

SiteSputnik. Локальные источники информации

По смыслу данная публикация дополняет [SiteSputnik Local](#) и наоборот.

Определение

- Локальные источники информации (ЛИИ) – это файлы локального компьютера или локальной сети, в содержании которых находятся ссылки на файлы глобальной сети интернета или на ЛИИ.
- В отличие от файлов-источников глобальной сети, имя которых начинается с префикса [http://](#) или [https://](#), файлы-источники локального компьютера начинаются с префикса [file:///](#), а файлы, размещенные в локальной сети с префикса [file://///](#).
- Это единственное отличие применения ЛИИ от «традиционных» источников, размещенных в интернете. В остальном применение всех трех видов источников практически идентично.
- Начиная с **SiteSputnik Local** версии 9.1.2 и выше, можно в одном Проекте применять все три вида Источников.
- Ограничение. На момент публикации в качестве ЛИИ допускаются только html-файлы или текстовые файлы. Но это задача, при необходимости, будет решена, потому что работа с более широким списком форматов файлов уже реализована на этапе рубрикации собранных материалов.

Как это выглядит на практике

- [file:///d:/WorkFilder/myname.htm](#) – это файл-источник, размещенный на локальном компьютере. Будут обработаны все ссылки на файлы, находящиеся в его контенте.
- [file:///c:/My/file01.txt +.htm; -site](#) – здесь на обработку попадут ссылки на html-файлы, несодержащие в своем имени лексему site.
- [file://///alex-1/ShareDocs/for.html +news](#) – это источник, размещенный в локальной сети. На обработку попадут ссылки на файлы, содержащие лексему news в имени.

Важное пояснение.

- Ничего сложного здесь нет. Если после имени файла через пробел идет [+лексема](#), то из содержащихся в контенте файла ссылок, попадают на обработку только те ссылки, в имени которых есть эта лексема. Если далее через пробел написано: [-лексема](#), то ссылка, её содержащая, игнорируется.
- Подробнее об этом уже пояснялось для источников из глобальной сети - на [ссылке](#).

Как сделать источниками сразу несколько файлов из локальной папки

- Это можно сделать, например, следующим образом:

[folder:///E:/BM/Help +.htm +forum;](#)

Пояснение.

[folder:///E:/BM/Help](#) – это абсолютно все файлы из папки Help и вложенных в неё папок.

[folder:///E:/BM/Help +.htm](#) – это фильтр: только html-файлы останутся в работе

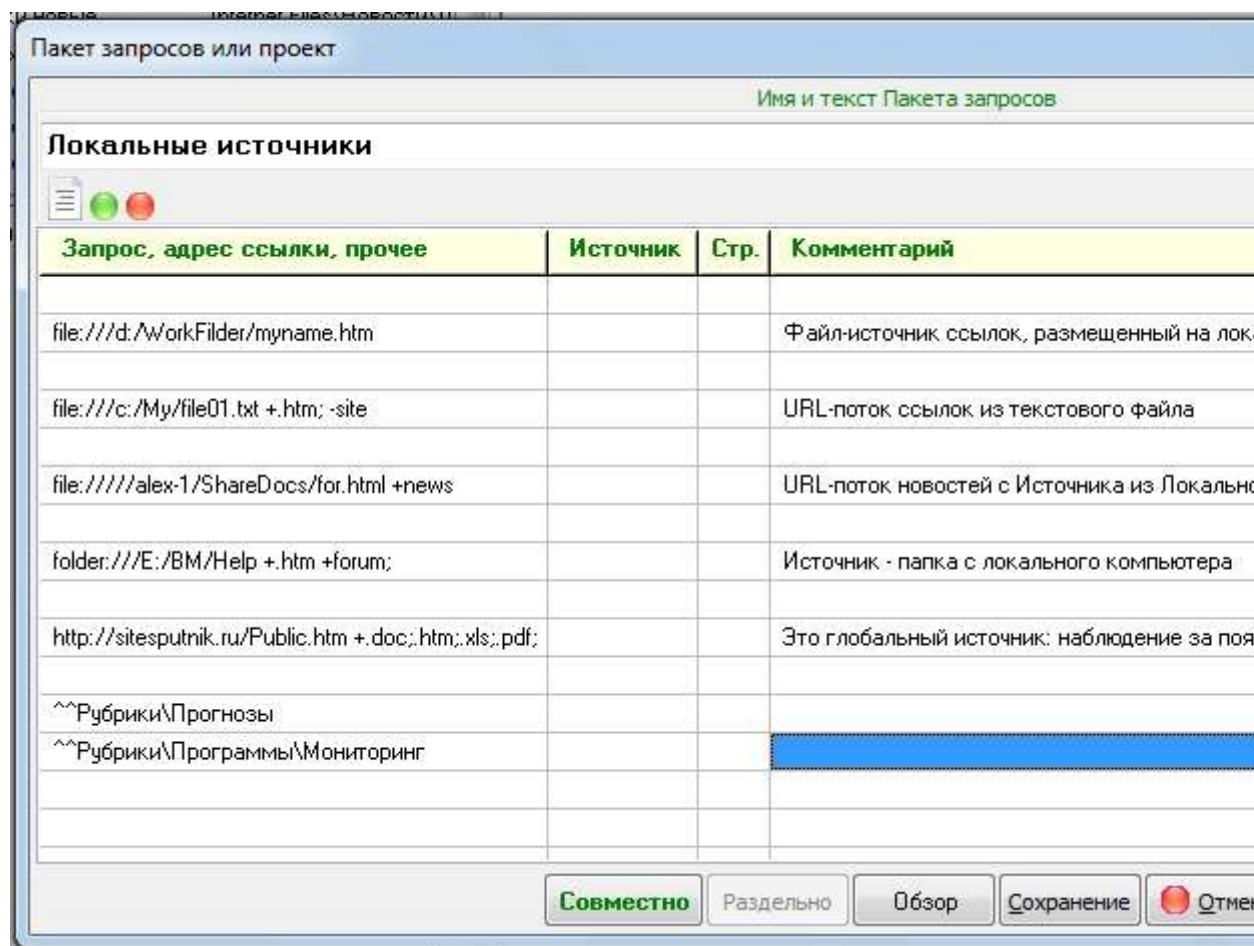
[folder:///E:/BM/Help +.htm +forum](#) – это из каждого оставшегося в работе html-файла на обработку отправляются только ссылки, содержащие лексему forum. Такой процесс может быть продолжен.

Другими словами, это в своем роде «очень специальный» спайдер. Это тоже уже было сделано для глобальной сети, подробнее – на [ссылке](#).

Где это можно применить.

- Для глобальных файлов – это прежде всего подключение не отдельного Источника, а целого списка Источников, находящегося в интернете. Теперь подобный список может быть размещен и на локальном компьютере или в локальной сети.
- Не важно каким образом были сделаны изменения в источнике: руками, другим программным комплексом или самим СайтСпутников, важно каким типом Проекта он наблюдается. Если это Проект типа News, то на обработку попадут только новые ссылки, относительно всех предыдущих обращений к Источнику, если это тип +WebSpider, то все ссылки будут проанализированы по группе критериев на предмет их обновления относительно последнего их состояния.

Пример проекта с упомянутыми источниками:



- В него добавлен для разнообразия глобальный источник:
 - <http://sitesputnik.ru/Public.htm+.doc;.htm;.xls;.pdf;> – это наблюдение за появлением новых публикаций о СайтСпутнике на сайте автора,
 - и анализ (рубрикация) материалов из всех Источников по правилам, прописанным в двух Рубриках и вложенных в них Рубрик.
- Любой или все источники из этого Проекта могут быть добавлены в Проект из этой [публикации](#) и наоборот.
- Материалы из данной публикации относятся к варианту комплектации **SiteSputnik Pro+News+Local**.

Все подробности:

- на [главной странице сайта о программе](#) и
- в [перечне публикаций о программе](#).

Замечание от 12.04.16.

- О том, как Рубрики можно прописать прямо в теле Проекта, написано на ссылке «[Динамические Рубрики](#)».

Обратная связь

- По всем вопросам можно обращаться на форум СПКР в раздел [о программе SiteSputnik \(СайтСпутник\)](#);
- к автору программы по E-mail или телефонам, указанным на <http://sitesputnik.ru/>

[Copyright © 2003-2019 Индивидуальный предприниматель Мильников Алексей Борисович. Все права защищены.](#)