Автор: Алексей Мыльников
https://sitesputnik.ru/
Дата выхода публикации: 03.09.2015
Дата последнего обновления: 07.10.2022

Инструкция по применению программы FileForFiles & SiteSputnik Pro

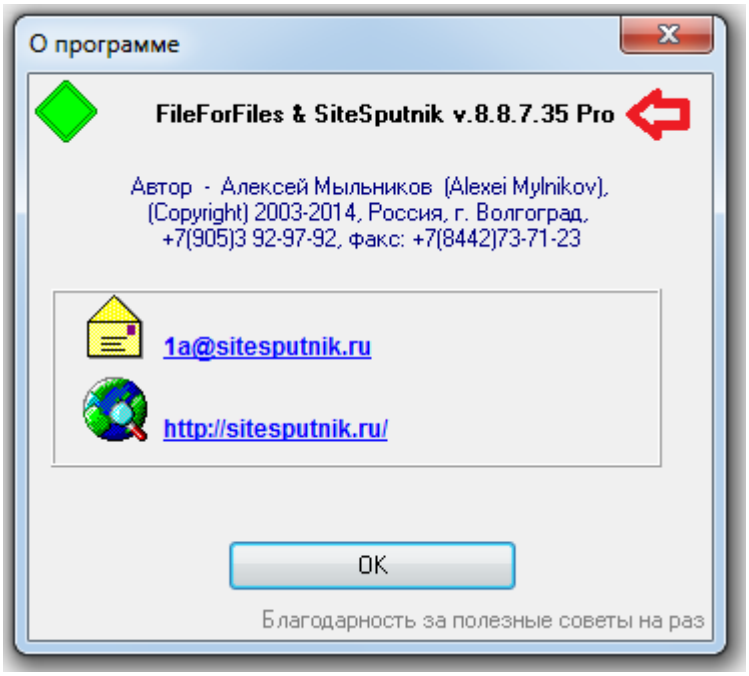
для *начинающих* Пользователей

Назначение программы

- Организация и автоматизация профессионального поиска и сбора информации из открытых источников интернета, мониторинг появления новых ссылок на заданную тему. Программа постоянно совершенствуется в направлении: «Найти всё и гарантированно». «Программа для допроса интернета», - ещё одно определение Пользователя о назначении этого модуля программы.

Установка программы

- Удалите демоверсию программы **FileForFiles & SiteSputnik** (сокращенно **SiteSputnik** или **СайтСпутник**), установленную на Вашем компьютере. Для этого достаточно прямо из Проводника удалить папку FileForFiles.
- Папку FileForFiles из персонального дистрибутива программы разместите, например, прямо в корне диска D:\.
- Откройте файл FileForFiles.exe и ответьте на два вопроса. Программа установится и откроется.
- Для того, чтобы убедиться, что Вы имеете дело с вариантом **SiteSputnik Pro**, выберите из главного меню пункт: **? – О программе**. Отрывшееся окно должно иметь вид:



Обновление программы и переход на новые версии

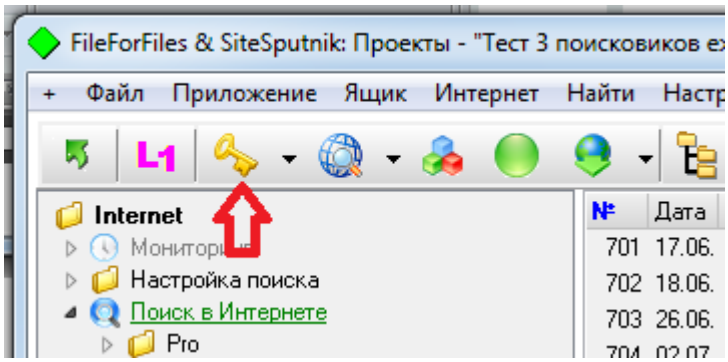
- Рекомендуется систематически обновляться: **? – Проверить наличие обновлений**.
- Обновления и переходы на новые версии программы бесплатные.

Несколько видеоуроков

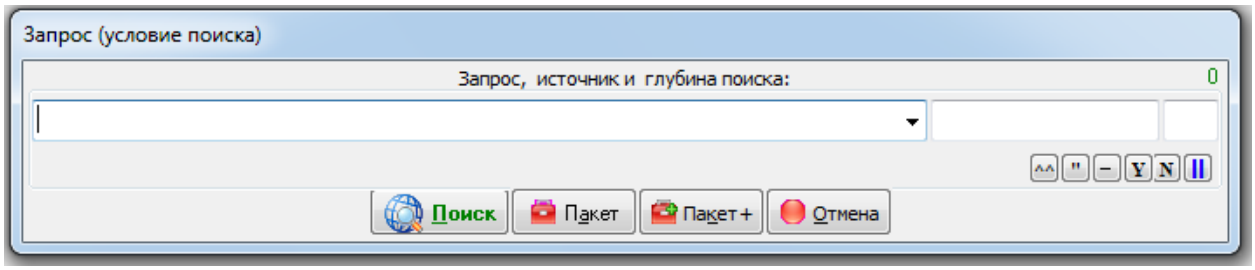
1. Сбор информации в интернете, простой пакет запросов - урок 02
2. Аналитическое объединение результатов поиска и сбора - урок 03
3. Настройки для списка "Ссылки новые" - урок 05
4. Поиск информации в интернете: очень простой запрос - урок 01
5. ~~Поиск информации в интернете: сложные запросы~~ - урок 13 **(не относится к Pro)**

Простой запрос

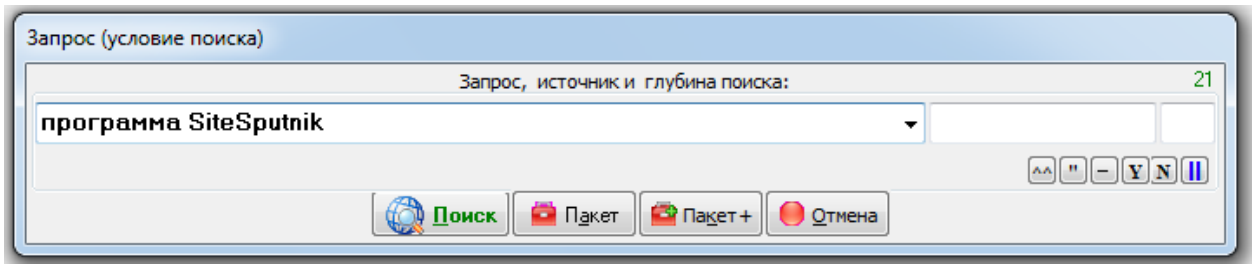
- На верхней панели инструментов нажмите на иконку «Золотой ключик»:



- В открывшемся окне:

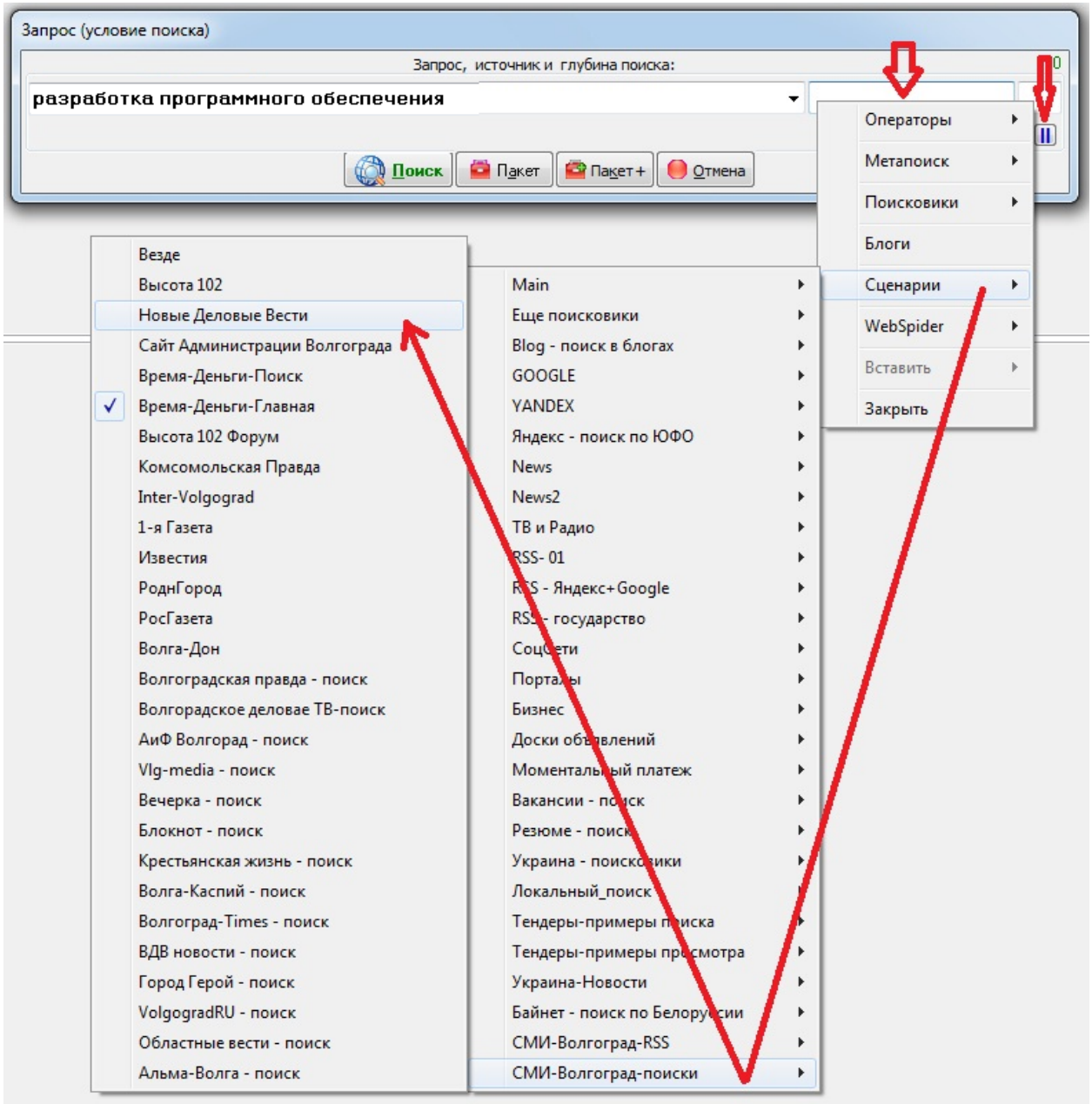


- Введите текст запроса, например, *программа SiteSputnik,* или свой запрос:

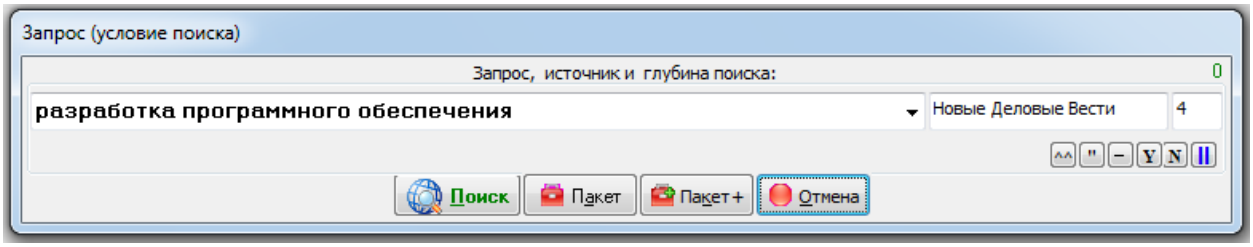


- Если Вы, больше ничего не делая, сразу нажмёте на кнопку «**Поиск**», то по умолчанию запрос выполнится в Яндексе, Google, Yahoo, Рамблере, Bing (MSN), Mail, в комментариях и блогах Яндекса и в блогах в Google на глубину трёх страниц поиска для каждого Источника. Дубли ссылок будут отброшены, выдача будет единой, как будто Вы задействовали не 9 Источников, а один. Попробуйте выполнить этот или другой поиск и ознакомиться со списком «Объединение», содержащим единую выдачу. О назначении остальных десяти документов, созданных **программой SiteSputnik** при выполнении этого запроса, можно прочитать в отдельной публикации: SiteSputnik. Документы, файлы, списки, таблицы и другая выдача программы.

- Если Вам нужно найти информацию в других Источниках, то наведите курсов на кнопку ||. Это приведет к открытию меню доступных на данный момент Источников. Выберите Источник, например, «Новые деловые вести»:



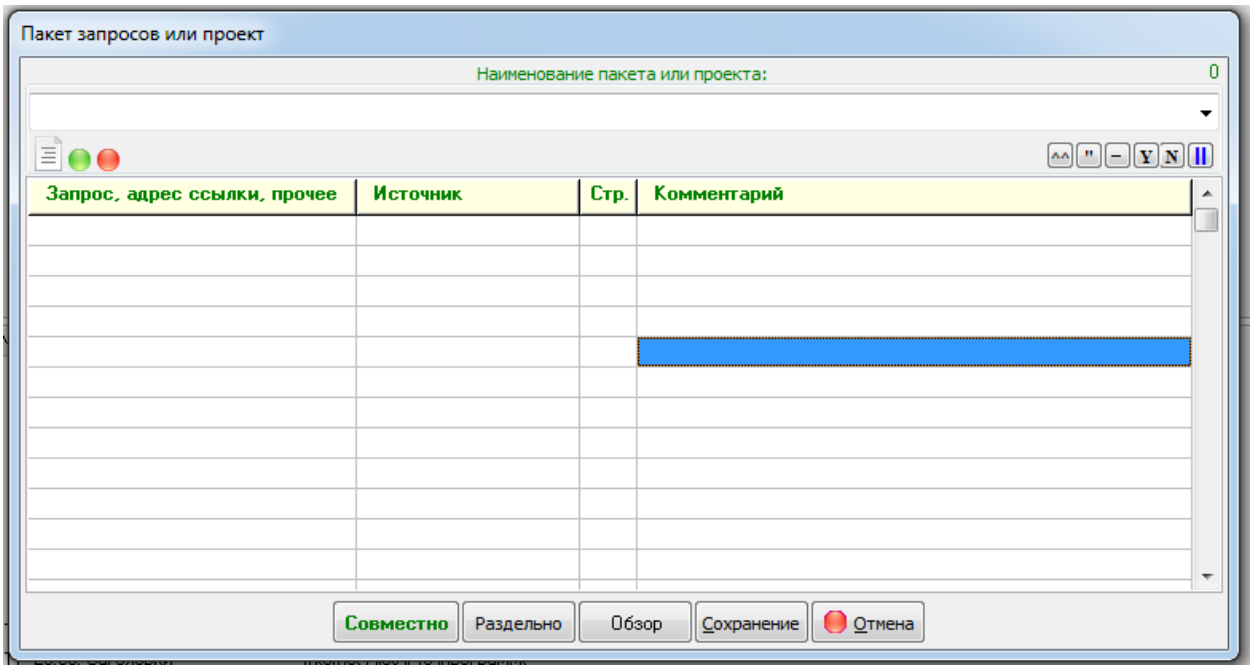
- После выбора Источника Вы увидите следующее:



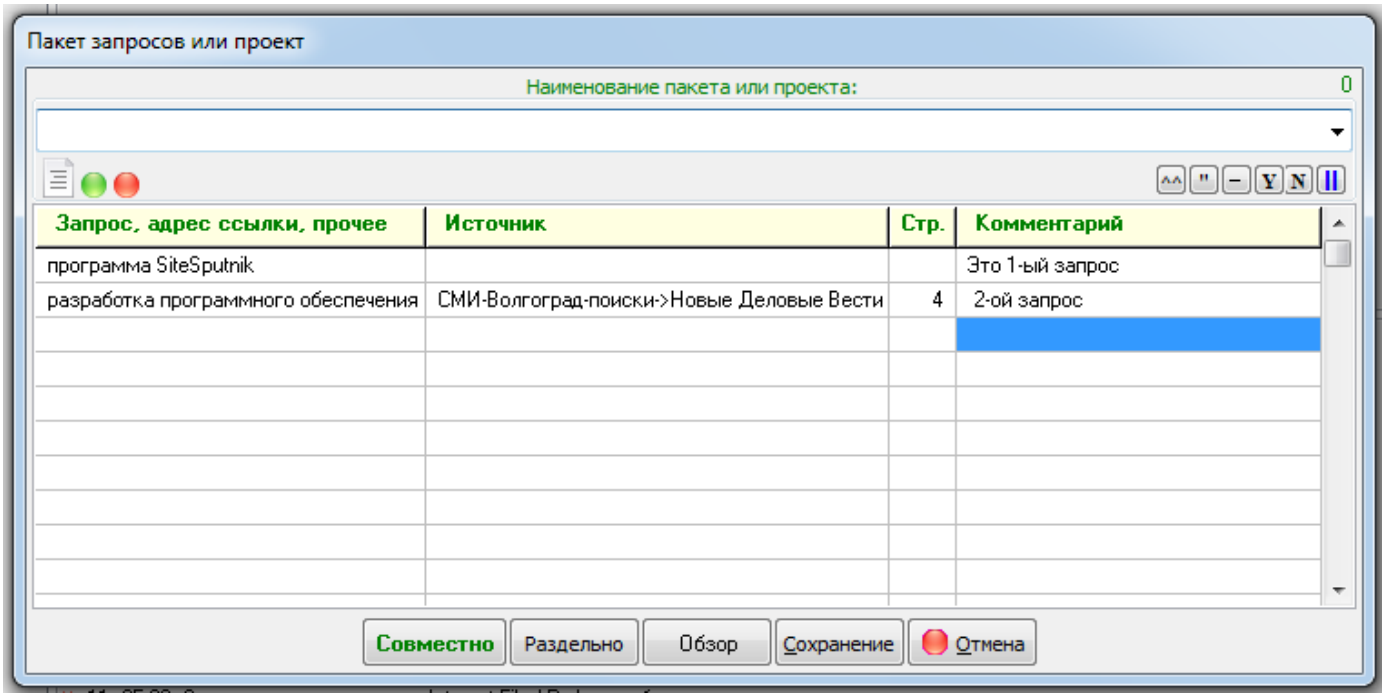
- Глубину поиска, размещенную в третьем поле, можно изменить вручную. Глубина – это количество поисковых страниц, автоматически открываемых СайтСпутником. После нажатия на кнопку «Поиск», заданный запрос выполнится по встроенному в «Новые Деловые Вести» внутреннему поиску, который, в подавляющем большинстве случаев, более полный, чем поиск по этому ресурсу через поисковик, например, Яндекс или Google.

Пакеты запросов !

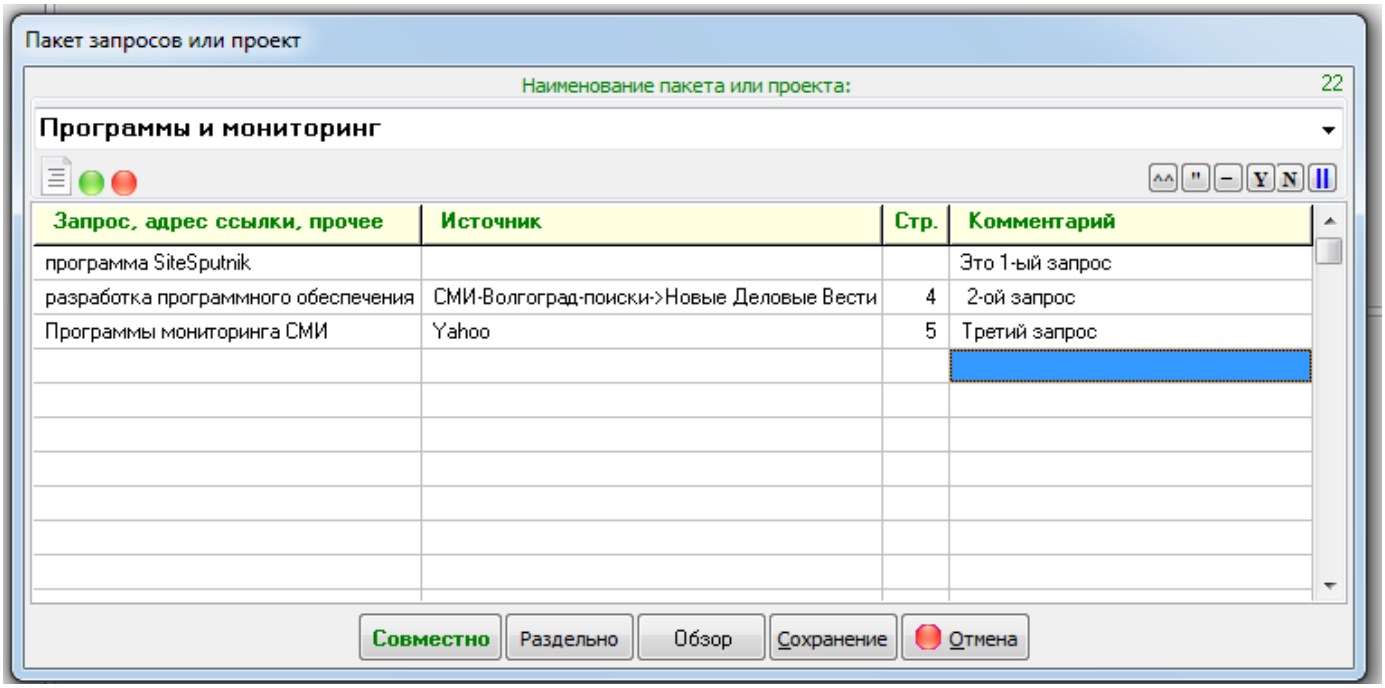
- Если после нажатия на иконку «Золотой ключик», Вы нажмёте на кнопку «Пакет», то увидите следующее:



В каждой строке таблички, состоящей из четырёх столбцов, можно прописать: запрос, Источник, на котором он выполняется, глубину поиска в Источнике в страницах и в четвёртом столбце произвольный комментарий. Например, если Вы возьмете предыдущие два запроса, то после добавления их в Пакет получите следующее:



Далее, если добавите в Пакет третий запрос, а именно: «*Программы мониторинга СМИ*», выполняемый на Yahoo на глубину 5 страниц, и назовёте Пакет «*Программы и мониторинг*», то получите:



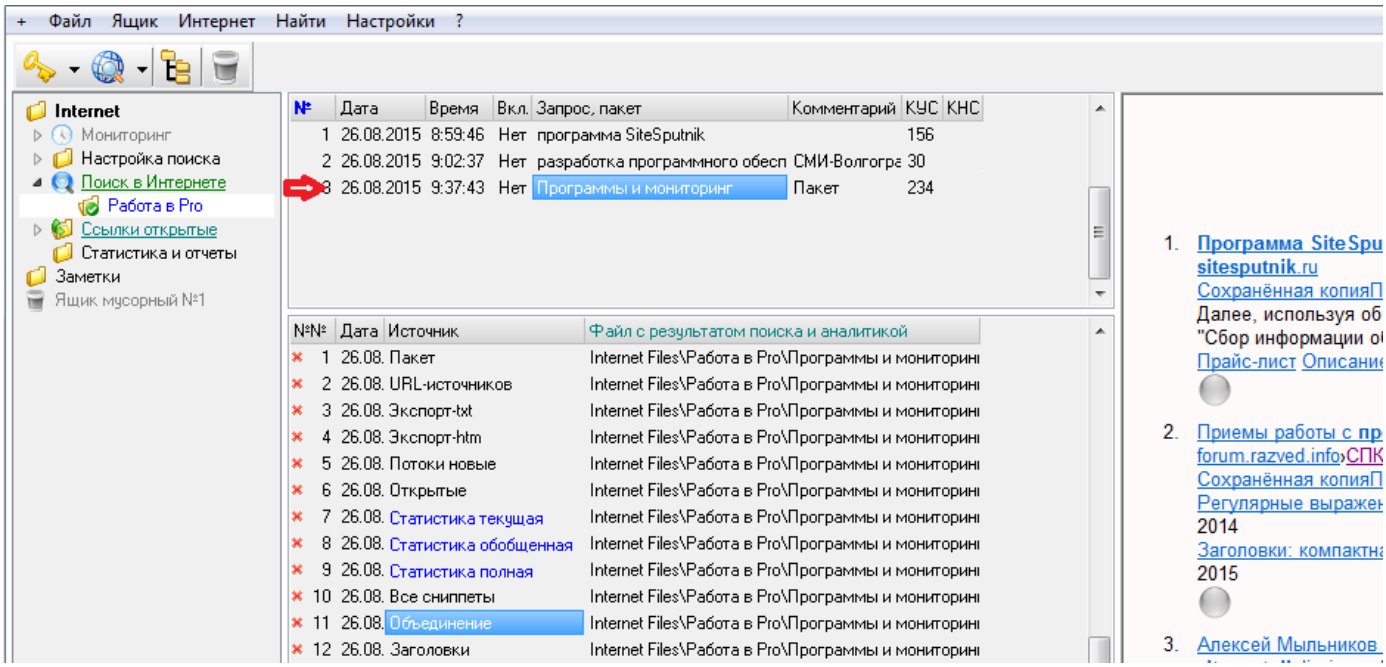
Далее можно добавить в Пакет столько запросов, сколько нужно. Закончив формирование Пакета, нажмите на кнопку «Сохранение», чтобы сохранить его в **библиотеке Пакетов запросов**. Выполнить Пакет можно по кнопке «Совместно» или «Раздельно».

Назначение кнопок таково.

- «Совместно» - все запросы работают на один объединенный результат.
- «Раздельно» - каждый запрос выполняется отдельно.
- «Обзор» - выбор Пакета из библиотеки Пакетов запросов.
- «Сохранение» - сохранение пакета в этой библиотеке.

Автоматическое архивирование, базы данных

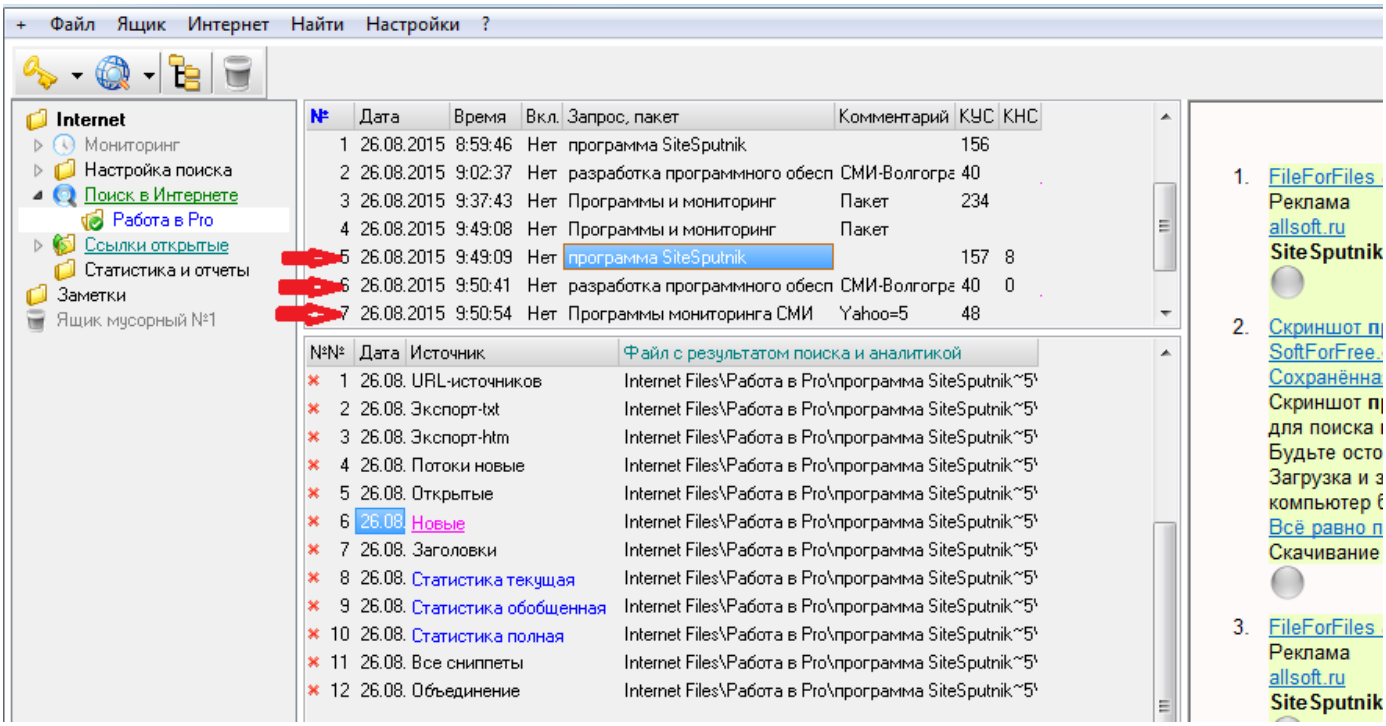
- Все Ваши действия и все результаты работы программы SiteSputnik запоминаются в библиотеках и во встроенных базах данных. Если Вы выполните предыдущий Пакет по кнопке «Совместно», то ниже на картинке в третьей строке запомнится вся информация, добывая СайтСпутником по Пакету:



Примечание:

в первых двух строках сохранились результаты самостоятельного выполнения двух простых запросов, описанных вначале данной инструкции в разделе **Простой запрос**.

- Если Вы прямо сейчас выберите ваш Пакет из библиотеки: **Золотой ключик – Пакет – Обзор – Программы и мониторинг**, и выполните его по кнопке «Раздельно», то увидите:

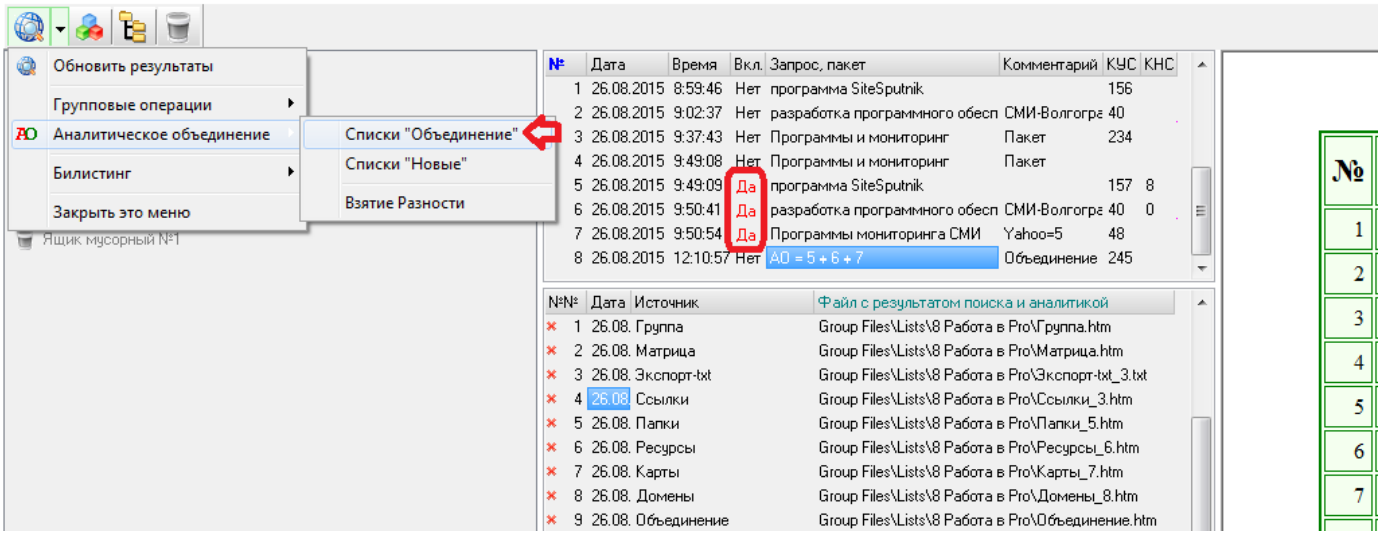


что в пятой, шестой и седьмой строках сохранились результаты работы каждого из трёх запросов из Пакета.

Далее в пунктах Аналитическое объединение и Список «Новые» приведена важная и уникальная информация.

(!) Аналитическое объединение (АО)

- Да, выполнили Вы в предыдущем пункте запросы «Раздельно». Да, бывает, что это удобно и полезно. Но, главное, что можно результаты выполнения отдельных запросов и **пакетов запросов** объединить в один результат, причем не просто объединить, отбросив дубли найденных ссылок, а выполнить упорядочивание выдачи по следующему признаку: чем большим количеством запросов была найдена ссылка, тем ближе к началу выдачи она размещена. Во-первых, это повышает релевантность выдачи. Во-вторых, это позволяет «вычислять» страницы, ресурсы и домены, удовлетворяющие нескольким запросам или пакетам. Частный случай применения АО – это установление связи объектов, точнее их упоминания на одних страницах и ресурсах.
- Для того чтобы применить операцию АО вручную (а её можно и прописывать в тексте Пакета), сделайте так как показано на картинке ниже: установите значение «Да» в столбце «Вкл» и выберите пункт: **Глобус с линзой – Аналитическое объединение – Списки “Объединение”**

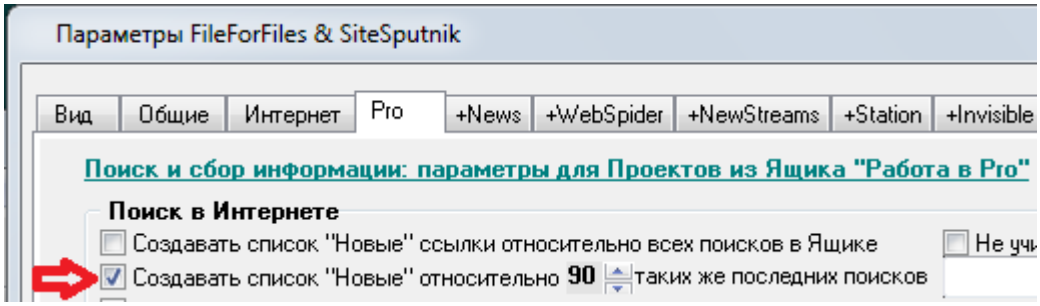


Объединенный и упорядоченный по релевантности результат появится в восьмой строке.

- Есть ещё операция взятия разности, когда из ссылок одного запроса или пакета, удаляются ссылки другого.

(!) Список Ссылки «Новые»

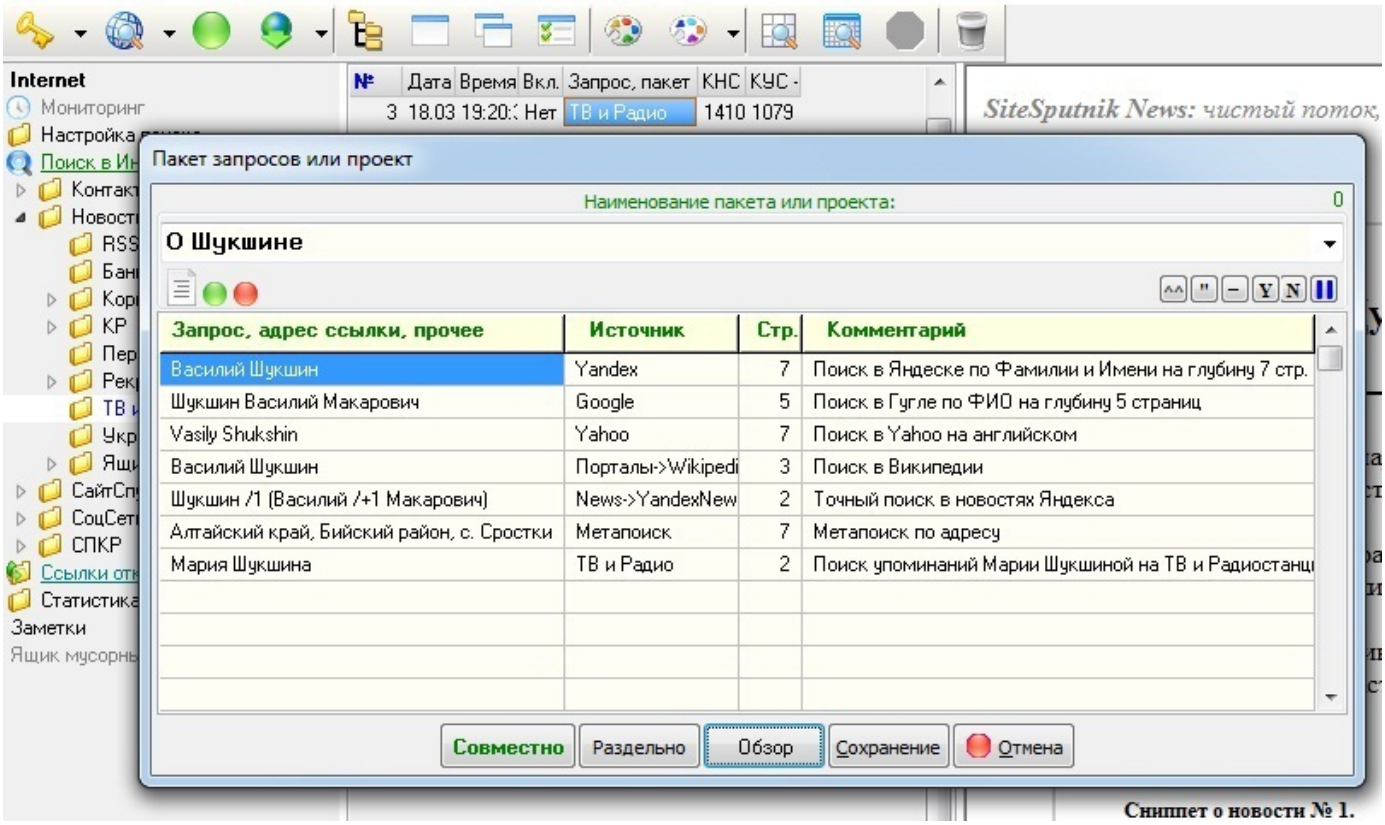
- Этот список – очень важная составляющая выдачи программы СайтСпутник. Обратите внимание на 1-ю и 5-ю строки на картинке выше. В них с интервалом 50 минут (дата и время выполнения указаны во 2 и 3 столбцах) сохранились результаты выполнения одного и того же запроса а именно: «Программа SiteSputnik». Для 1-ой строки в столбце КУС (количество универсальных ссылок) стоит 156, а в столбце КНС (количество новых ссылок) пусто. Для 5-ой строки КУС = 157, а КНС = 8. Последнее означает, что второе выполнение запроса принесло 8 новых ссылок относительно первого его выполнения. В общем случае, правильнее было бы сказать не новых, а не найденных ранее ссылок. Если бы запрос выполнялся в 11 раз, то список «Новые» содержал бы ссылки не найденные ни в одном из его предыдущих 10-ти выполнений. Если бы примененные Источники выдавали информацию, упорядоченную по дате, то это бы были действительно новые ссылки. В любом случае, довольно часто при наблюдении за объектом достаточно просматривать только списки «Новые».
- На картинке ниже показана настройка, прописывающая поиск новых ссылок только для одинаковых поисков. Под одинаковыми поисками понимаются одни и те же запросы или одноименные Пакеты запросов:



Если Вам нужно искать новые ссылки для разных запросов или Пакетов, то установите галочку у параметра выше. Простой случай, когда Вам это может понадобиться: Вы ищете новую информацию по физическому лицу, задавая разные запросы: ФИО, номера телефонов, ИНН, е-майлы, какие-нибудь другие реквизиты, в том числе, появляющиеся при анализе только что полученных результатов. В этом случае Вам, напрмер, не придется вторично просматривать ссылку, найденную по е-майлу, если она ранее была найдена по номеру телефона. Эта ссылка попадет в «Объединение», но не попадет в «Новые».

Подобный поиск новых ссылок будем называть сокращенно **Мониторингом поиска**. Он входит в состав самой минимальной комплектации программы, а именно в **SiteSputnik Pro**. Никаких скачиваний и никакого анализа ссылок при мониторинге поиска не производится.

- Ещё пример пакета запросов, взятый с главной страницы сайта www.sitesputnik.ru:



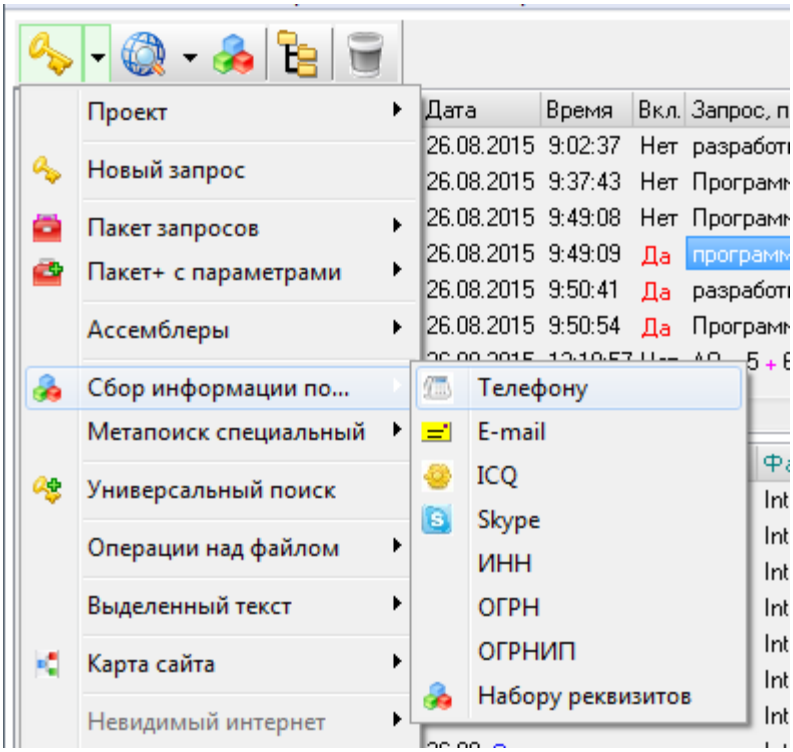
Он состоит из семи разных запросов, собирающих информацию о Василие Шукшине из нескольких Источников, среди которых есть поисковики, Википедия, точный поиск в новостях Яндекса, метапоиск и поиск упоминаний на ТВ и радиостанциях. В сценарий ТВ и Радио входят: "Первый канал", "ТВ Россия", НТВ, РБК ТВ, "Эхо Москвы", радиокомпания "Маяк", ... и другие Источники информации. Для каждого Источника прописана своя глубина поиска или просмотра в страницах. Она указана в третьем столбце.

Спецзапросы

На верхней панели инструментов нажмите на иконку: **Золотой ключик – Сбор информации по… -**

- **Телефону**
- **E-mail**
- **ICQ**
- **Skype**
- **ИНН**
- **ОГРН**
- **ОГРНИП**
- **Набору реквизитов**

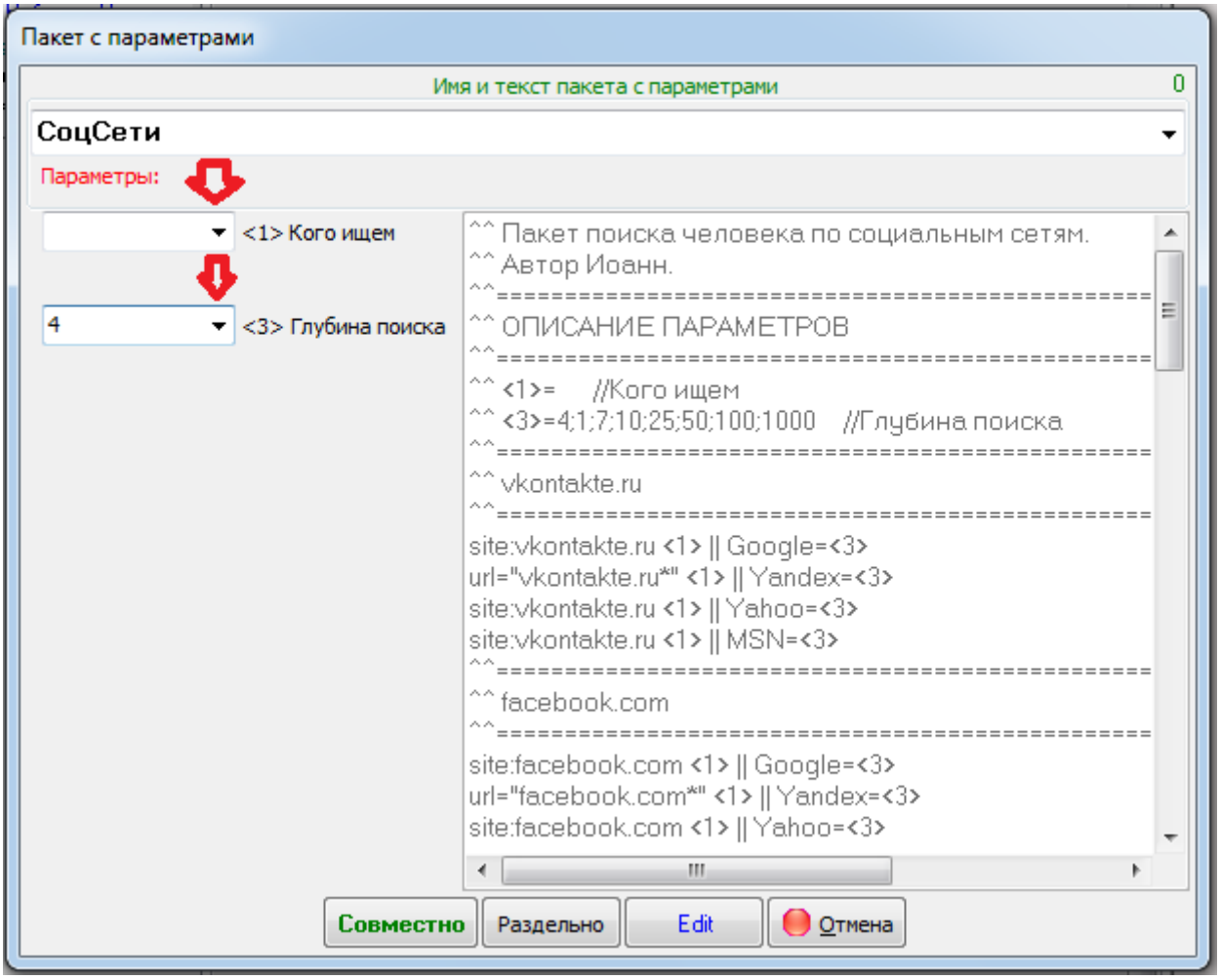
Смотрите на картинку:



Это обращение к готовым, встроенным в программу СайтСпутник поискам, которые очень тщательно «перероют» открытые источники интернета в поисках реальной, а не официальной информации, «подвешенной» на соответствующий реквизит. Вам не придется обращаться к разным поисковикам, задавать разные формы записи реквизитов, находить и просматривать по несколько раз одни и те же ссылки. «Мусорных» ссылок также будет на порядок меньше.

Пакеты с параметрами

- Предельно просто об этом написано в инструкции.
- Любые пакеты запросов, предназначенные для решения стандартных поисковых задач, могут быть параметризованы, сохранены в **Библиотеке пакетов с параметрами** и выполнены с подстановкой фактических (нужных) значений параметров.
- Чтобы выбрать готовый пакет с параметрами из библиотеки, нажмите на иконку: **Золотой ключик - Пакет+ - Имя пакета - Открыть**. Посмотреть как это делается для поиска разных вариантов записи e-mail можно на видео.
- На картинке ниже показано, что увидит Пользователь, если откроет Пакет с параметрами с именем «СоцСети». Смотрите на левую часть картинки, где вводятся два параметра:



1) Кого ищем, например, Алексей Мыльников.
2) На какую глубину в страницах ищем, например, 4

Справа размещен текст пакета, в котором прописан поиск в соцсетях: ВКонтакте, FaceBook, Одноклассники, МирТесен, МойКруг, МойМир, Привет.Ру, habrahabr.ru, ruspace.ru, ЖдиМеня, Профессионалы.Ру, с использованием четырех поисковиков: Google, Yandex, Yahoo, Bing (MSN), индексирующих эти сети. Пакет можно редактировать по кнопке «**Edit**».

Чтобы выполнить аналогичный поиск для другого человека достаточно ввести его ФИО. При необходимости в пакет можно добавить поиски, встроенные в сами социальные сети.

- В папке «Петровские», доступной по: **Золотой ключик - Пакет+ - Петровские - … - Открыть**, находится много пакетов с параметрами, разработанных Олегом Петровым из Санкт-Петербурга.
- Начинающим Пользователям не обязательно перегружать себя информацией и прямо сейчас осваивать технологию создания пакетов с параметрами. Это можно сделать позднее – смотрите специальную инструкцию, отдельную публикацию и обсуждение на форуме.
- По сути, пакеты с параметрами – это самостоятельные созданные формы расширенного поиска.

Ассемблеры

- Предназначены для сборки сложного задания из готовых: простых запросов, пакетов запросов и пакетов с параметрами.
- Начинающему Пользователю рекомендую приступить к освоению Ассемблеров позднее, обратившись к инструкции, к обсуждению на форуме или видео, где продемонстрирована сборка. Необходимость в Ассемблере может возникнут позднее.

Запрос+Пост-запрос (PostQuery)

- Расстояние между ключевыми словами в запросе, нахождение их в одном предложении или абзаце, скобочная структура запроса для Яндекса, Google, встроенных в сайты поисков и любых других подключенных источников, - об этом в новой публикации SiteSputnik PostQuery. Не входит в модуль **Pro**.

Мастер подключения источников информации

- Мастер как раз и предназначен для подключения встроенных в сайты внутренних поисков, таких как рассмотренный в предыдущем пункте. При помощи него так же можно подключить имеющиеся на сайтах RSS-потоки или создать (эмулировать) свои RSS-потоки, что очень востребовано в **SiteSputnik News**. О Мастере – в отдельной публикации. С одной стороны, сказать что применение Мастера является простым делом было бы неправильно, с другой стороны, он вполне доступен для «среднего» Пользователя, и, главное, позволяет применить практически любые Источники без помощи программистов. Это отдельный «элемент SiteSputnik Pro», и, если Вам не хватит включенных в персональный дистрибутив встроенных Источников, то Вы можете самостоятельно применить Мастер или обратиться за помощью на форум СПКР в раздел Библиотека источников и инструменты её создания и ведения. Источники, подключенные другими Пользователями, могут быть импортированы в программу на Вашем компьютере по нажатию одной кнопки.

Разное

- Так же в **SiteSputnik Pro** доступны следующие возможности:
  - Метапоиск, метапоиск по сайту, метапоиск в офисных документах, метапоиск кеш-копий ссылок;
  - Глубокий поиск;
  - Универсальный поиск;
  - Построение карт сайтов, ресурсов, веб-папок, искомых объектов на основе найденных в открытых источниках интернета принадлежащих им ссылок;
  - Подробнее:
    - на главной странице сайта о программе и
    - в перечне публикаций о программе.
- Например, карту сайта https://sitesputnik.ru/, построенную **программой СайтСпутник** без единого обращения к самому сайту, можно увидеть здесь. Обращение к функции построения карты сайта: **подменю Золотого ключика – Карта сайта – Стандартная**. Для начала можете построить карту своего корпоративного или личного сайта.
- Важное о карте сайта.
  - Если под невидимым интернетом понимать непроиндексированные основными поисковиками ссылки, то карту сайта можно рассматривать как материализацию видимого интернета. Именно она является базой, относительно которой и с использованием которой производится поиск невидимых ссылок в **SiteSputnik Invisible**.

Инструкции для *начинающих* по другим вариантам комплектации программы

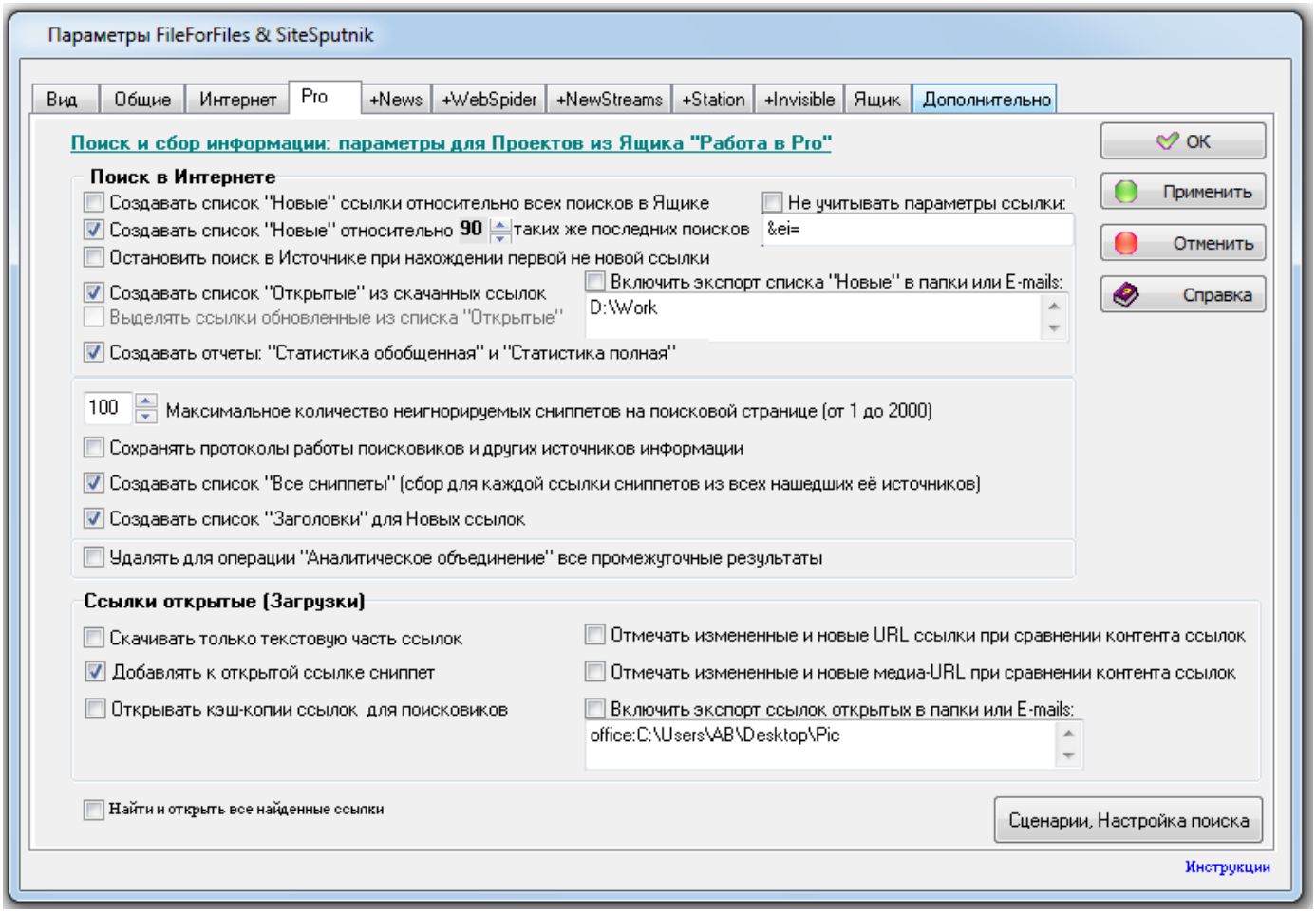
Вариант комплектации SiteSputnik Pro является базовым. Он входит во все другие варианты комплектации программы и делает их более мощными. Это одно из отличий продвинутых вариантов программы SiteSputnik от похожих программ. В настоящее время существуют следующие варианты программы:

- +News – мониторинг **новостных потоков**: СМИ, соцсетей и других сайтов,
  - инструкция для начинающих Пользователей в стадии написания.
- +WebSpider - мониторинг удовлетворяющих критерию **обновлений источников**.
  - инструкция для начинающих Пользователей в стадии написания.
- +Comments – мониторинг комментариев к новостям, сообщениям, публикациям
  - инструкция для начинающих Пользователей в стадии написания.

- +Station – мониторинг с использованием нескольких компьютеров.
    - инструкция для начинающих Пользователей в стадии написания.
- +NewStreams (Бета версия), **поиск Источников**, отслеживание их появления.
    - инструкция для начинающих Пользователей в стадии написания.
- +Invisible - базовые функции поиска информации в невидимом интернете.
    - инструкция для начинающих Пользователей в стадии написания.
- +Objects - объекты: физические и юридические лица - сбор информации и связи.
    - инструкция для начинающих Пользователей в стадии написания.
- +Contacts - поиск контактной информации.
- Corp - ориентирован на корпоративных клиентов, функционально подобен Pro.

Основные настройки для SiteSputnik Pro

- Они доступны из главного меню: **Настройки – Параметры – Вкладка Pro**:



- Создавать список “Новые” относительно всех поисков в Ящике:
    - Найденные в текущем задании ссылки проверяются на предмет нахождения их в любом другом задании, выполненном ранее в текущем Ящике. В список “Новые” попадают только те ссылки, которые встретились впервые.
- Создавать список “Новые” относительно NN таких же последних поисков:
    - Найденные в текущем задании ссылки проверяются на предмет нахождения их в NN таких же предыдущих заданиях, выполненных ранее в текущем Ящике. В список “Новые” попадают только те ссылки, которые встретились впервые. Под «такими же заданиями» понимаются один и тот же запрос или одноименные Пакеты запросов.
- Не учитывать параметры ссылки:
    - Чекбокс включает или отключает учитывание при определении новизны ссылки перечисленные через знак «точка с запятой» параметры в адресе ссылки.

- Остановить поиск в Источнике при нахождении первой не новой ссылки:
    - Как только в очередном источнике попадается ранее найденная ссылка, так сразу осуществляется переход к работе со следующим Источником. Параметр требует особого внимания при применении. Он бывает полезен, если все Источники в задании выдают информацию в обратном хронологическом порядке, а сам поиск происходит на большую глубину.
- Всключить экспорт списка ”Новые” в папки или E-mails:
    - Чекбокс включает или отключает отправку списка ”Новые” в перечисленные через знак «точка с запятой» папки локального компьютера, или локальной сети, или e-mail адреса.
- Создать список ”Открытые” из скачанных ссылок:
    - Найденные в текущем задании ссылки, которые были ранее открыты или скачаны в программе, попадают в этот список.
- Создать отчеты ”Статистика обобщенная” и ”Статистика полная”:
    - ”Статистика текущая” создается всегда. Она содержит информацию об эффективности работы Источников для данного выполнения задания.
    - ”Статистика обобщенная” – это «усредненные» результатов из ”Статистик текущих” для всех выполнений текущего задания.
    - ”Статистика полная” содержит «усредненные» результаты для всех разных заданий, выполненных в текущем Ящике.
- Максимальное количество неигнорируемых сниппетов на поисковой странице:
    - Если на странице найдено больше ссылок, чем в этом ограничителе, то они (сниппеты, содержащие эти ссылки) не попадут в выдачу.
- Сохранять протоколы работы поисковиков и других источников информации:
    - Можно сохранить в выдаче СайтСпутника каждую открываемую страницу для всех задействованных в задании источников.
- Создать список «Все сниппеты» (сбор для каждой ссылки сниппетов из всех нашедших её источников)
    - Создается список, в котором для каждой ссылки формируется большой сниппет из всех сниппетов из нашедших ссылку источников.
- Создать список «Заголовки» для Новых ссылок:
    - Этот список содержит только заголовки страниц, попавших в список ”Новые”.
- Удалять для операции “Аналитическое объединение” все промежуточные результаты:
    - В выдаче останется только результат операции АО, сами операнды будут удалены.
- Скачивать только текстовую часть ссылок
    - Напрямую скачивается файл ссылки, а не выполняется открытие ссылки встроенным браузером.
- Всключить экспорт ссылок открытых в папки или E-mails:
    - Чекбокс включает или отключает отправку открываемых или скачиваемых ссылок в перечисленные через знак «точка с запятой» папки локального компьютера, или локальной сети, или e-mail адреса. Префикс ***office:*** перед папкой или e-mail означает, что осуществляется отправка только офисных файлов.

Все подробности:

- на главной странице сайта о программе и
- в перечне публикаций о программе.

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 https://sitesputnik.ru/

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