

# Интеграция журнала LiveJournal в персональный сайт

## 1 Вступление

Все большую популярность среди пользователей интернета приобретает ведение личных онлайн-дневников. Существует множество сайтов, которые позволяют публиковать свои откровения в интернете легко и непринужденно, без какого-либо знания специальных языков программирования и разметки гипертекста. Один из таких сайтов - LiveJournal ([www.livejournal.com](http://www.livejournal.com)), интернет-сервис, где каждый желающий может завести себе личный дневник.

## 2 Задача

Создать программный комплекс, который позволит интегрировать журнал из Livejournal в персональный сайт, предусмотреть возможность отображения постов и комментариев с использованием предустановленного шаблона.

## 3 Существующие решения

Существует несколько методов решения данной задачи, которые описаны непосредственно на сайте LiveJournal (<http://www.livejournal.com/developer/embedding.bml?method=all>):

1. Самый простой путь встроить ваш журнал в персональный сайт - это вставить JavaScript в HTML-код страницы. В браузерах, где JavaScript отключен, возможно отображение ссылки на журнал.

```
<script language="JavaScript"
src="http://www.livejournal.com/customview.cgi?username=zhupanenko&styleid=101&enc=js">
<noscript><a href="http://zhupanenko.livejournal.com/">View my
LiveJournal</a></noscript>
</script>
```

2. Встраивание Livejournal при помощи фреймов. Данный метод будет работать только в том случае, если браузер поддерживает отображение фрейм-структур.

```
<center>
<iframe name="livejournal"
src="http://www.livejournal.com/users/zhupanenko/"
frameborder="0"
scrolling="auto" height="400" width="460">
<a href="http://zhupanenko.livejournal.com/">View my LiveJournal</a>
</iframe>
</center>
```

3. Интеграция при помощи CGI-скриптов или PHP. Средствами языка программирования производится чтение данных с удаленного ресурса и затем последующий вывод в браузер без обработки информации.

```
<?php
$journalURL = "http://www.livejournal.com/".
             "customview.cgi?username=zhupanenko&styleid=101";

if ($fh = fopen($journalURL,"r")) {
    fpassthru($fh);
} else {
    echo "<p>Unable to load journal.</p>\n";
}
?>
```

Очевидное преимущество всех существующих решений – простота реализации. Подобная интеграция займет не более 10 минут с минимальным увеличением скорости загрузки страницы.

Недостатки: все вышеописанные способы интеграции доступны только для пользователей с платными экаунтами. В случае же их модификации под бесплатные журналы на вашей странице появится реклама которую необходимо будет дополнительно фильтровать. Так же появляется проблема несоответствия дизайна станицы и встраиваемого журнала. Вывод: ни один из способов внедрения от Livejournal не предоставляет гибкости в настройках отображения записей и комментариев журнала.

#### **4 Решение**

Суть предлагаемого метода – парсинг записей и комментариев журнала Livejournal и последующий вывод результатов с использованием предустановленного шаблона.

#### **5 Используемые технологии**

PHP, MySQL, RSS, Sockets

#### **6 Кросс-браузерность**

Internet Explorer 5.0,6.0,7.0; FireFox 1.0,2.0; Opera 9.0; Safari 2.0;

#### **7 Входные и выходные параметры**

##### Входные параметры:

1. \$user – имя пользователя Livejournal, журнал которого будет интегрироваться в персональный сайт;
2. \$postCount – количество записей, отображаемых на странице;

##### Выходные параметры:

1. Структура данных, содержащая в себе записи и комментарии журнала Livejournal.
2. Фрагмент HTML-кода, отображающий записи и комментарии журнала Livejournal с использованием предустановленного шаблона.

#### **8 Реализация**

Для начала нам необходимо получить ленту записей (постов) пользователя. К счастью Livejournal выдает по RSS-каналу всю необходимую информацию о дневнике и последних 25 записях. Ниже приведен фрагмент RSS-потока одного из журналов

```
<?xml version='1.0' encoding='utf-8' ?>
<rss version='2.0' xmlns:lj='http://www.livejournal.org/rss/lj/1.0/'>
<channel>
  <title>второе, что пришло в ро...</title>
  <link>http://shstepka.livejournal.com/</link>
  <description>второе, что пришло в ро... - LiveJournal.com</description>
  <lastBuildDate>Thu, 24 May 2007 12:26:15 GMT</lastBuildDate>
  <generator>LiveJournal / LiveJournal.com</generator>

  <image>
    <url>http://userpic.livejournal.com/61827888/6441793</url>
    <title>второе, что пришло в ро...</title>
    <link>http://shstepka.livejournal.com/</link>
    <width>100</width>
    <height>100</height>

  </image>

<item>
  <guid isPermaLink='true'>http://shstepka.livejournal.com/196544.html</guid>
  <pubDate>Thu, 24 May 2007 12:26:15 GMT</pubDate>
```

```

<title>Вопрос</title>
<link>http://shstepka.livejournal.com/196544.html</link>
<description>В Windows нельзя создать файл или папку под названием &quot;Con&quot;;,
ибо у Билла Гейтса в детстве была прозвище &quot;Con&quot;; то есть
&quot;ботаник&quot;. И он постарался чтобы в его системе отсутствовали такие файлы и
папки.&lt;br /&gt;&lt;br /&gt;есть ли этому РАУМНОЕ объяснение?</description>

<comments>http://shstepka.livejournal.com/196544.html</comments>
<lj:security>public</lj:security>
</item>
<item>
<guid isPermaLink='true'>http://shstepka.livejournal.com/196104.html</guid>
<pubDate>Wed, 23 May 2007 13:12:18 GMT</pubDate>
<title>Ветер знает, где меня искать!</title>

<link>http://shstepka.livejournal.com/196104.html</link>
<description>На обеденном перерыве пускали змея с мальчиком Феем или мальчиком
Тимой, ну&lt;br /&gt;вообщем с Тимофеем из Мира Детства!!! Мы шли на речку, несли змея
в руках с катушкой, очень похож он на удочки, а у прохожих возникал вопрос при виде
змея и нас: &quot;Что ловить будете?&quot;;- а мы хором думали:
&quot;Ветер&quot;.&lt;br /&gt;Тима поймал ветер, а ветер поймал змея, а змей поймал
Тиму и мы понеслись по мосту над рекой вниз, расправив руки и крылья, а под нами
летали чайки, именно под нами и именно ЧАЙКИ, оказывается эти птицы бывают и в
Москве, только надо очень хотеть их увидеть и услышать.</description>
<comments>http://shstepka.livejournal.com/196104.html</comments>

<lj:security>public</lj:security>
</item>
<item>
<guid isPermaLink='true'>http://shstepka.livejournal.com/195985.html</guid>
<pubDate>Tue, 22 May 2007 12:52:17 GMT</pubDate>
<title>игра слов</title>
<link>http://shstepka.livejournal.com/195985.html</link>

<description>Пойду домой... дойду - помой!&lt;br /&gt;&lt;span
class=&apos;ljuser&apos;; lj:user=&apos;zero_result&apos;; style=&apos;white-space:
nowrap&apos;;&gt;&lt;a href=&apos;http://zero-
result.livejournal.com/profile&apos;;&gt;&lt;img
src=&apos;http://stat.livejournal.com/img/userinfo.gif&apos;; alt=&apos;[info]&apos;;
width=&apos;17&apos;; height=&apos;17&apos;; style=&apos;vertical-align: bottom; border:
0&apos;; /&gt;&lt;/a&gt;&lt;a href=&apos;http://zero-
result.livejournal.com/&apos;;&gt;&lt;b&gt;zero_result&lt;/b&gt;&lt;/a&gt;&lt;/span&gt;
</description>

<comments>http://shstepka.livejournal.com/195985.html</comments>
<category>слова</category>
<lj:security>public</lj:security>
</item>
...
</channel>
</rss>

```

Поскольку одним из самых распространенных языков написания скриптов в сети является PHP, мы воспользуемся именно этим интерпретатором для интегрирования дневника в сайт.

Широко распространенных способов обработки XML-документов существует два — Event-based APIs и Document Object Model (DOM) APIs. В PHP стандартная поддержка XML организована с помощью Event-based API (основана на событиях).

Сперва создадим класс RSSParser, внутри которого будет выполняться вся работа по разбору XML. После создания класса, получим RSS-данные от сервиса LiveJournal и инициализируем обработчик XML, который будет использовать для событийной обработки (Event-based API) класс RSSParser. Код класса RSSParser приведен ниже:

```

//RSSParser class
class RSSParser
{
    var $postCurrentNumb = 0;
    var $postCount = 0;
    var $insideItem = false;
    var $tag = "";
    var $maxPostSize = 999999;

    var $imageURL = "";
    var $lastBuildDate = "";
    var $lbd = "";
    var $userPic = "";

    var $title = "";
    var $dt = "";
    var $text = "";
    var $category = "";
    var $comments = "";

    var $lj = array();

    function startElement($parser, $tagName, $attrs)
    {
        if($this->insideItem)
        {
            $this->tag = $tagName;
        }
        elseif($tagName == "ITEM")
        {
            $this->insideItem = true;
        }
        elseif($tagName == "IMAGE")
        {
            $this->insideItem = true;
        }
        elseif($tagName == "LASTBUILDDATE")
        {
            $this->tag = $tagName;
            $this->insideItem = false;
        }
    }

    function endElement($parser, $tagName)
    {
        if($tagName == "IMAGE")
        {
            $this->userPic=$this->imageURL;
            $this->title = "";
            $this->insideItem = false;
        }

        if($tagName == "ITEM")
        {
            if ($this->postCurrentNumb<$this->postCount)
            {
                $this->lj[$this->postCurrentNumb][title]=UtoW($this-
>title);
                $this->text=UtoW($this->text);
                if (strlen($this->text)>$this->maxPostSize)
                $this->text=substr($this->text,0,$this-
>maxPostSize);
                $this->lj[$this->postCurrentNumb][text]=$this->text;
                $this->lj[$this-
>postCurrentNumb][comments]=trim($this->comments);
                $this->lj[$this->postCurrentNumb][tag]=UtoW($this-
>category);
                $this->lj[$this-
>postCurrentNumb][dt]=dateConvert($this->dt);
            }
        }
    }
}

```



Теперь создадим функцию, которая будет выполнять инициализацию обработчика RSS и его запуск.

```
//generate posts
function generatePostData($postCount)
{
    $this->postsCount=$postCount;
    $xml_parser=xml_parser_create("UTF-8");
    $rss_parser=new RSSParser();
    $rss_parser->postCount=$this->postsCount;
    $rss_parser->maxPostSize=$this->maxPostSize;
    xml_set_object($xml_parser, &$rss_parser);
    xml_set_element_handler($xml_parser,          "startElement",
"endElement");
    xml_set_character_data_handler($xml_parser, "characterData");

    $fp = fopen("http://".$this->user.".livejournal.com/data/rss",
"r")
        or die("Error reading RSS data from http://".$this->
user.".livejournal.com/data/rss !");

    while(($data = fread($fp, 4096)) && ($rss_parser->
postCurrentNumb<$rss_parser->postCount))
    {
        xml_parse($xml_parser, $data, feof($fp)
or die("Error parsing RSS data from http://".$this->
user.".livejournal.com/data/rss !");
    }
    fclose($fp);

    xml_parser_free($xml_parser);

    $this->posts=$rss_parser->lj;
    $this->userPic=$rss_parser->userPic;
    $this->lastBuildDate=$rss_parser->lbd;

    if ($this->delcuts)
        $this->clearLivejournalCuts();
}
}
```

Данная функция является одним из методов класса Livejournal. Параметр `$this->postsCount` определяет количество записей, которые будут считаны, `$this->maxPostSize` - максимальный размер поста в байтах. Как результат выполнения функции `generatePostData()` мы получаем двухмерный массив `$this->posts`, в котором находятся записи журнала разбитые на элементы: Заголовок, Дата, Текст, Ссылка на комментарии.

Следующий шаг – получение комментариев для каждой из записей. Все было бы так же просто, если Livejournal выдавал ленту комментариев по RSS. К сожалению, такой услуги нет ни для платных, ни для бесплатных экаунтов. Комментарии придется получать при помощи HTML-парсера.

Существует две системы стилей (<http://www.livejournal.com/customize/>) Livejournal при помощи которых происходит отображение журналов - S1 (для пользователей, знакомых с CSS и HTML) и S2 (для остальных). Кроме того, существует несколько стилей для просмотра журналов Livejournal (<http://www.livejournal.com/manage/settings/>): Horizon, XCalibur, Dystopia, Lunx. Самым "легким" из них является Lunx. Именно в этом стиле мы будем считывать страницы с комментариями пользователей. Кроме того, это позволит нам написать универсальный парсер для всех стиливых отображений журналов.

Метод `getSocketData()` считывает HTML-данные со страницы с комментариями. Для того, чтоб использовать стиль Lunx необходимо в конце адресной строки дописать параметр `"?format=light"`.

```

//get socket data
function getSocketData($host,$request)
{
    if ($fp = fsockopen($host, 80, $errno, $errstr, 5))
    {
        fputs($fp,$request);
        $data="";
        while(!feof($fp))
            $data.=fgets($fp,2048);
        fclose($fp);
    }
    return ($data);
}

//get livejournal comments
function getLJCommentsSock($ljUser, $commentsId)
{
    $commentsPath=$commentsId.".html?format=light";
    $host="livejournal.com";
    $commonRequest=
        "Accept: */*"\r\n".
        "Accept-Language: ru"\r\n".
        "User-Agent: Mozilla/4.0 (compatible;
MSIE 5.0; Windows NT) "\r\n".
        "Connection: Keep-Alive"\r\n";

    // Get request

    $request=
        "GET /".$commentsPath." HTTP/1.0"\r\n".
        $commonRequest.
        "Cookie: ".$cookie."\r\n".
        "Host: ".$ljUser.".livejournal.com\r\n"\r\n".
        "Pragma: no-cache"\r\n";

    $data=$this->getSocketData($host, $request);

    $temp1=strpos($data,"/><a name='")+2;
    $temp2=strpos($data,"<hr />",$temp1);

    if ($temp1 && $temp2)
        $data = substr($data, $temp1, $temp2-$temp1);
    else
        $data="";

    $this->commentsData=UtoW($data);
}

```

Необходимый для нас блок информации находится между фрагментами текста `"/><a name='"` и `"<hr />"`

Следующий шаг – передаем текстовый блок с комментариями для парсинга методу `parseComments()`. Обработку HTML выполняет экземпляр класса `HtmlParser`.

```

//parse livejournal comments
function parseComments()
{
    $j=0;
    $prevName="";
    $parser = new HtmlParser($this->commentsData);

    while ($parser->parse())
    {
        if ($parser->iNodeName=="span" && $parser->iNodeType ==
NODE_TYPE_ELEMENT)
        {
            $attrValues = $parser->iNodeAttributes;
            $attrNames = array_keys($attrValues);
            $size = count($attrNames);
            for ($i = 0; $i < $size; $i++)

```

```

        {
            $name = $attrNames[$i];
            if ($attrNames[$i]=="lj:user")
                $this-
>comments[$j][user]=UtoW($attrValues[$name]);
        }
    };

    if ($parser->iNodeName=="Text" && ($parser->iNodeType ==
NODE_TYPE_TEXT || $parser->iNodeType == NODE_TYPE_COMMENT))
    {
        if (strpos($parser->iNodeValue,"UTC"))
            $this->comments[$j][dt]=dateConvert($parser-
>iNodeValue);

        if ($prevName=="td" && trim($parser->iNodeValue))
            $this->comments[$j][text].=$parser->iNodeValue;

        if ($prevName=="br")
            $this->comments[$j][text].="<BR>".$parser-
>iNodeValue;

        if ($prevName=="b" && !$this->comments[$j][user] &&
$parser->iNodeValue!="(")
            $this->comments[$j][subj]=$parser->iNodeValue;
    };

    if ($parser->iNodeName=="img" && $parser->iNodeType ==
NODE_TYPE_ELEMENT)
    {
        $attrValues = $parser->iNodeAttributes;
        $attrNames = array_keys($attrValues);
        $size = count($attrNames);
        for ($i = 0; $i < $size; $i++)
        {
            $name = $attrNames[$i];
            if ($name=="src" &&
strpos($attrValues[$name],"userpic"))
                $this-
>comments[$j][pic]=$attrValues[$name];
        }
    }

    if ($parser->iNodeName=="a" && $parser->iNodeType ==
NODE_TYPE_ELEMENT)
    {
        $attrValues = $parser->iNodeAttributes;
        $attrNames = array_keys($attrValues);
        $size = count($attrNames);
        for ($i = 0; $i < $size; $i++)
        {
            $name = $attrNames[$i];
            if ($name=="href" &&
strpos($attrValues[$name],"replyto"))
            {
                $this-
>comments[$j][reply]=$attrValues[$name];
                $j++;
                $this->comments[$j][text]="";
            }
        }
    }

    $prevName=$parser->iNodeName;
};

$this->commentsCount=$j;
}

```

Как результат выполнения функции `parseComments()` мы получаем двухмерный массив `$this->comments`, в котором находятся комментарии к текущей записи журнала разбитые на элементы: Имя пользователя, Адрес юзерпика, Заголовок, Дата, Текст, Ссылка на ответ.

В качестве дополнительных параметров можно установить максимальный размер записи (`var $maxPostSize = 999999`) и возможность удаления катов (`var $delcuts = true`);

## 9 Пример работы

В качестве примера работы программного комплекса создадим произвольный шаблон для отображения записей и комментариев журнала Livejournal.

```
<?php

require_once('inc/function.inc.php');
require_once('class/rssparser.class.php');
require_once('class/htmlparser.class.php');
require_once('class/lj.class.php');

connect_to_db($host,$login,$pass,$db_name);

$user="shstepka";
$postCount=10;

$lj=new Livejournal($user);

if ($_GET[comment])
{
    $lj->generateCommentsData($lj->user, $_GET[comment]);

    $query="SELECT value FROM info WHERE name='userpic' OR
name='lastbuilddate' ORDER BY name ASC";
    $result=mysql_query($query);
    $row=mysql_fetch_array($result);
    $lbd=$row[value];
    $row=mysql_fetch_array($result);
    $userpic=$row[value];

    $postOut='<table width="600px" align="center" cellpadding="10">';

    $query="SELECT text, textfull FROM posts WHERE
commentid='".$_GET[comment].'";
    $result=mysql_query($query);
    $row=mysql_fetch_array($result);

    if ($row[textfull])
        $row[text]=$row[textfull];

    $postOut.='<tr><td><table width="100%" style="border: 1px dashed #000000;
background: url(\'/img/back2.jpg\');"><tr>';
    $postOut.='<td style="padding: 20px" valign="top"></td>';
    $postOut.='<td width="100%" style="padding: 20px 20px 20px 0px"><p
class="text">'.str_replace("&rsquo;", "", $row[text]).'</p></td>';
    $postOut.='</tr><tr><td colspan="2" align="right" style="padding-right:
10px"><A href="http://'. $lj-
>user.'.livejournal.com/'. $_GET[comment].'.html?mode=reply"
target="_blank">reply...</A></td></tr></table></td></tr>';

    for ($i=0; $i<$lj->commentsCount; $i++)
    {
        $imgOut = ($lj->comments[$i][pic]) ? '<BR><A href="http://'. $lj-
>comments[$i][user].'.livejournal.com" target="_blank" style="line-height:
0"><b>'. $lj->comments[$i][user].'</b></A>' : '<b>anonymous</b>';

        $postOut.='<tr><td><table width="100%" style="border: 1px dashed
#000000; background: url(\'/img/back.gif\'); height: 100%">';

        $postOut.='<tr><td style="padding: 20px"
align="center">'. $imgOut.'</td><td width="100%" valign="top" style="height:
100%"><table width="100%" style="height: 100%">';
```



```

                $dateOut=substr($lj->posts[$i][dt],8,2)."-" .substr($lj-
>posts[$i][dt],5,2)."-" .substr($lj->posts[$i][dt],0,4)." | ".substr($lj-
>posts[$i][dt],11,5);

                $postOut.='<tr><td align="right" style="padding: 4px"><span
class="dt">['.$dateOut.']</span></td></tr>';

                $postOut.='<tr><td colspan="2" align="center"><span
class="title">'.$lj->posts[$i][title].</span></td></tr>';
                $postOut.='<tr><td colspan="2" style="padding: 10px"><p
class="text">'.$lj->posts[$i][text].</p></td></tr>';
                $postOut.='<tr><td colspan="2" align="right" style="padding-right:
10px"><A href="/'.$lj->path.'/'.$commentId.'">comments...</A></td></tr>';

                $postOut.='</table></td></tr>';

        }

        $postOut.='</table>';

}

?>

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN">
<html>

<head>
<title>Livejournal Integrator Sample</title>
<meta http-equiv="Content-Type" content="text/html; charset=windows-1251">
<style type="text/css">
    html, body
    {
        margin: 0;
        padding: 0;
        height: 100%;
        background: url('/img/back.jpg');
    }
    .dt
    {
        font: 13px Arial;
    }
    .text
    {
        font: 14px Tahoma;
        color: #FFF;
    }
    .title
    {
        font: 14px Tahoma;
    }
    a
    {
        color: #1C1FA7;
    }
    a:hover
    {
        color: #1C1FA7;
        text-decoration: none;
    }
</style>
</head>

<body>

<?=$postOut?>

</body>
</html>

```

Livejournal Integrator Sample - Windows Internet Explorer

http://lj/livejournal/ Google

Livejournal Integrator Sample Page Tools

shtepka's livejournal

[24-05-2007 | 15:26]

Вопрос

В Windows нельзя создать файл или папку под названием "Соп", ибо у Билла Гейтса в детстве была прозвище "Соп" то есть "ботаник". И он постарался чтобы в его системе отсутствовали такие файлы и папки.

есть ли этому РАУМНОЕ объяснение?

[comments](#)

[23-05-2007 | 16:12]

Ветер знает, где меня искать!

На обеденном перерыве пускали змея с мальчиком Феем или мальчиком Тимой, ну вообще с Тимофеем из Мира Детства!!! Мы шли на речку, несли змея в руках с катушкой, очень похож он на удочки, а у прохожих возникал вопрос при виде змея и нас: "Что ловить будете?"; - а мы хором думали: "Ветер".

Тима поймал ветер, а ветер поймал змея, а змей поймал Тиму и мы понеслись по мосту над рекой вниз, расправив руки и крылья, а под нами летали чайки, именно под нами и именно ЧАЙКИ, оказывается эти птицы бывают и в Москве, только надо очень хотеть их увидеть и услышать.

[comments](#)

[22-05-2007 | 15:52]

игра слов

Пойду домой... дойду - помой!

[zero result](#)

[comments](#)

Internet 100%

Livejournal Integrator Sample - Windows Internet Explorer

http://lj/livejournal/196544/

Google

Livejournal Integrator Sample

Tools

shtepka

В Windows нельзя создать файл или папку под названием "Соп", ибо у Билла Гейтса в детстве была прозвище "Соп" то есть "ботаник". И он постарался чтобы в его системе отсутствовали такие файлы и папки.

есть ли этому РАУМНОЕ объяснение?

[reply...](#)

[24-05-2007 | 17:00]

 Как оказалось, есть. Знающие люди просветили. Соп - это сокращение от слова консоль, так что ботаник ни причем.

[reply...](#)

[kirdya](#)

[24-05-2007 | 17:04]

shtepka

за консоль - спасибо!  
а почему файлы с таким именем не создаются?  
вот этому я ищу разумное объяснение)))

[reply...](#)

[shtepka](#)

[<<< back](#)

Done

Internet

100%

## 10 Ссылки

1. Обработка LiveJournal RSS стандартными средствами PHP - <http://bikman.ru/texts/techarticle/ljrssphp/>
2. HTML Parser - <http://php-html.sourceforge.net/>
3. Livejournal FAQ - <http://www.livejournal.com/support/faq.bml>

© Жупаненко Андрей  
[research@zhupanenko.com](mailto:research@zhupanenko.com)  
<http://research.zhupanenko.com/livejournal/rus/>