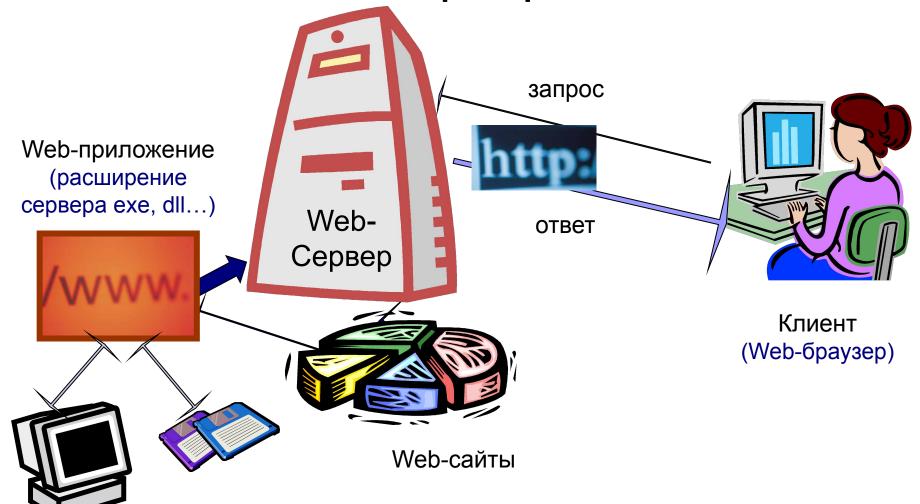
Основные службы Internet

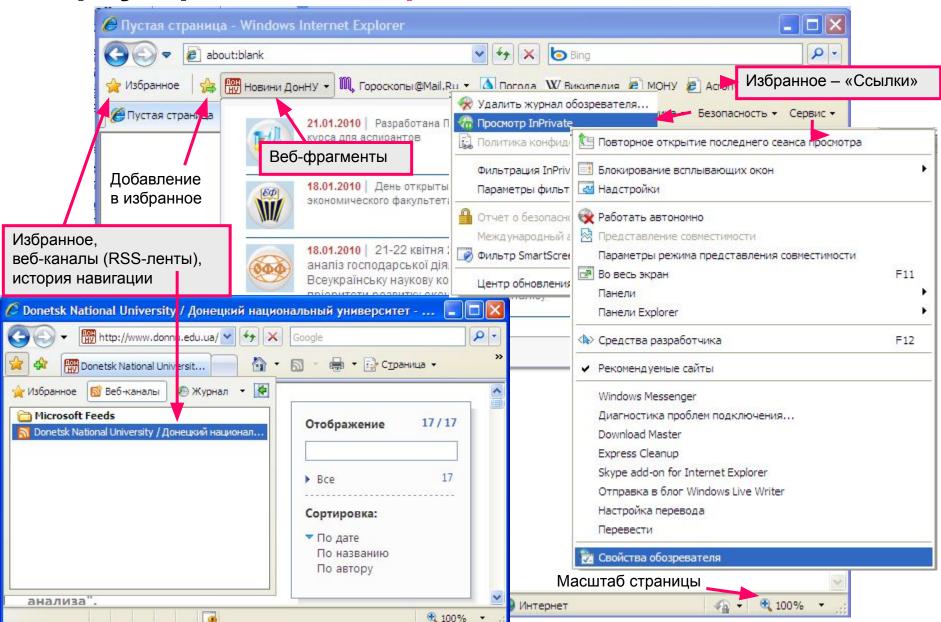
- Коммуникация:
 - Почта E-mail. Телеконференции (Newsgroups), чаты, пейджеры (IRC, ICQ, Skype...)
- Информация:

Всемирная паутина WWW (Web).

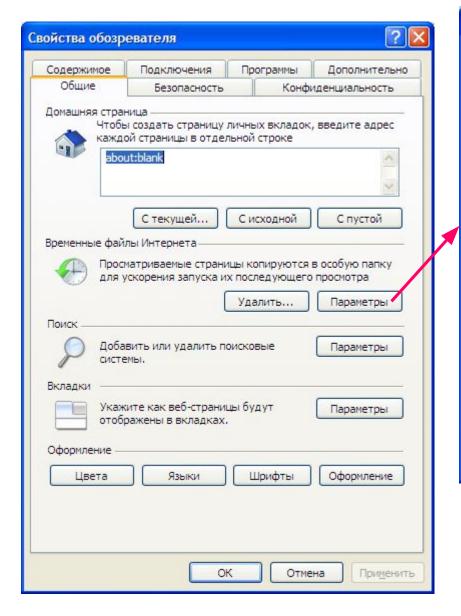
Взаимодействие Клиент – Web-сервер

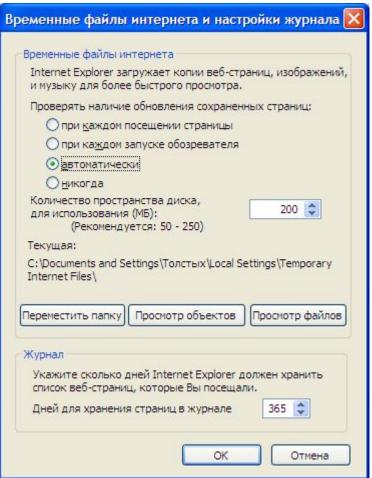


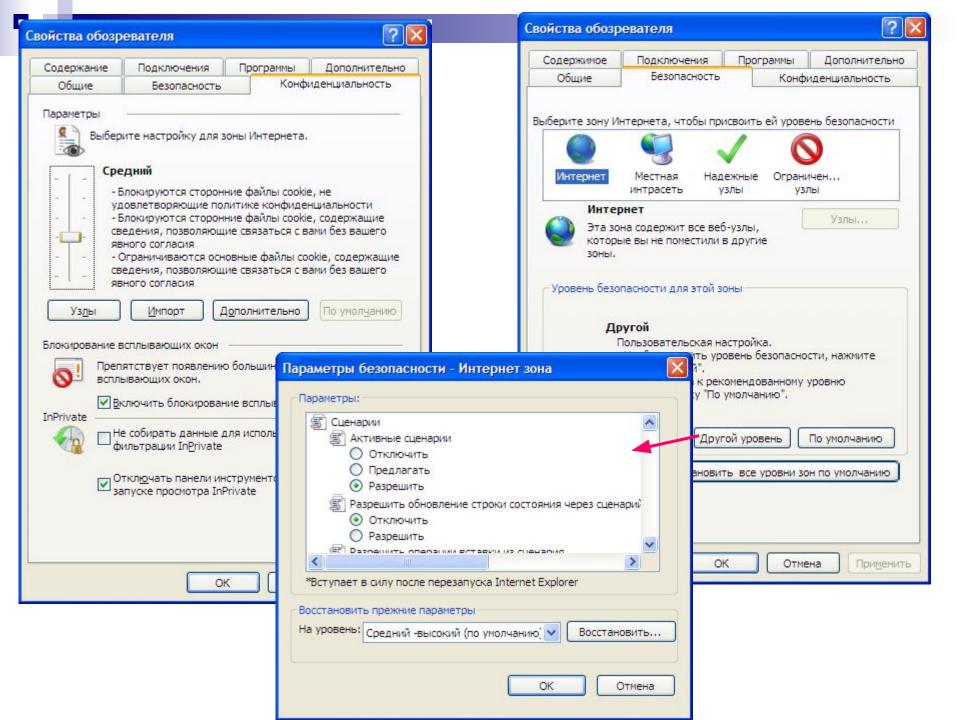
Браузер Internet Explorer 8.0

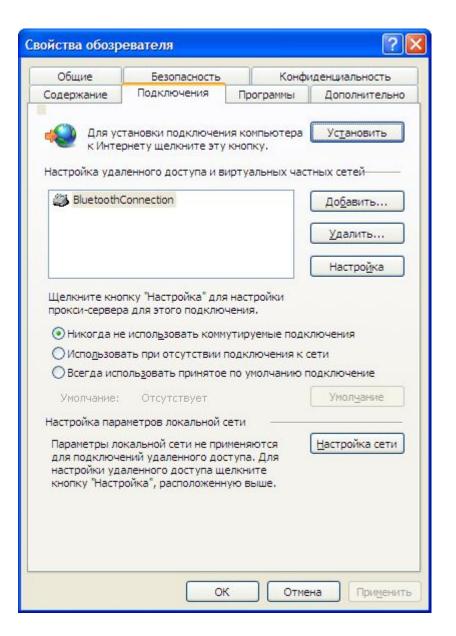


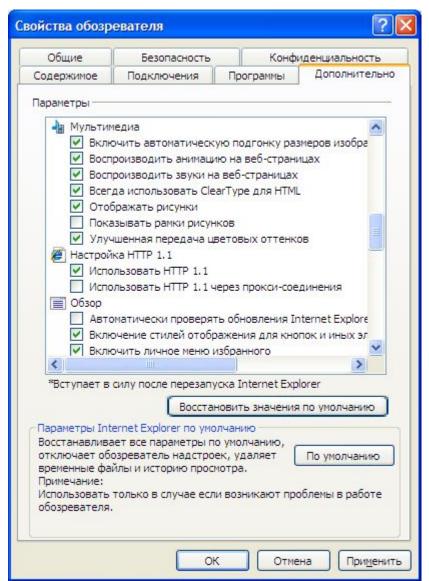
Свойства браузера Internet Explorer 8.0

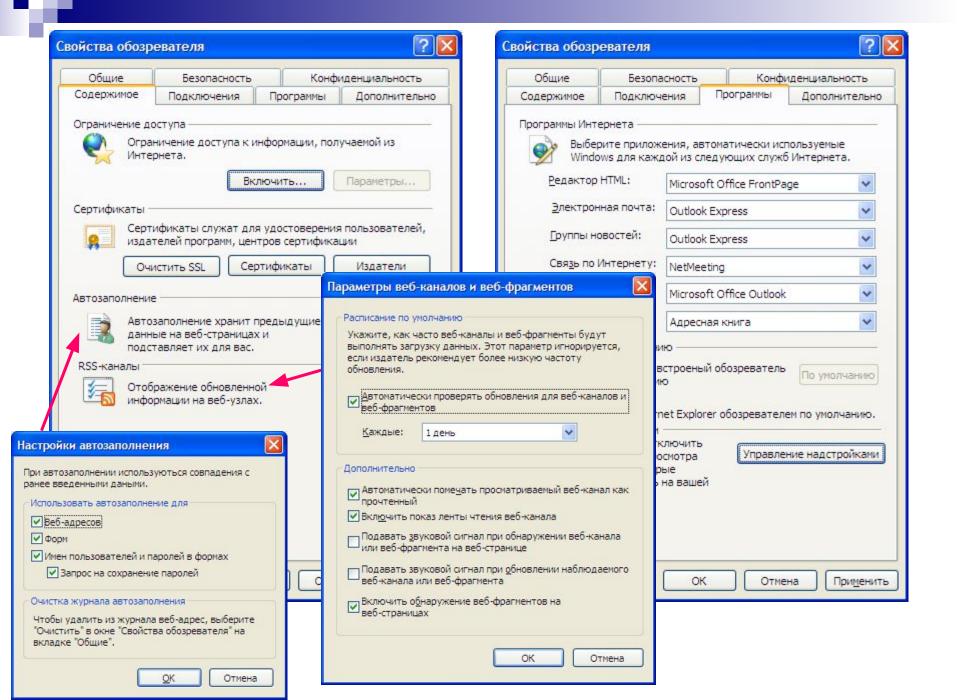






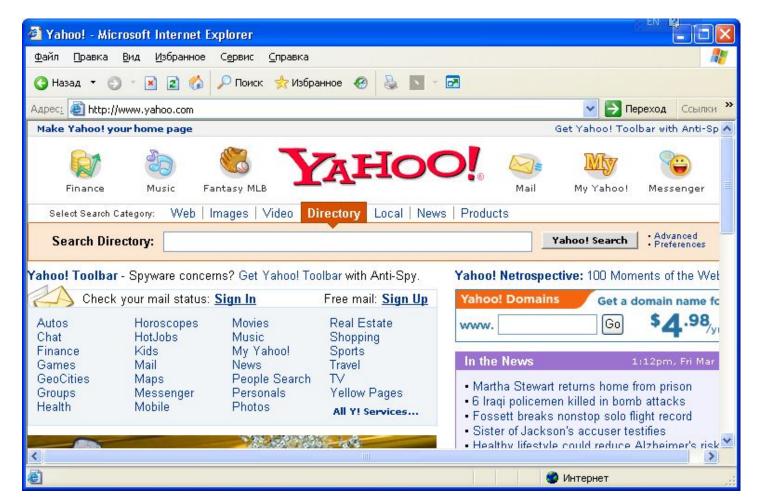






Индексированные каталоги:

классификаторы развивают систематический каталог, систематизаторы читают Web-страницы и систематизируют по разделам каталога



- Словарные информационно-поисковые системы (ИПС) База данных слов → к каждому слову приписывается адрес Webстраницы, где «паук» нашел это слово. Поиск адресов в базе по: AND (&, и) OR (|, или), NOT (!, -, не), "фраза", (), near и др...
- Метапоисковые машины (рассылка запросов по поисковым машинам)
 meta-ukraine.com, www.metacrawler.com, www.highway61.com
- Контекстные поисковые машины www.oingo.com, www.gurunet.com, www.simpli.com

Как работают ИПС

Поисковая машина — это комплект программ:

- **Spider** («паук») программа, которая загружает в поисковую машину Webстраницы. Работает аналогично браузеру, но ничего не отображает ни на каком экране.
- Crawler («червяк», или «путешествующий паук») программа, способная найти на Web-странице все ссылки на другие страницы. Ее задача определить, куда дальше должен ползти «паук», руководствуясь ссылками или заранее заданным списком адресов.
- Indexer (индексатор) программа, которая «разбирает» страницу на составные части и анализирует их. Вычленяются и анализируются заголовки Web-страниц, заголовки документов, ссылки, текст документов, отдельно текст, выделенный полужирным шрифтом, курсивом и т.д. Глубина индексации может быть разной. Полные тексты документов в базу данных копируют не все поисковые роботы некоторые ограничиваются лишь заголовками.
- **Database** (база данных) хранилище всех данных, которые поисковая система загружает и анализирует. Требует огромных ресурсов как для хранения, так и для последующей обработки.
- Search Engine Results Engine (система выдачи результатов поиска) решает, какие страницы удовлетворяют запросу пользователя и в какой степени. Именно с этой частью поисковой системы «общается» пользователь.



Международные:

- http://www.google.com
- http://scholar.google.com.ua
- http://www.bing.com
- http://search.yahoo.com
- http://www.ask.com
- http://www.alltheweb.com
- http://www.lycos.com
- http://www.excite.com
- http://www.altavista.com

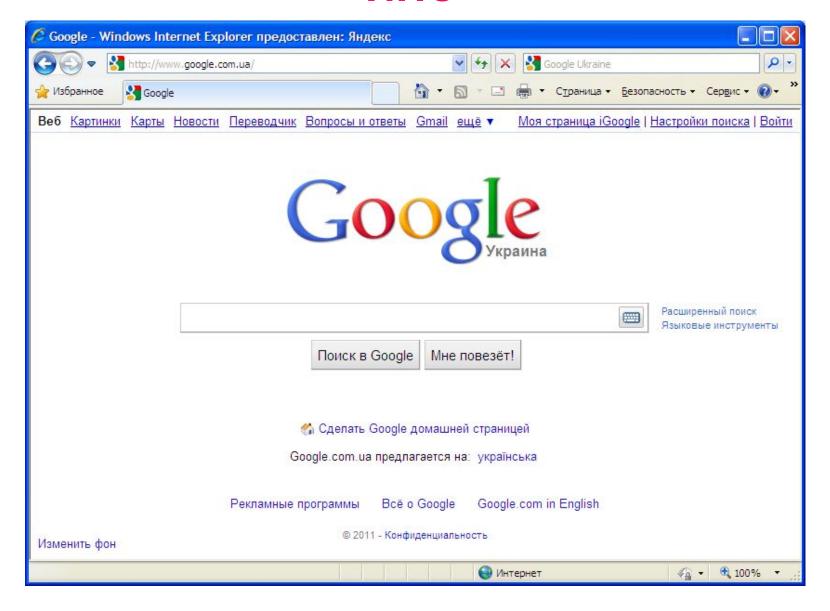
Российские:

- http://www.yandex.ru
- http://www.rambler.ru
- http://www.aport.ru

Украинские:

- http://meta.ua
- http://uaport.net

ИПС



🌈 Справка по расширенному поиску : Основные сведения о Поиске Google - Справка - Веб-поиск - Windows Internet Exploi





http://www.google.com/support/websearch/bin/answer.py?answer=136861



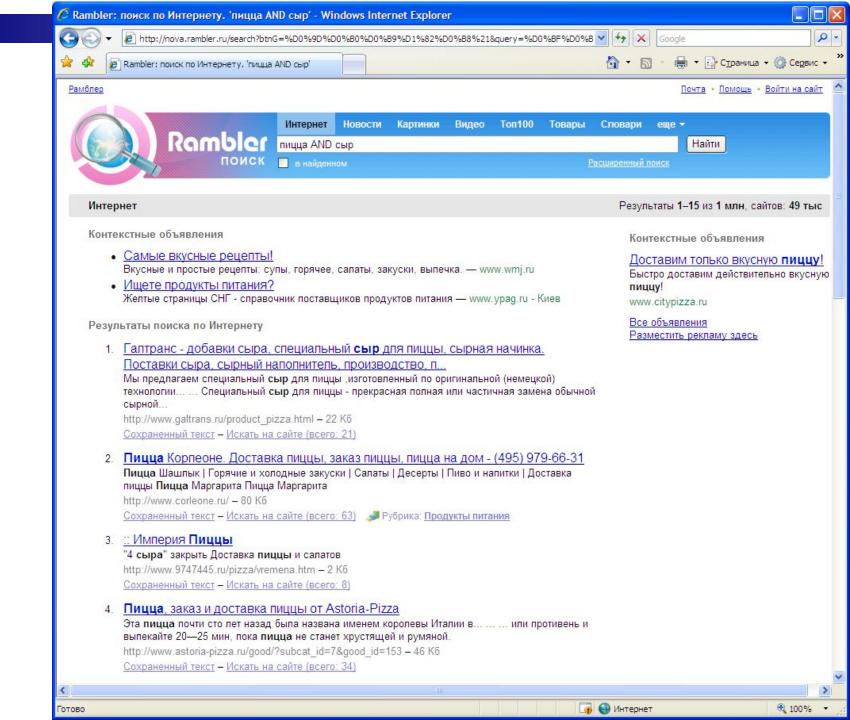
Справка по расширенному поиску: Основные све...

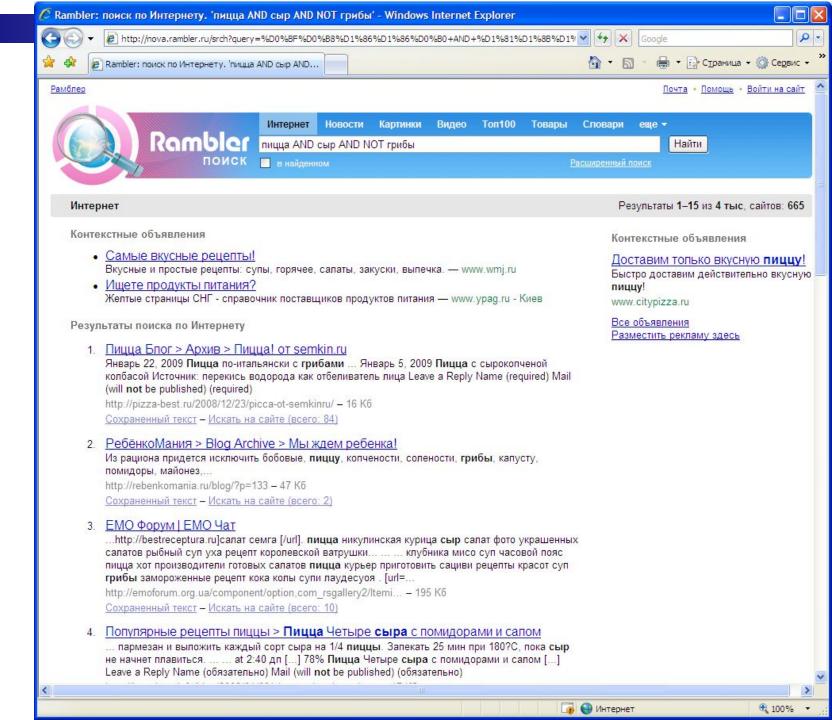
Основные сведения о Поиске Google: Справка по расширенному поиску



В статье Справка по базовому поиску освещены все основные приемы выполнения поиска, но иногда требуется чуть больше возможностей. В данном документе мы рассмотрим расширенные функции веб-поиска Google. При этом не стоит забывать, что даже такие продвинутые пользователи, как члены группы поиска в компании Google, используют эти функции менее 5% от общего числа раз. Простого базового поиска часто бывает вполне достаточно. Как всегда, мы используем квадратные скобки [] для обозначения запросов, так что [быть или не быть] - это пример запроса; [быть] или [не быть] - это два примера запросов.

- Поиск по словосочетанию ("")
 - Если вы заключаете набор слов в двойные кавычки, то тем самым даете команду рассматривать указанные слова именно в таком порядке, без изменений. Google уже использует этот порядок слов и просто так от него отклоняться не будет, поэтому кавычки в большинстве случаев излишни. Настаивая на поиске по словосочетанию, вы можете случайно пропустить важные результаты. Например, при поиске ["Александр Пушкин"] (в кавычках) будут пропущены страницы, в которых упоминается Александр Сергеевич Пушкин.
- Поиск в пределах определенного веб-сайта (site:)
 - Google позволяет указывать, что результаты поиска должны быть с указанного веб-сайта. Например, запрос [ирак site:kommersant.ru] возвратит страницы об Ираке, но только с сайта kommersant.ru. Более простые запросы [ирак kommersant.ru] или [ирак Коммерсант] обычно бывают так же эффективны, хотя могут возвратить результаты с других сайтов, упоминающих "Коммерсант". Можно указать целый класс веб-сайтов. Например, [ирак site:.ru] возвратит результаты только из домена .ru, а [ирак site:.iq] возвратит результаты только с иракских веб-сайтов.
- Поисковые слова, которые нужно исключить (-)
- Добавление знака минуса прямо перед словом приведет к тому, что страницы, содержащие это слово, не будут появляться в ваших результатах поиска. Указывайте знак минуса непосредственно перед словом, а перед минусом ставьте пробел. Например, в запросе [кисло-сладкий соус]знак минуса используется как дефис и не будет считаться символом исключения. А в запросе [кисло-сладкий -соус] будет выполнен поиск слов "кислосладкий", но будут исключены ссылки на соус. Можно исключить любое число слов, ставя знак - перед каждым из них, например [ягуар -автомобили -футбол -ос]. Знак - можно использовать для исключения не только слов. Например, поставьте его перед оператором "site:" (без пробела), чтобы исключить определенный сайт из результатов поиска.





PageRank

PageRank — это алгоритм, позволяющий оценить, насколько данная интернет-страница популярна, то есть это функция от интернет-страницы, которую можно сосчитать заранее.

Сергей Брин в 1998 году предложил следующую идею: определять рейтинг страницы через количество ведущих на нее ссылок и рейтинг ссылающихся страниц.

Google PageRank учитывает не все ссылки. Поисковая система отфильтровывает ссылки с искусственно созданных сайтов, специально предназначенных для скопления ссылок. Некоторые ссылки могут не только не учитываться, но и отрицательно сказаться на ранжировании ссылающегося сайта.

Также в Интернете используются и другие методы оценки популярности страницы:

- учет частоты обновляемости страницы (чем чаще страница обновляется, тем она «лучше»);
- учет посещаемости (чем больше пользователей посещают страницу, тем она «лучше»);
- учет регистрации в каталоге-спутнике поисковой системы.

SEO – Search Engine Optimization

Поисковая оптимизация направлена на увеличение количества посетителей Web-сайта за счёт повышения ранга сайта (без оплаты поисковым компаниям).

Релевантность

Релевантность – соответствие информации отправленному запросу:

$$P = \frac{N1}{N} 100\%$$

где N1 — количество документов, поинформационная сути соответствующих запросу, N — всего полученных документов.

Пертинентность – соотношение объема *полезной для пользователя* информации к общему объему полученной информации (имеет решающее значение)



Релевантность и пертинентность

М

Примеры профессиональных запросов к ИПС

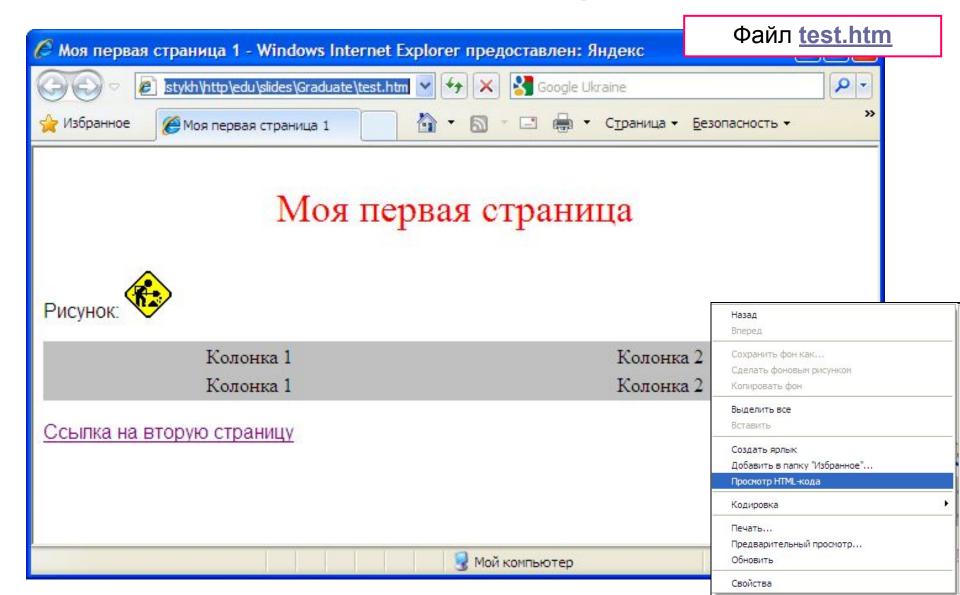
Запрос к системе "Интегрум" по теме" Услуги связи:

"услуги связи" или "междугородные переговоры" или "телефонные переговоры" или "мобильная связь" или "фиксированная связь" или "сотовая связь" или "сотовый оператор" или "средства связи" или "телефонная связь" или "спутниковая связь" или "космическая связь" или GPRS или ростелеком или связьинвест или госкомсвязь или госкомтелеком или госсвязьнадзор или телекоммуникации или электросвязь или АТС или ГТС или минсвязи или "министерство связи" или "волоконно-оптическая линия связи" или ВОЛС

Запрос к системе InfoStream по теме "Мобильная связь":

```
(((мобильн~связ) | (мобільн~зв'яз) | (сотов~связ) | (стільник~зв'яз) | (беспроводн~связ) | (бездрот~зв'яз) | (бесперебойн~связ) | (безперебійн~зв'яз) | j2me] | ems] | 3g] | gprs] | ggsn] | sgsn] | sms] | mms] | ems] | bluetooth] | mms] | tdma] | multipoint] | pcs] | cdma] | ofdm] | vpn] | wap] | umts] | gsm)&((моб~телефон)) | (стільник~телефон) | (сотов~телефон))) ! this.is
```

Подготовка Web-страниц



Исходный код страницы

```
<html>
 <head>
 <title>Moя первая страница 1</title>
 </head>
         Внутренние стили
 <body>
 Моя первая страница
 Рисунок: <img src="atwork.gif" width="38" height="38">
 <del>Внут</del>ренние стили
      Колонка 1
      Koлoнка 2
   Koлoнка 1
      Koлoнка 2
   18 
 <a href="test2.htm">Ссылка на вторую страницу</a>
26 snbsp;
21 </body>
 </html>
```

Заголовок страницы со стилями

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w</pre>
<html xmlns="http://www.w3.org/1999/xhtml" xmlns:v="urn:schemas-microsoft-co</pre>
<head>
<meta http-equiv="Content-Language" content="ru" />
<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
<title>Moя первая страница 1</title>
<style type="text/css">
                                                                  Таблица
.Head1 {
                                                             глобальных стилей
           text-align: center;
                                                                  (классы)
            font-size: xx-large;
            color: #FF0000:
.Normal {
            text-align: left;
            font-family: Arial, Helvetica, sans-serif;
</style>
</head>
```

