[**https://sitesputnik.ru/Help/I/i\_SiteSputnik\_Text\_Projects.pdf**](https://sitesputnik.ru/Help/I/i_SiteSputnik_Text_Projects.pdf)

Автор: Алексей Мыльников

<https://sitesputnik.ru/>

Дата выхода публикации: 12.04.2017

Дата последнего обновления: 28.08.2023

**SiteSputnik News. Текстовая или рукописная форма новостного Проекта**

**Командный способ формирования Проектов**

**Часть 1**

* В данной публикации показано, как без особого труда можно создавать Проекты для мониторинга открытых Источников интернета.
* Проекты своими руками можно создать, ознакомившись с содержанием только **Части 1** данной публикации, то есть, буквально, через 10 минут после установки программы.
* Для этого достаточно освоить всего два типа команд, которые по формату аналогичны командам, используемым в ОС Windows: ИмяКоманды [Операнды].
* Рассматриваемая форма Проекта называется **текстовой** или **рукописной**. Все действия по проектированию выполняются на одном листе.
* Пояснения сделаны на примере **готового Проекта**, взятого из сообщения на форуме СПКР о [Динамических рубриках СайтСпутника](https://sitesputnik.ru/forum/Sputnik-t-5811-p-56516.html#pp56516).
* Этот Проект выглядит следующим образом:

http://news.vdv-s.ru/rss.xml
https://xml.v102.ru/zagolovki-v102.php
<https://vd-tv.ru/feed>
https://krestyane34.ru/ +.html

Рубрика://Россия/ЮФО/Волгоград/Областная администрация
губернатор /2 (Волгоградской /1 области | Бочаров)

Рубрика://Россия/ЮФО/Волгоград/Сельское хозяйство/Весенние работы
Волгоград /7 (озимые | яровые | посевная |сельхоз\*)
КП: C:\Волгоград\Апрель

* В Проекте задействованы два типа команд: «Определение Источника» и «Определение Рубрики».
	+ Первые четыре строки – это четыре команды типа «Определение Источника».
		- 1-3 строки – это живые RSS-потоки. Команды имеют только имя. Операндов в данном случае нет.
		- 4 строка – это эмуляция RSS-потока (в настоящее время эмуляция уже делается **[автоматически](https://sitesputnik.ru/Help/SSMasNews.htm%22%20%5Cl%20%22robot%22%20%5Co%20%22%D0%9C%D0%B5%D0%BD%D0%B5%D0%B4%D0%B6%D0%B5%D1%80%20%D0%BF%D0%BE%D0%B4%D0%BA%D0%BB%D1%8E%D1%87%D0%B5%D0%BD%D0%B8%D1%8F%20%D0%BD%D0%BE%D0%B2%D0%BE%D1%81%D1%82%D0%BD%D1%8B%D1%85%20%D0%98%D1%81%D1%82%D0%BE%D1%87%D0%BD%D0%B8%D0%BA%D0%BE%D0%B2)**, а именно, указывать +.html не обязательно). Со страницы https://krestyane34.ru/ выбираются и отправляются на мониторинг (скачивание, рубрикацию и некоторые другие операции) только те страницы, в адресе которых есть лексема .html. Эта лексема является операндом. Для того, чтобы понять какой операнд следует добавить в команду, в большинстве случаев достаточно открыть в браузере или СайтСпутнике ссылку-источник и посмотреть какие лексемы характерны для **адресов ссылок на полное содержание статей или сообщений**. Если это не понятно (или [**автоматика**](https://sitesputnik.ru/Help/SSMasNews.htm#robot) не сработает), то пришлите Вашу ссылку-источник на E-mail 1a@sitesputnik.ru, и Вам подскажут полный формат команды. Одной-двух подсказок будет достаточно для того, чтобы далее самостоятельно формулировать команды данного типа. Еще раз повторю, что в настоящее время это делается автоматически, но при желании можно поправить «автоматику».
	+ 5-6 и 7-9 строки – это две команды типа «Определение Рубрики».
		- Имя первой команды - Рубрика://Россия/ЮФО/Волгоград/Областная администрация. Оно определяет название Рубрики. Операнд всего один - губернатор /2 (Волгоградской /1 области | Бочаров). Он определяет Правило попадания информации в Рубрику. В данном случае это сведения о губернаторе области. Примеры основных Правил и пояснения к ним приведены в Приложении 1.
		- Имя второй команды: Рубрика://Россия/ЮФО/Волгоград/Сельское хозяйство/Весенние работы. Операндов два. Второй из них - КП: C:\Волгоград\Апрель. Он определяет дополнительную папку локального компьютера (клиентскую папку), в которую будет отправлено содержание данной Рубрики.
* Собственно, все «начальные трудности» позади, можно перейти к созданию своих Проектов. Последовательность действий такова:
	+ Нажмите на «Золотой ключик», затем на кнопку «Пакет».
	+ Если Вы увидите табличную форму формирования Проекта, то лучше её перевести в текстовую форму: нажмите на кнопку-иконку с подсказкой «К текстовой форме Проекта».
	+ Введите свои команды. Можно ввести или живьём скопировать приведенные выше команды, чтобы протестировать программу.
	+ По кнопке «Сохранение» сохраните сделанный Проект в библиотеке.
	+ По кнопке «Совместно» отправьте его на выполнение.
	+ Далее совершенствуйте созданный готовый Проект:
		- добавляйте в него новые Источники и Рубрики,
		- оттачивайте Правила попадания в Рубрики,
		- подробнее на ссылке «[**О мониторинге СМИ, соцсетей и сайтов**](https://sitesputnik.ru/Help/I/i_SiteSputnik_News.pdf)»
		- изучайте [другие возможности СайтСпутника](https://sitesputnik.ru/) и добавляйте их в Проект.

\* \* \*

**Часть 2**

**О саморазвивающемся Проекте**

* Можно применить [модуль +NewStreams](https://sitesputnik.ru/Help/I/i_SiteSputnik_NewStreams.pdf), который в автоматическом режиме будет находить новые Источники и проверять: содержат ли они интересующую Вас информацию. Эта проверка осуществляется следующим образом. Все попадающие в поле зрения Проекта ссылки анализируются СайтСпутником на предмет нахождения в их контенте адресов ссылок, являющихся потоками, приносящими информацию в Рубрики Проекта. Проект будет сам наблюдать за изменяющимся информационным пространством и предлагать Вам новые потоки для включения их в Проект. Вам достаточно будет контролировать действия автоматики. Освоение этого модуля также не требует больших усилий.

**Подробнее о двух основных командах**

**Первая команда: «Определение Источников»**

* Имя команды - это URL ссылка, далее могут следовать операнды, например:

http://krestyane34.ru/ +.html /fast

* Имя команды начинается с префикса http:// или https://.
* В Проекте таких команд может быть несколько. Находиться они могут в любом месте Проекта.
* Ещё подробнее о команде - в приведенной ниже Таблице 1.

**Вторая команда: «Определение Рубрики»**

* Это имя рубрики и от одного до трех операндов, например:

Рубрика://Россия/ЮФО/Волгоград/Сельское хозяйство/Весенние работы
Волгоград /7 (озимые | яровые | посевная |сельхоз\*)
КП: C:\Волгоград\Апрель

Искать только в родительской

Здесь к двум операндам, пояснения к которым даны выше, добавлен третий операнд «Искать только в родительской»

* В данном случае третий операнд имеет смысл, если в Проекте задействована родительская Рубрика:

Рубрика://Россия/ЮФО/Волгоград/Сельское хозяйство

В нашем Проекте она не задействована.

* Имя команды начинается с префикса Рубрика:// или rubric://. Префиксы эквивалентны.
* В Проекте таких команд может быть несколько. Находиться они могут в любом месте Проекта.

\* \* \*

**Таблица 1. Основные команды SiteSputnik News**

Красным цветом выделены две основные команды: «Определение источников» и «Определение Рубрик».

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **№** | **Префикс** | **Корень** | **Операнды** | **Пояснения** | **Способ\*** |
| **1** | **http://****https://** | Ссылка конкретная | +L1;L2;LN –m1;m2;mN | L1,L2,LN – лексемы, хотя бы одна из которых содержится в адресе ссылки, попадающей на мониторинг. Лексема может содержать знаки **&&**, например, **news&&tv/** - это означает, что в ссылке должны быть обе лексемы: и **news**, и **tv/**.m1,m2,mN – запрещённые лексемы.  | /fast /slow/ie /edge |
| 2 | file:/// | Имя файла локального | Аналогично пункту № 1, только рассматривается не Ссылка конкретная, а конкретный файл на Вашем компьютере. Это [локальные источники информации](https://sitesputnik.ru/Help/I/i_SiteSputnik_LocSource.pdf).Ограничение временное: файл может быть или html, или текстового типа. В файле должны быть ссылки. |
| 3 | file:///// | Имя файла локальной сети | Аналогично пункту № 2, только рассматривается конкретный файл из локальной сети.Это [источники информации, размещённые в Вашей локальной сети](https://sitesputnik.ru/Help/I/i_SiteSputnik_LocSource.pdf). |
| 4 | folder:/// | Имя папки локальногокомпьютера | Аналогично пункту № 1, только рассматривается не Ссылка конкретная, а конкретная папка на Вашем компьютере. И лексемы ищутся не в адресах ссылок, а в именах файлов, находящихся в папке и всех вложенных в неё папок. Это мониторинг [файлов на Вашем компьютере](https://sitesputnik.ru/Help/I/i_SiteSputnik_Local.pdf). |
| 5 | folder:///// | Имя папки локальной сети | Аналогично пункту № 4, только рассматривается папка из локальной сети.Это мониторинг [файлов из Вашей локальной сети](https://sitesputnik.ru/Help/I/i_SiteSputnik_Local.pdf). |
| **6** | **rubric://****Рубрика://** | Имя рубрики | Правило попадания в Рубрику | Смотрите ниже Приложение 1. | - |
| КП: f1;f2;f3 | f1,f2,f3 – папка, папка локальной сети, E-mail или папка сайта |
| Искать только в родительской | Правила попадания в Рубрику проверяются только для ссылок, попавших в родительскую Рубрику |
| **7** |  | Поиск по запросу | Имя Источника из «Настройка поиска» | Поиск по запросу в поисковики и встроенные в сайты поиски, например: президент России || Yandex=7 || news=3 | /fast /slow/ie /edge |
| **8** |  | В качестве команды может использоваться другой готовый Проект. Подробнее на [ссылке](http://forum.razved.info/index.php?t=2344&p=47342#pp47342) по коричневой стрелке. |

**\*** Параметры /fast и /slow – явная прописка быстрого или обычного способа открытия Источника. Подробнее на [ссылке](http://forum.razved.info/index.php?t=5804).

/ie – браузер Internet Explorer 11, /edge – браузер Edge.

**О макрокомандах**

* Макрокоманды действуют на все команды, заданные в Проекте.
* Они имеют более высокий приоритет, чем установленные в Настройках параметры, определяющие аналогичные функции.
* Запоминать как пишутся макрокоманды не обязательно, все они доступны в диалоге из контекстного меню.

**Таблица 2. Основные макрокоманды**

|  |  |  |  |
| --- | --- | --- | --- |
| **№** | **Название** |  | **Пояснения** |
| 1 | ^^Новости; |  | Проект типа «Новости». Определяются новые ссылки, они скачиваются и отправляются на Рубрикацию. Если в Проекте прописаны Рубрики, то эта макрокоманда может быть опущена. |
| 2 | ^^WebSpider; |  | Проект типа «Обновление». Все ссылки скачиваются, сравниваются с их предыдущим состоянием, выделяются новые фрагменты, которые и отправляются на Рубрикацию. |
| 3 | ^^Рубрики |  | Указывается имя статической (созданной заранее) Рубрики. Подробнее на [ссылке](https://sitesputnik.ru/Help/SSRubriki.htm). |
| 4 | ^^Комментарии; |  | Выделяются комментарии к найденным ссылкам. Они отправляются на Рубрикацию на те же Рубрики, что и сами ссылки. Замечание: выделить комментарии удается не всегда. |
| 5 | ^^Комментарии+Рубрики |  | Аналогично пункту 4, но Рубрикация на «своих» Рубриках. |
| 6 | ^^Сигналы |  | Подключение Сигнализации о значимых событиях. Подробнее на [ссылке](http://forum.razved.info/index.php?t=3579&p=37385#pp37385). |
| 7 | ^^RSS-потоки новые=3^^URL-потоки новые=4 |  | Указывает максимальную глубину [поиска RSS-потоков новых](https://sitesputnik.ru/Help/I/i_SiteSputnik_NewStreams.pdf), несущих искомую информацию. Указывает максимальную глубину [поиска URL-потоков новых](https://sitesputnik.ru/Help/I/i_SiteSputnik_NewStreams.pdf), несущих искомую информацию. |
| 8 | ^^Одинаковый контент+^^Одинаковый контент- |  | Отсеивать новости и сообщения с одинаковым контентом.Анализ на одинаковый контент не выполняется. |
| 9 | ^^Экспорт;^^Импорт; |  | [Экспорт](http://forum.razved.info/index.php?t=2205&p=24812#pp24812) в Клиентские папки Чистого потока.[Импорт](http://forum.razved.info/index.php?t=2205&p=24812#pp24812) из Клиентских папкок Чистых потоков. |
| 10 | ^^Пауза=4 |  | Задает 4-х секундную паузу после окончания обработки очередного элементарного Источника. |
| ^^Пауза.Yandex=3 ^^Пауза.Google=4 ^^Пауза.Yahoo=2 | «Персональные паузы» после открытия страниц Яндекса, Google и Yahoo. |

**Приложение 1. Некоторые Правила попадания в Рубрики на примерах**

* Со всеми Правилами можно ознакомиться на [ссылке](https://sitesputnik.ru/Help/SSRubriki.htm).
* Для начала работы рекомендуется ознакомиться со следующими правилами:
	+ Сергей Иванов - где-то в тексте обязательно есть словоформы от Сергей и Иванов
	Сергей /2 Иванов - между этими словоформами может быть не более 2-х пробелов (не более 1-го слова).
	Сергей & Иванов - эти словоформы должны быть в одном предложении.
	Сергей && Иванов - эти словоформы должны быть в одном абзаце.
	'Сергей Иванов' - строго в тексте должно быть написано то, что Вы заключили в одинарные кавычки.
	(Сергей | Иванов) - хотя бы одно из двух слов в тексте.
	Сергей ~Иванов - Сергей есть, а Иванов нет.
	!стройка - отсечение слева, слово перестройка - не учитывается.
	Алексе\* - справа вместо \* - что угодно, Александр - не учитывается.
	!Путин\* - не учитываются: путь, пути, Распутин.
	(!Путин\* | Медведев | Нарышкин) /4 (приехал | прилетел | прибыл) /3 Волгоград - хотя бы один из трёх посетил Волгоград.
	+ Допускается многоуровневая скобочная структура.
	+ Слова в одной строке Правила (если нет специальных операторов) - логическое "И", то есть, должны быть все в тексте.
	+ Строки соединяются логическим "ИЛИ", то есть, текст учитывается, если удовлетворяет хотя бы одной строке Правила.
	+ Регулярные выражения могут быть прописаны в любом месте Правила.

**Все подробности:**

* на [главной странице сайта о программе](https://sitesputnik.ru/) и
* в [перечне публикаций о программе](https://sitesputnik.ru/Public.htm).

**Обратная связь**

* По всем вопросам можно обращаться на форум СПКР в раздел [о программе SiteSputnik (СайтСпутник)](http://forum.razved.info/index.php?ct=7&f=0);
* к автору программы по Е-mаil или телефонам, указанным на <https://sitesputnik.ru/>

Copyright © 2003-2024 Индивидуальный предприниматель Мыльников Алексей Борисович. Все права защищены.