственности» и носит характер тематического ресурса по истории Смутного времени, сделанного на основе публикации исторических архивных документов. Если считать автоссылки с других страниц сайта Росархива, на проект имеется более 40 ссылок. Они размещены на сайтах архивов (в виде баннеров) <http://www.yar-archives.ru> и <http://arhiv.adm44.ru> и в каталоге образовательных ресурсов <http://window.edu.ru>.

Если не считать ссылку на архивный фонд и на листы архивного дела, научно-справочный аппарат отсутствует. Навигации практически нет, но в данном случае она и не требуется, потому что весь предложенный текст умещается на одном экране.

Тематический ресурс имеет образовательное и научно-популярное значение, но этот электронный документ вряд ли может быть полезен для научной работы. прежде всего потому, что текст дан в отрывках, что сильно уменьшает его значение, по крайней мере в сравнении с бумажной публи-

кацией СГГД (электронная копия которой, между прочим, тоже доступна в Интернете, в библиотеке «Руниверс»: <http://www.runivers.ru/lib/book6820/174865/>). Таким образом, можно сказать, что данная электронная публикация не имеет самостоятельного научного значения.

В задании также требуется оформить ссылку на электронный документ в соответствии с требованиями библиографического стандарта (см. Приложение 5.2). Для данного документа ссылка будет такой:

1584 г. мая 31. Чин венчания на царство царя Федора Ивановича // [Проект «Преодоление Смуты (конец XVI – начало XVII вв.) и укрепление российской государственности»] / [М., 2012]. URL: <http://rusarchives.ru/smuta/01-001-venchanie-na-carstvo-fedora-ivanovicha.shtml> (дата обращения: 01.01.2014)

Приложение 5.2.

Выдержки из ГОСТа Р 7.0.5-2008, имеющие отношение к ссылкам на электронные ресурсы.

Особенности составления библиографических ссылок на электронные ресурсы

10.1 Объектами составления библиографической ссылки также являются электронные ресурсы локального и удаленного доступа. Ссылки составляют как на электронные ресурсы в целом (электронные документы, базы данных, порталы, сайты, веб-страницы, форумы и т. д.), так и на составные части электронных ресурсов (разделы и части электронных документов, порталов, сайтов, веб-страниц, публикации в электронных сериальных изданиях, сообщения на форумах и т. п.).

(Российские правила каталогизации. Ч. 1. Основные положения и правила [Электронный ресурс]/Рос. библ. ассоц., Межрегион. ком. по каталогизации. — М., 2004. — 1 CD-ROM. — Загл. с этикетки диска)

*Галина Васильевна Старовойтова, 17.05.46 — 20.11.1998: [мемор. сайт] / сост. и ред. Т. Лиханова. [СПб., 2004]. URL: <http://www.starovoitova.ru/rus/main.php> (дата обращения: 22.01.2007).

¹⁰ Справочники по полупроводниковым приборам // [Персональная страница В. Р. Козака] / Ин-т ядер, физики. [Новосибирск, 2003]. URL: <http://www.inp.nsk.su/%7Ekozak/start.htm> (дата обращения: 13.03.2006).
Ильянц Г. Создание телевидения // QRZ.RU: сервер радиолюбителей России. 2004. URL: <http://www.qrz.ru/articles/article260.html> (дата обращения: 21.02.2006).

176. Паринов С. И., Ляпунов В. М., Пузырев Р. Л. Система Соционет как платформа для разработки научных информационных ресурсов и онлайн-сервисов // Электрон, б-ки. 2003. Т. 6, вып. 1.

URL: <http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2003/part1/PLP/> (дата обращения: 25.11.2006).

10.2. Ссылки на электронные ресурсы составляют по правилам, изложенным в разделах 4-9, с учетом следующих особенностей.

10.3. Если ссылки на электронные ресурсы включают в массив ссылок, содержащий сведения о документах различных видов, то в ссылках, как

правило, указывают общее обозначение материала для электронных ресурсов.

² Жизнь прекрасна, жизнь трагична... [Электронный ресурс]: 1917 год в письмах А. В. Луначарского, А. А. Луначарской / отв. сост. Л. Роговая; сост. Н. Антонова; Ин-т «Открытое о-во». М., 2001. URL: <http://www.auditorium.ru/books/473/> (дата обращения: 17.04.2006).

65. Авилова Л. И. Развитие металлопроизводства в эпоху раннего металла (энеолит — поздний бронзовый век) [Электронный ресурс]: состояние проблемы и перспективы исследований // Вестн. РФФИ. 1997. № 2. URL: <http://www.rfbr.ru/pics/22394ref/file.pdf> (дата обращения: 19.09.2007).

10.4. В примечании приводят сведения, необходимые для поиска и характеристики технических спецификаций электронного ресурса. Сведения приводят в следующей последовательности: системные требования, сведения об ограничении доступности, дату обновления документа или его части, электронный адрес, дату обращения к документу.

10.4.1. Сведения о системных требованиях приводят в тех случаях, когда для доступа к документу требуется специальное программное обеспечение (например, Adobe Acrobat Reader, PowerPoint и т. п.).

⁸ Беглик А. Г. Обзор основных проектов зарубежных справочных служб: програм. обеспечение и тех-нол. подходы // Использование Интернет-технологий в справочном обслуживании удаленных пользователей: материалы семинара-тренинга, 23-24 нояб. 2004 г. / Рос. нац. б-ка, Виртуал. справ. служба. СПб., 2004. Систем, требования: PowerPoint. URL: <http://vss.nlr.ru/about/seminar.php> (дата обращения: 13.03.2006).

53. Волков В. Ю., Волкова Л. М. Физическая культура: курс дистанц. обучения по ГСЭ 05 «Физ. Культура» / С.-Петерб. гос. политехи, ун-т, Межвуз. центр по физ. культуре. СПб., 2003. Доступ из локальной сети Фундамент, б-ки СПбГПУ. Систем, требования: Power Point. URL: <http://www.unilib.neva.ru/dl/local/407/oe/oe.ppt> (дата обращения: 01.11.2003).

10.4.2. Примечание об ограничении доступности приводят в ссылках на документы из локальных сетей, а также из полнотекстовых баз данных, доступ к которым осуществляется на договорной основе или по подписке

(например, «Кодекс», «Гарант», «КонсультантПлюс», «EBSCO», «ProQuest», «Интегрум» и т. п.).

⁵ О введении надбавок за сложность, напряженность и высокое качество работы [Электронный ресурс]: указание М-ва соц. защиты Рос. Федерации от 14 июля 1992 г. № 1-49-У. Документ опубликован не был. Доступ из справ.-правовой системы «КонсультантПлюс».

10.4.3. При наличии сведений о дате последнего обновления или пересмотра сетевого документа, их указывают в ссылке, предваряя соответствующими словами «Дата обновления» («Дата пересмотра» и т. п.). Дата включает в себя день, месяц и год.

114. Экономический рост//Новая Россия: [библиогр. указ.]/сост.: Б. Берхина, О. Коковкина, С. Канн; Отд-ние ГПНТБ СО РАН. Новосибирск, [2003 —]. Дата обновления: 6.03.2007. URL: <http://www.prometeus.nsc.ru/biblio/newrus/egrowth.ssi> (дата обращения: 22.03.2007).

Если по экранной титульной странице электронного ресурса удаленного доступа (сетевого ресурса) невозможно установить дату публикации или создания, то следует указывать самые ранние и самые поздние даты создания ресурса, которые удалось выявить.

10.4.4. Для электронных ресурсов удаленного доступа приводят примечание о режиме доступа, в котором допускается вместо слов «Режим доступа» (или их эквивалента на другом языке) использовать для обозначения электронного адреса аббревиатуру «URL» (Uniform Resource Locator — унифицированный указатель ресурса).

Информацию о протоколе доступа к сетевому ресурсу (ftp, http и т. п.) и его электронный адрес приводят в формате унифицированного указателя ресурса.

После электронного адреса в круглых скобках приводят сведения о дате обращения к электронному сетевому ресурсу: после слов «дата обращения» указывают число, месяц и год:

⁵ Весь Богородский уезд: форум // Богородск — Ногинск. Богородское краеведение: сайт. Ногинск, 2006. URL: <http://www.bogorodsk-noginsk.ru/forum/> (дата обращения: 20.02.2007).

46. Инвестиции останутся сырьевыми//PROGNOSIS.RU: ежедн. интернет-изд. 2006.25января. URL: <http://www.prognosis.ru/print.html?id=6464> (дата обращения: 19.03.2007).

9. Жилищное право: актуальные вопросы законодательства: электрон, журн. 2007. № 1. URL: <http://www.gilpravo.ru> (дата обращения: 20.08.2007).