

Поисковая система  
*Avalanche 2.5*



Начало работы

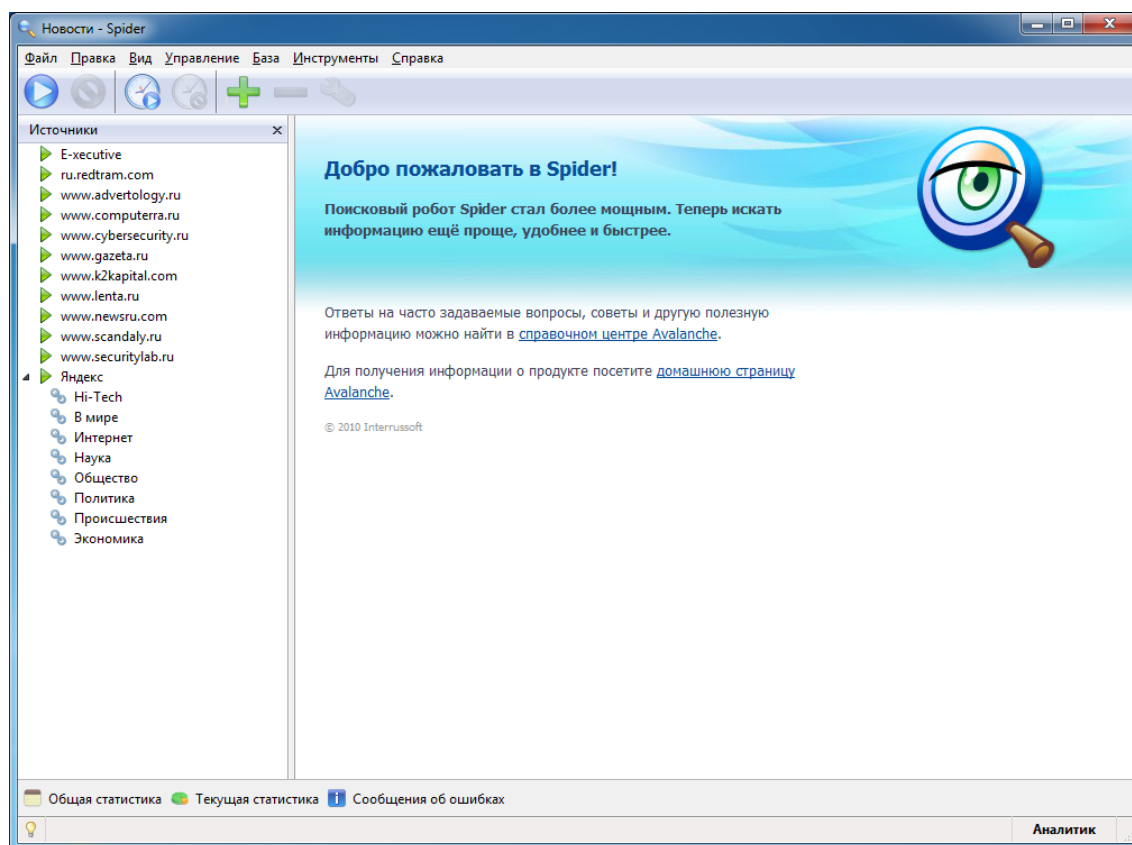
## Содержание

<b>1</b>	<b>Первый запуск</b>	<b>2</b>
<b>2</b>	<b>Просмотр результатов загрузки</b>	<b>5</b>
<b>3</b>	<b>Настройка новых источников</b>	<b>7</b>
<b>4</b>	<b>Помощник по тонкой настройке</b>	<b>12</b>
<b>5</b>	<b>Настройка новых рубрик</b>	<b>18</b>

# 1 Первый запуск

После установки на рабочем столе появятся две новые иконки Spider и Avalanche. Первая Spider - это поисковый паук, который загружает документы из ваших источников. Вторая Avalanche - это менеджер документов, благодаря этой программе вы сможете просматривать результаты загрузки в удобном оконном интерфейсе.

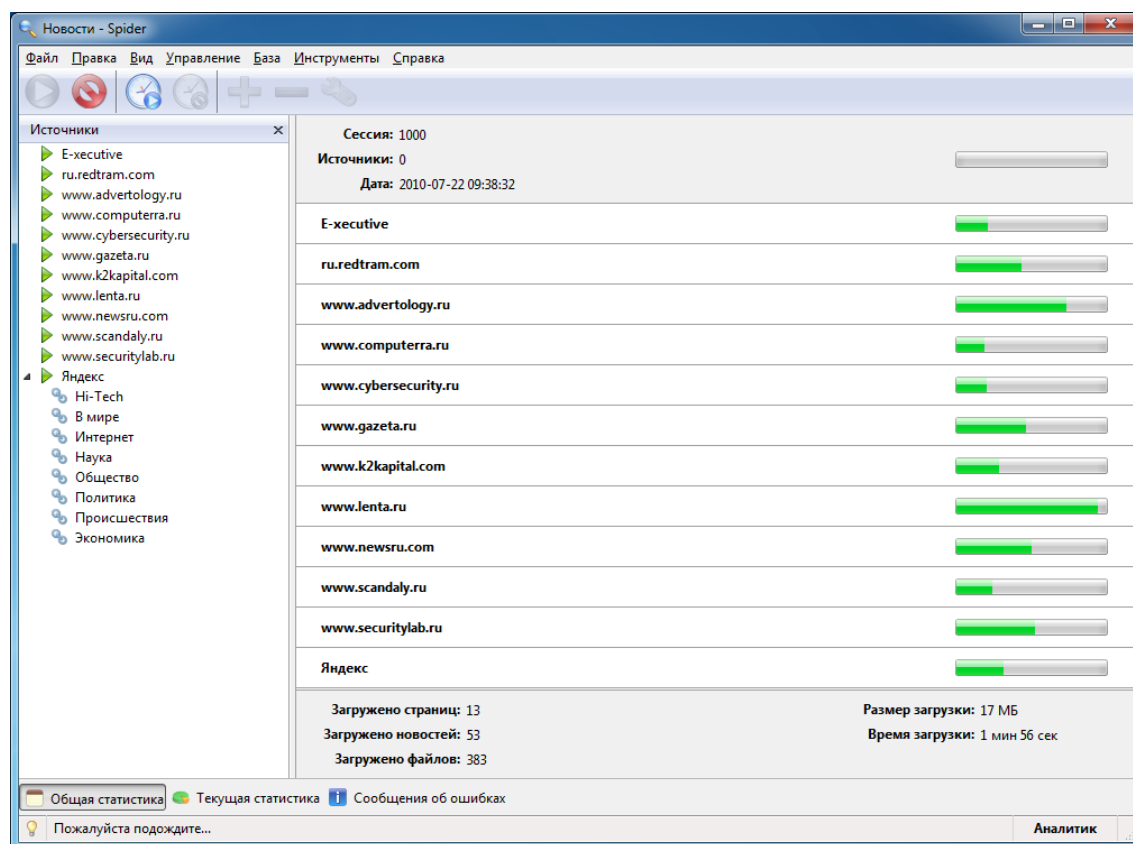
Итак в результате первого запуска Spider вы увидите следующий экран.



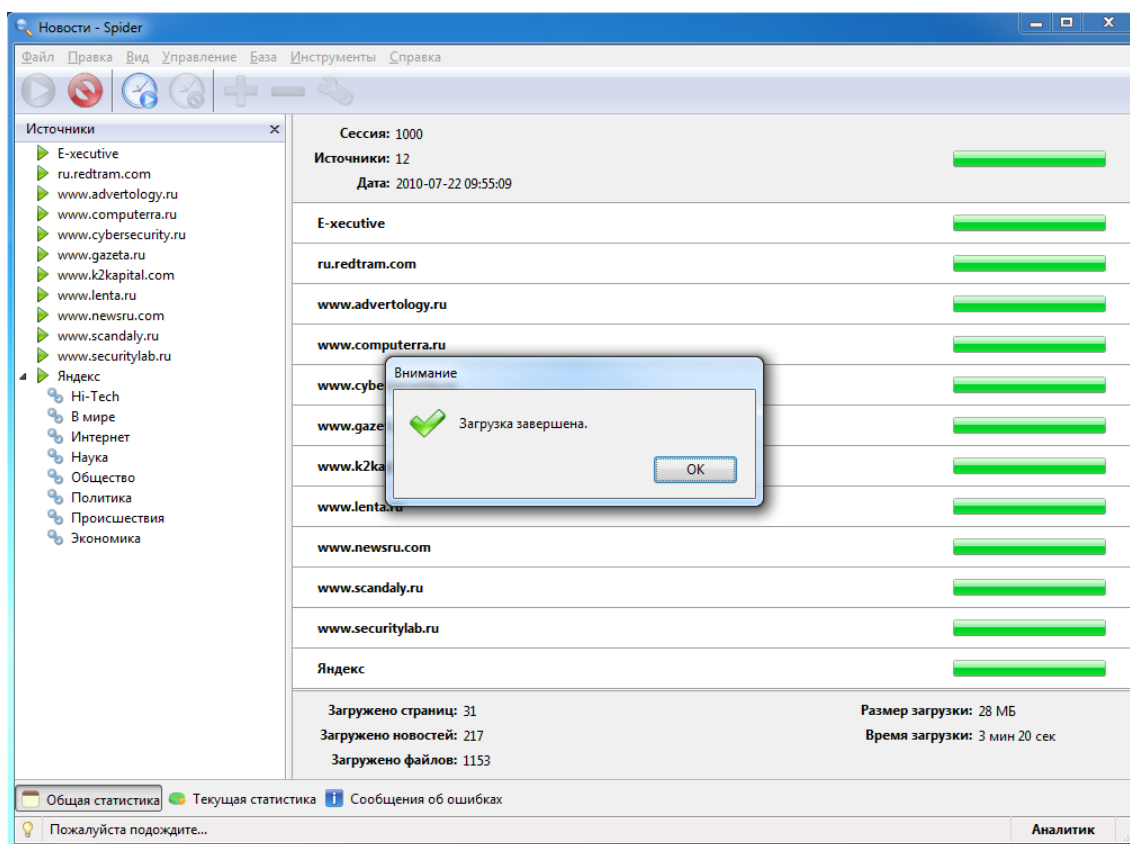
Слева задаются источники для загрузки, в комплекте поставки 12 предустановленных источников.

Теперь сделаем пробный запуск робота, нажав синюю стрелочку **старт** и посмотрим, что будет происходить.

Вы увидите справа список загрузки со статусом загрузки для каждого источника. Постепенно в процессе загрузки статус начнет меняться и через некоторое время вы увидите следующее.



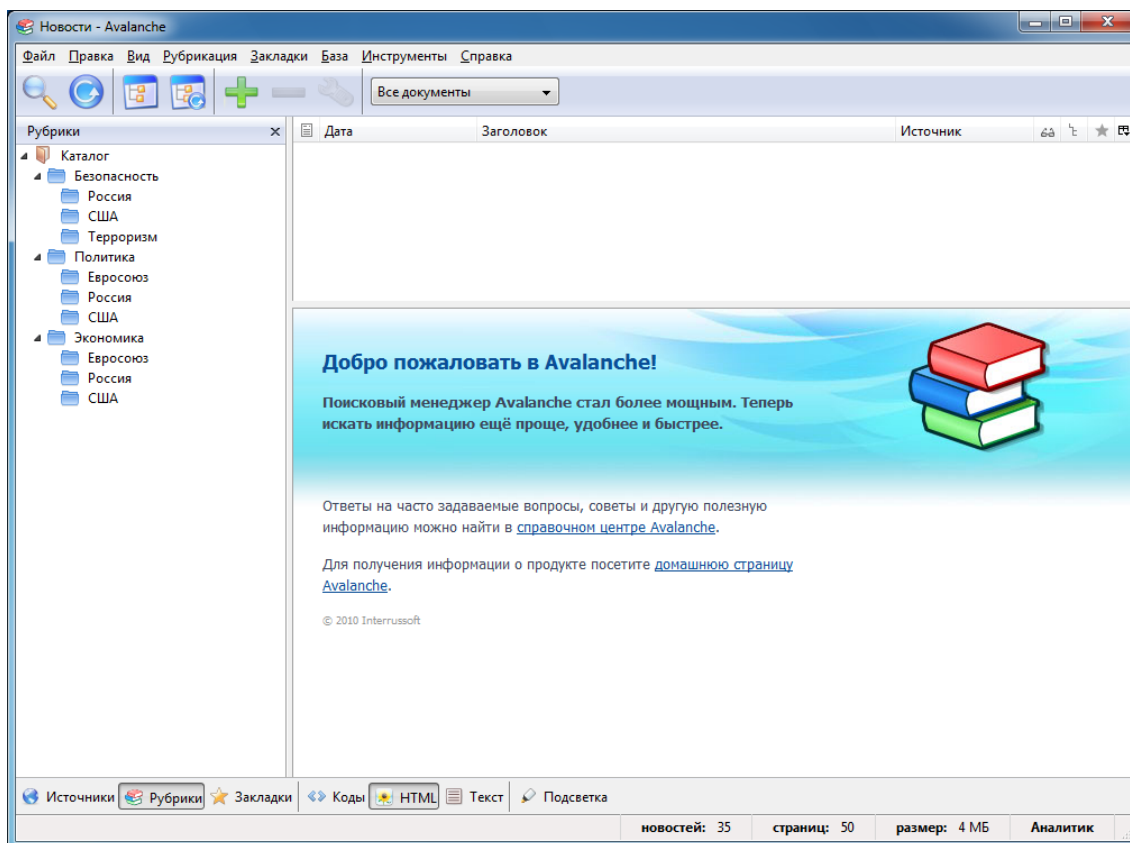
В конце загрузки произойдет рубрикация загруженных документов и появится сообщение об успешном окончании загрузки.



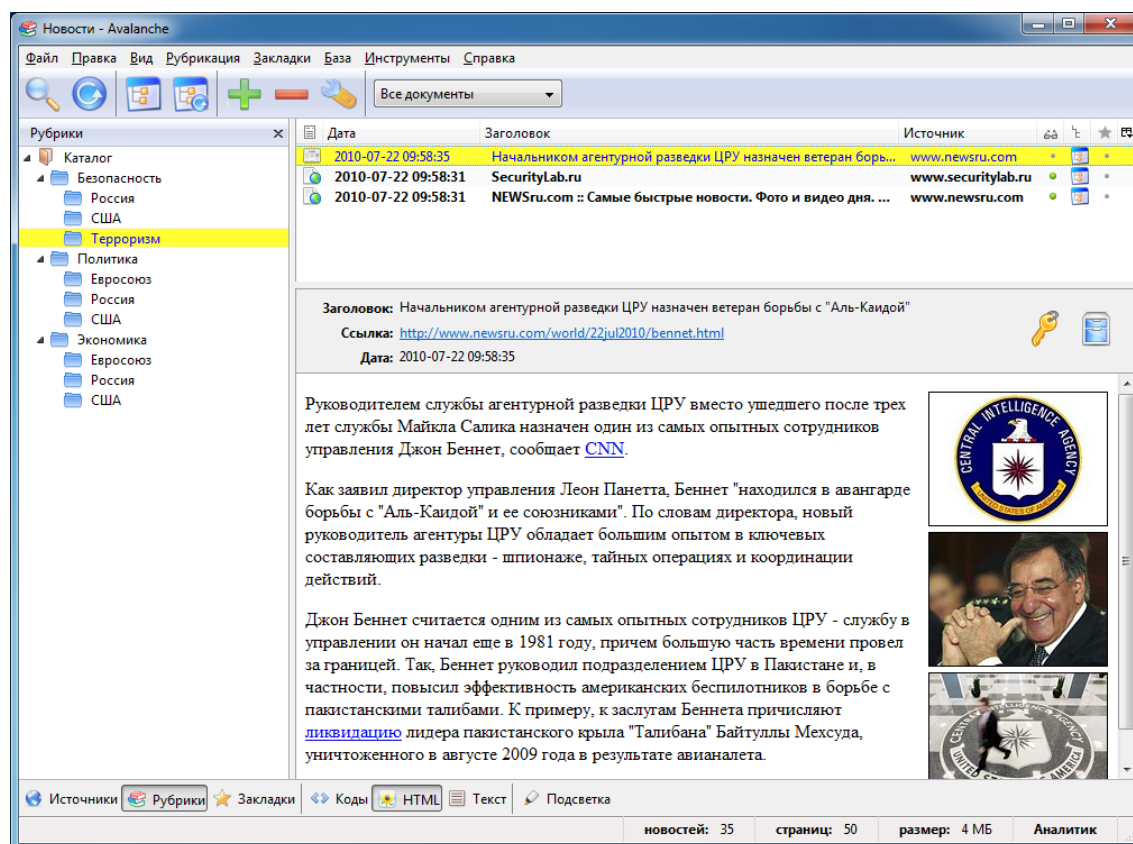
Теперь перейдем к главе посвященной просмотру загруженных результатов.

## 2 Просмотр результатов загрузки

Чтобы просмотреть результаты загрузки запустите программу Avalanche. Программа Avalanche запустится и на экране вы увидите.



Теперь пощелкайте по разным рубрикам и посмотрите попавшие в них документы.

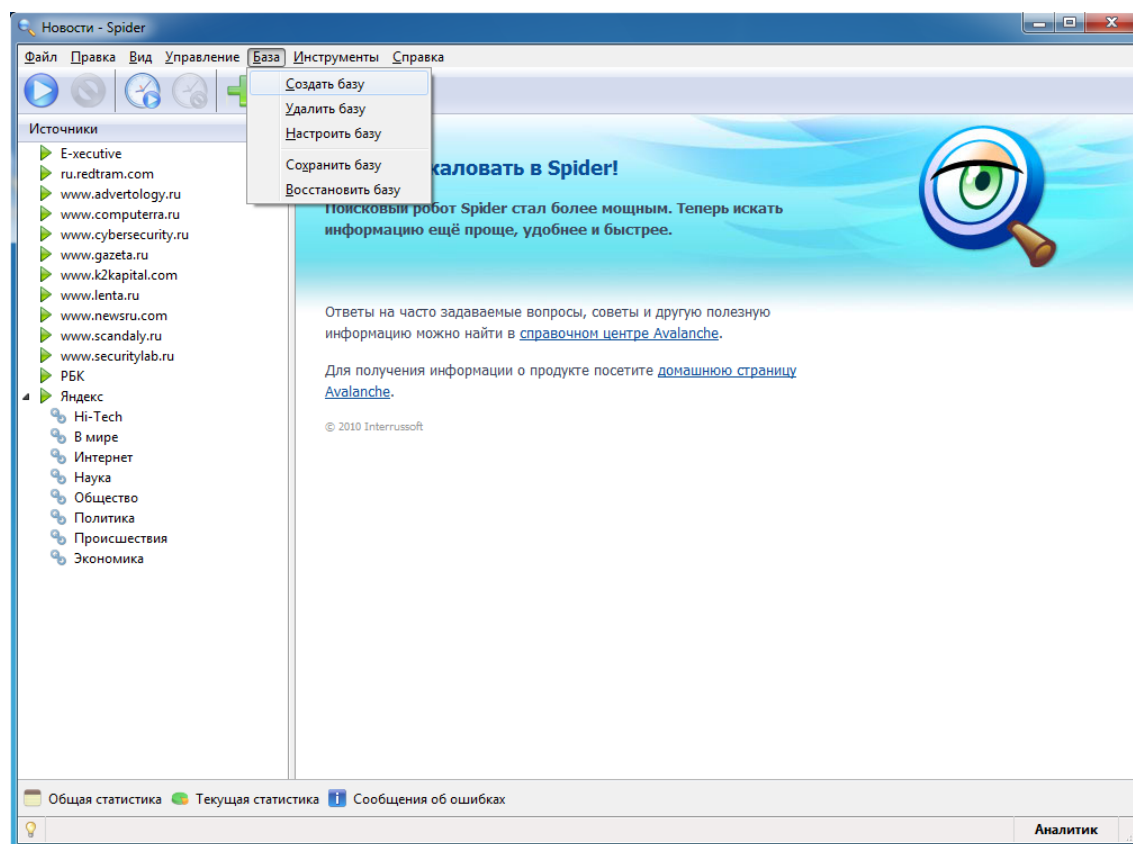


В следующей главе вы научитесь создавать и настраивать собственный источник загрузки.

### 3 Настройка новых источников

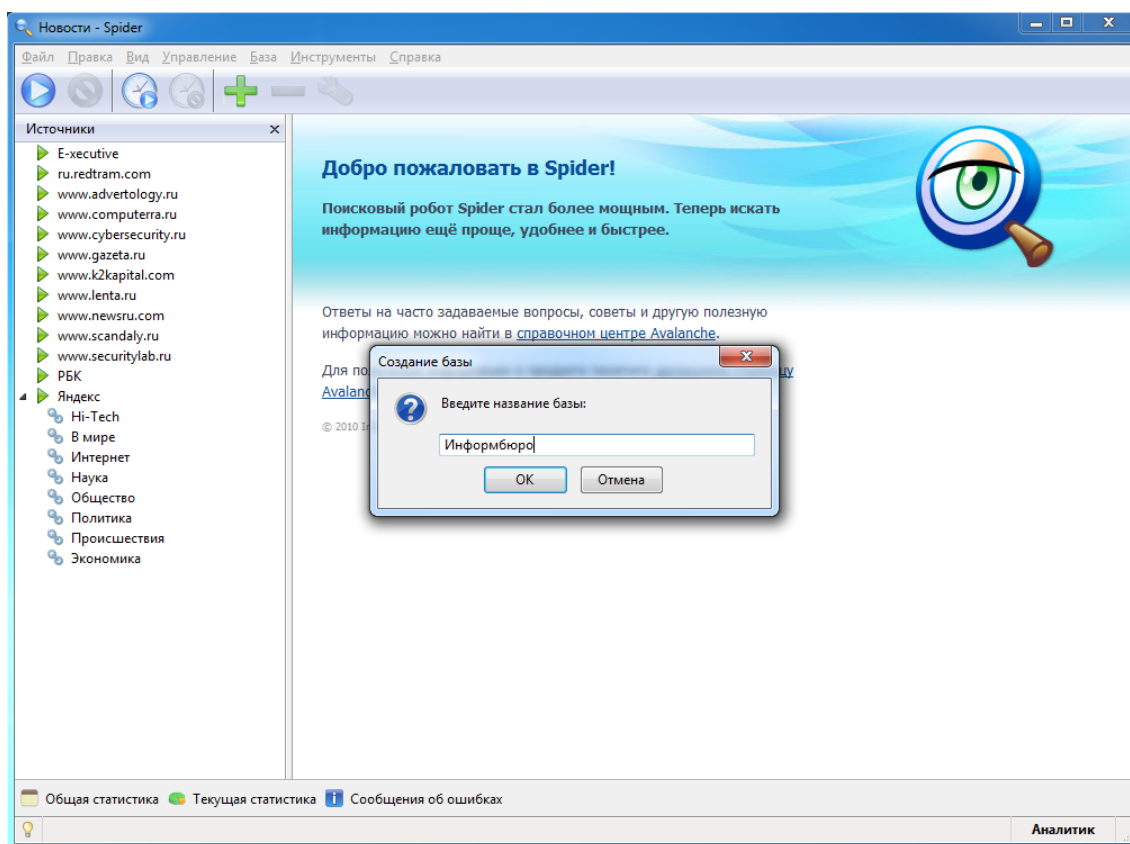
Предназначение Аваланча в персональной настройке поиска, то есть в задании ваших собственных источников. Например мы хотим создать базу для загрузки новостей популярных информационных агентств.

Первым шагом будет создание новой базы загрузки:

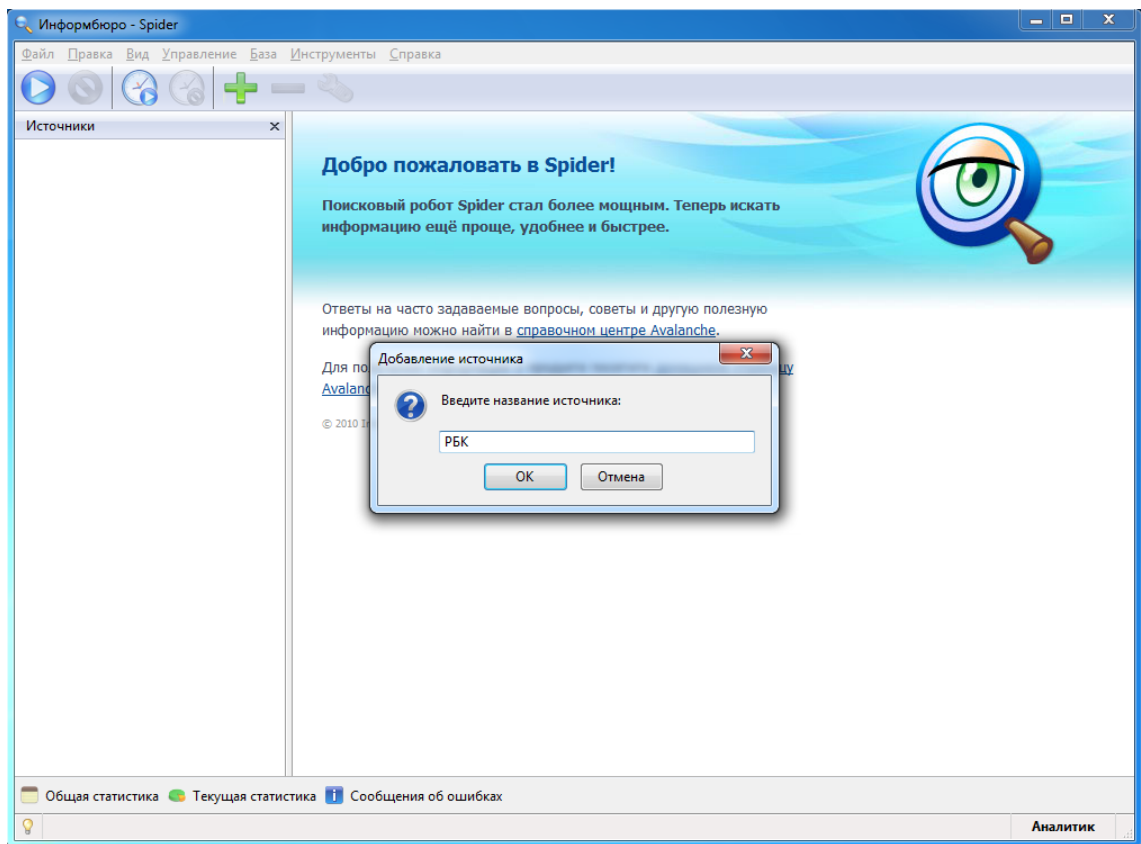




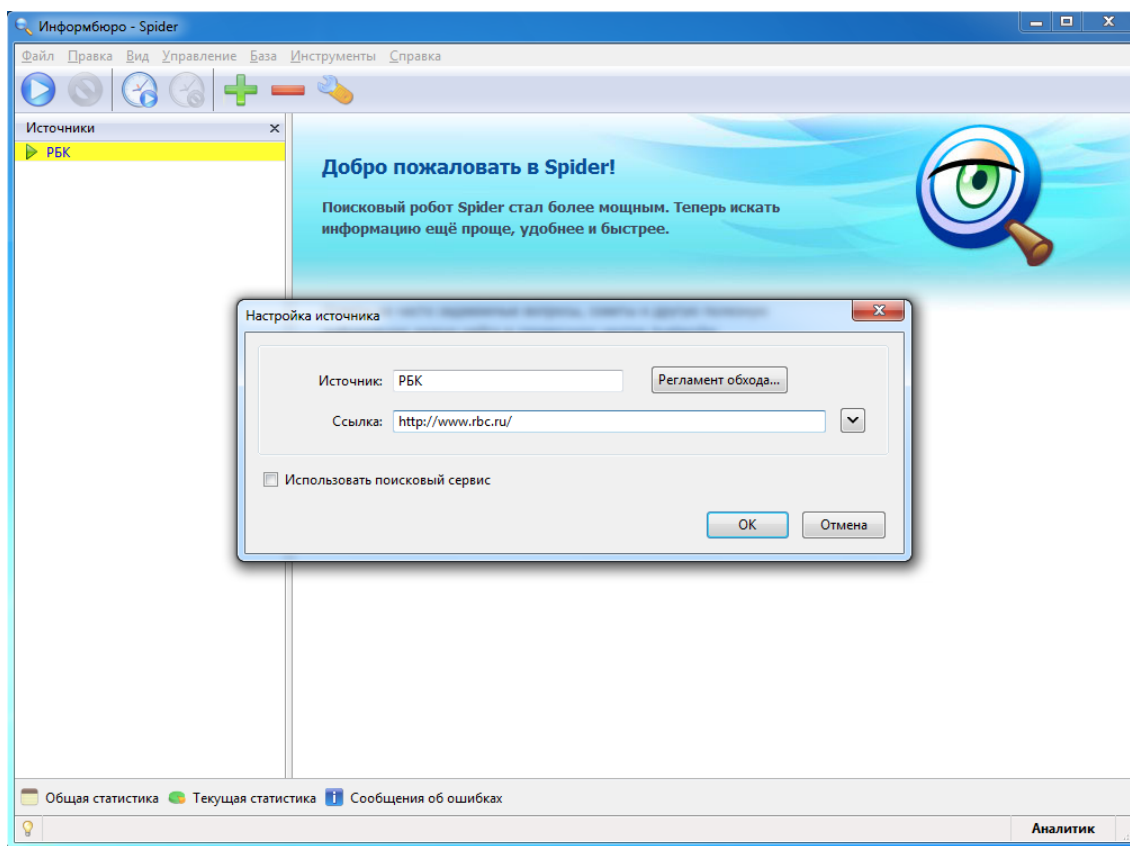
Задайте название вашей новой базы, например Информбюро.



Теперь добавьте первый источник - информационное агентство РБК. Для этого кликните на плюсик и введите в окне его название:

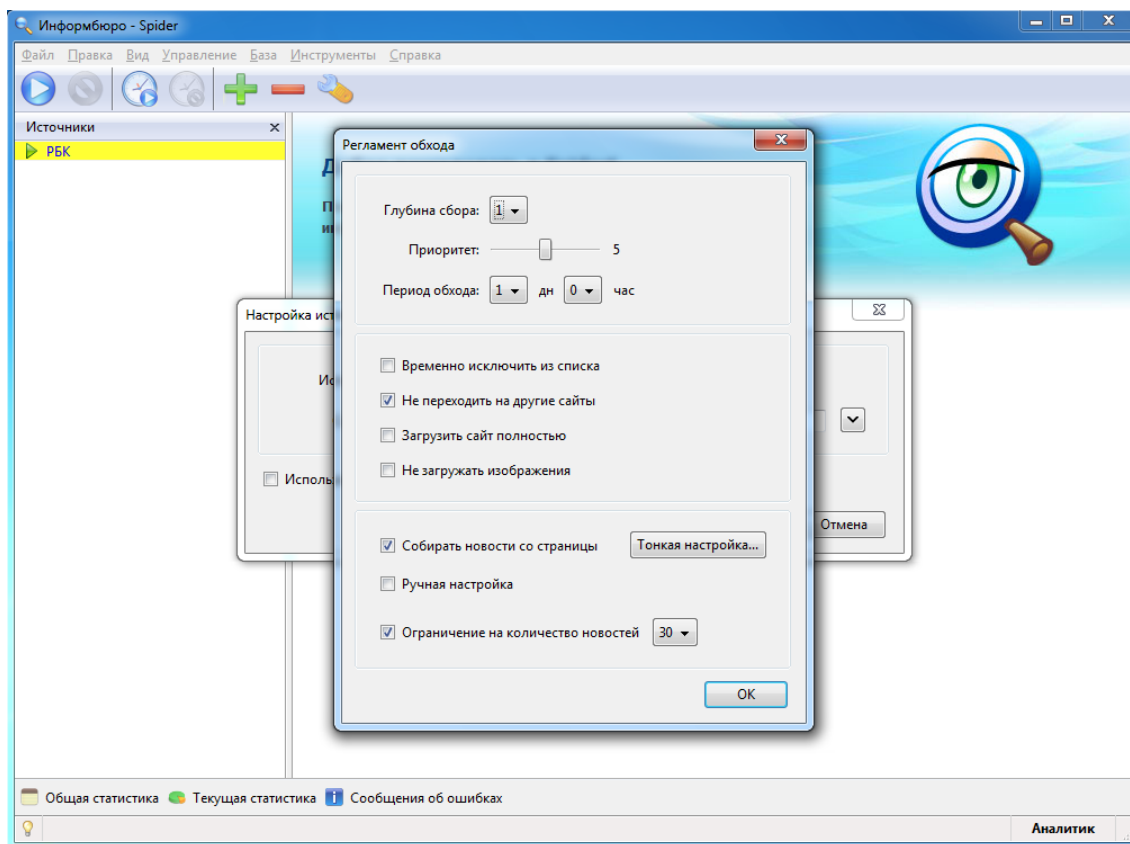


Теперь настроим источник, выберите только, что созданный источник РБК и нажмите на кнопку с изображением гаечного ключа, в открывшемся окне введите ссылку.



Обязательно используйте префикс `http://` иначе будет ошибка и вы не сможете сохранить изменения в диалоге.

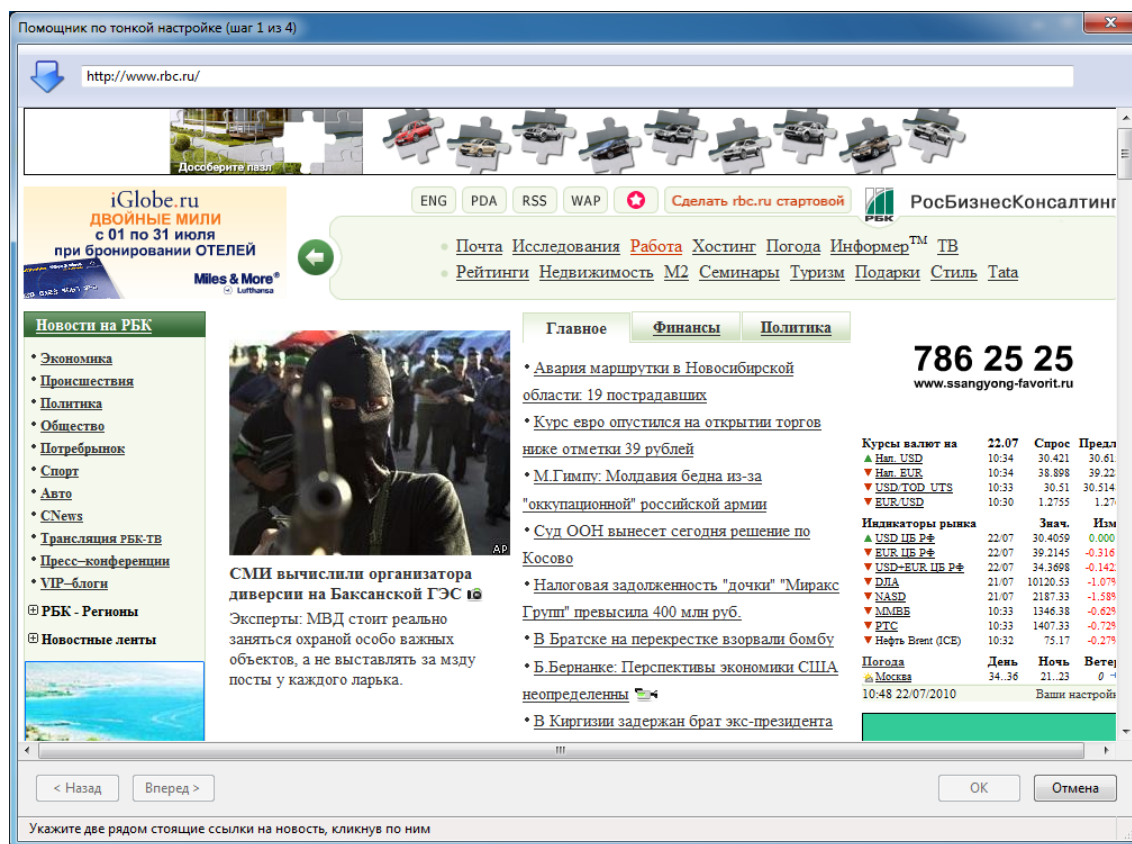
Теперь настроим сборщик новостей, для этого не выходя из диалога настройки источника нажмите Регламент обхода. Появится вот такой диалог.



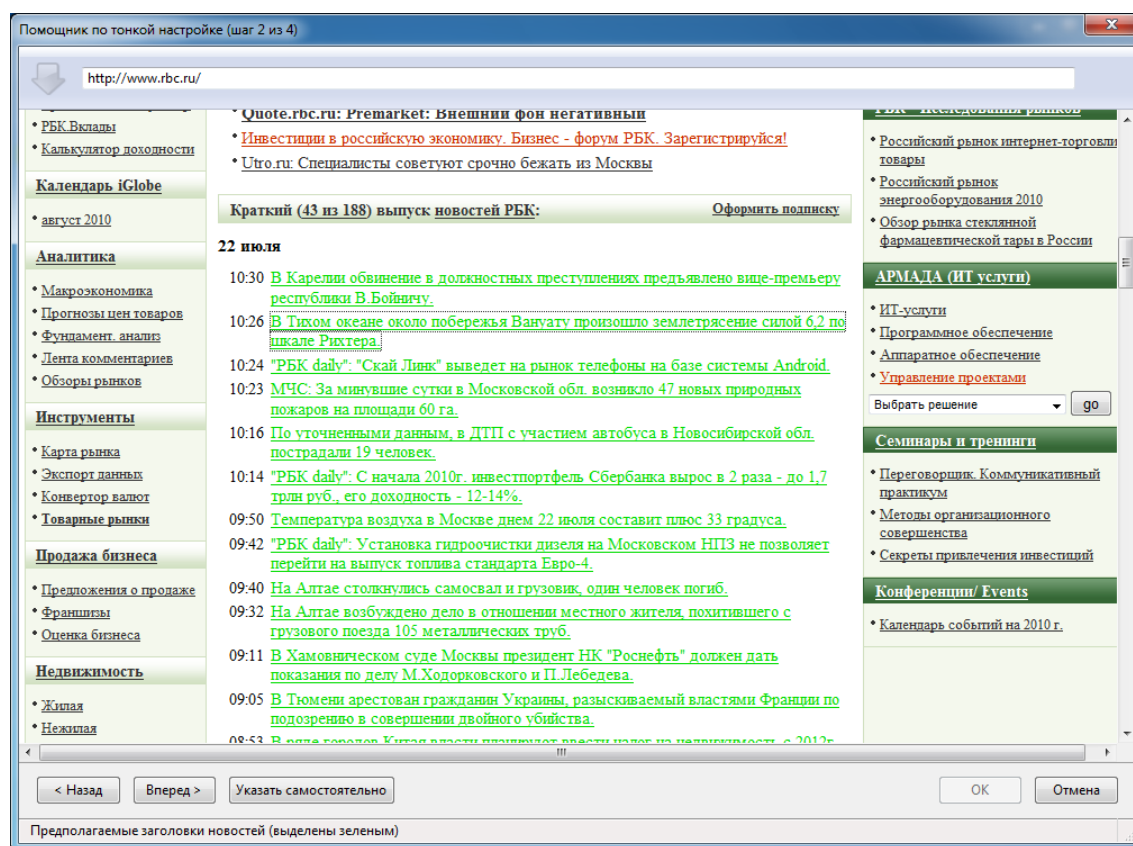
Не закрывайте этот диалог и сразу переходите к чтению следующей главы, в которой речь пойдет о тонкой настройке на новости.

## 4 Помощник по тонкой настройке

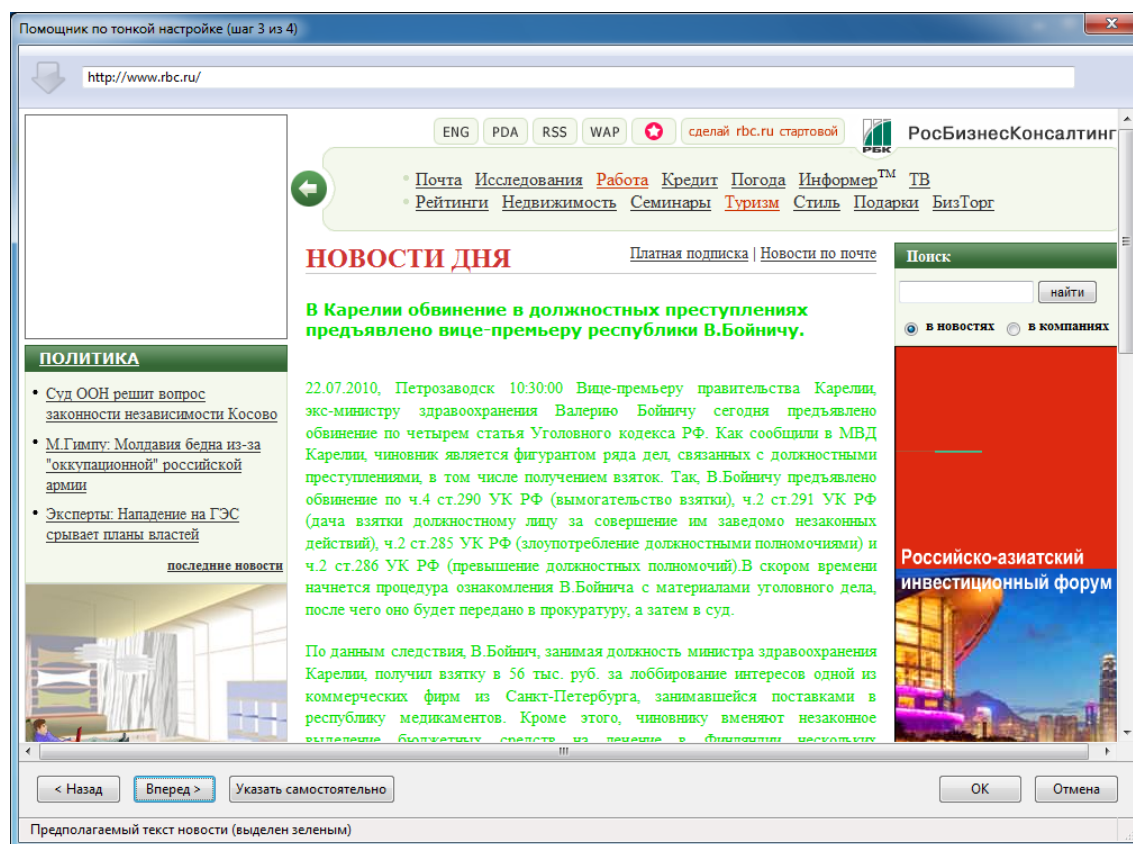
Нажмите на кнопку Тонкая настройка из диалога с регламентом при настройке источника. Вы увидите, как по указанной ранее ссылке загрузится страница и вы попадете на первый шаг автоматизированной визуальной настройки на новости.



На первом шаге укажите две рядом стоящие ссылки блока новостей который вам нужен. Вот на следующем экране показан результат этой операции, можно заметить, как помощник по тонкой настройке подсветил предполагаемые заголовки новостей зеленым цветом.

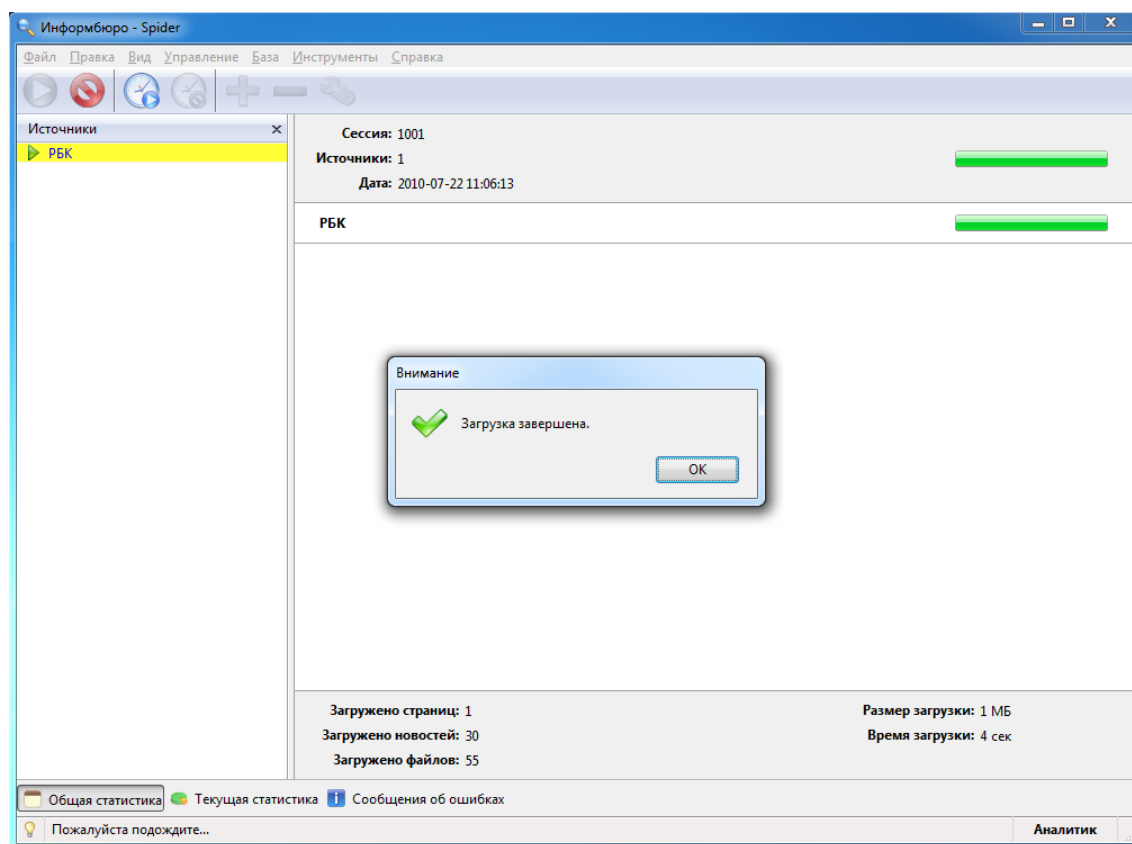


Нажмите вперед и подождите пока настройщик загрузит новость из подсвеченного списка. И тут настройщик уже подсветит предполагаемый текст самой новости.



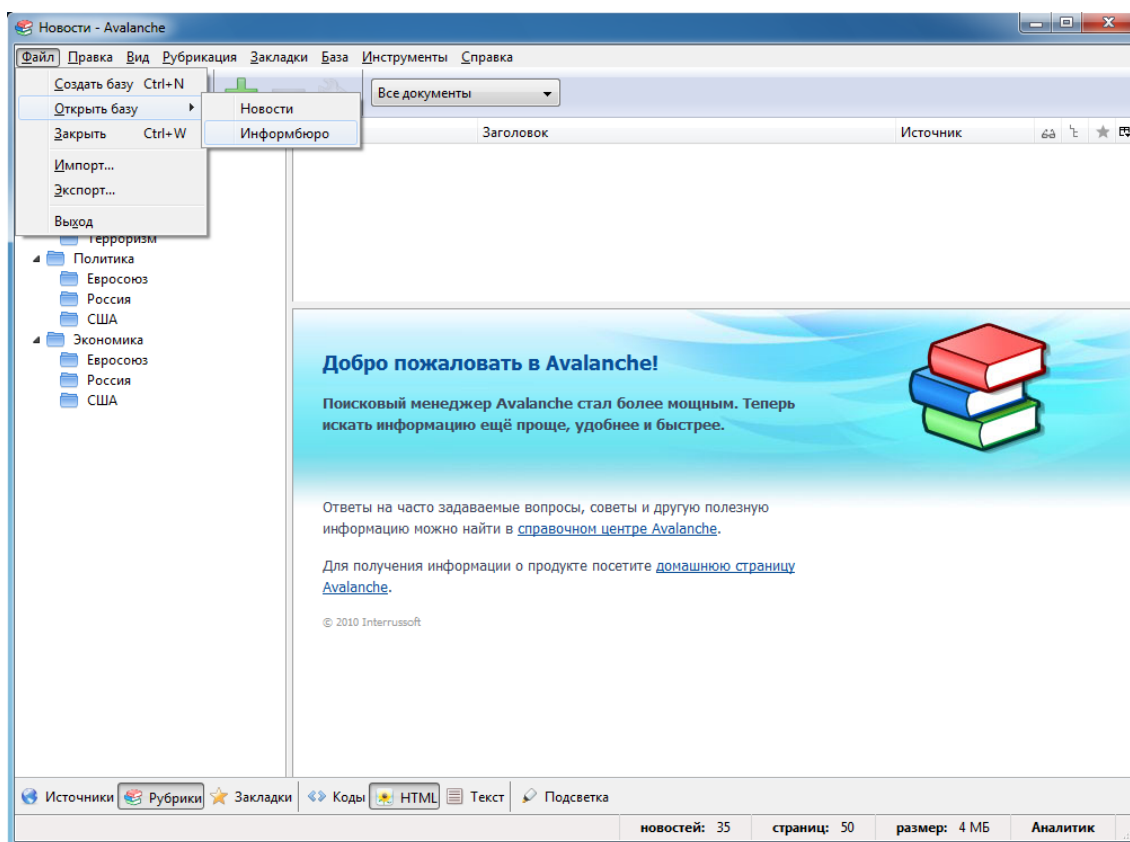
Убедившись в правильности такой подсветки, нажмите ОК. И далее ОК в регламенте и затем в диалоге с настройкой источника.

Теперь можно загрузить все новости из заново настроенного источника РБК, кликнув на стрелочку Старт.





Теперь открыв программу Аваланче мы откроем нашу новую базу.



И перейдем на страницу с Источниками, где и просмотрим список загруженных новостей.

The screenshot shows the 'Информбюро - Avalanche' application window. The interface includes a menu bar (Файл, Правка, Вид, Рубрикация, Закладки, База, Инструменты, Справка) and a toolbar with icons for search, refresh, and other functions. A dropdown menu shows 'Все документы'.

The main content area is divided into two sections. The top section displays a list of news items with columns for 'Дата' (Date), 'Заголовок' (Title), and 'Источник' (Source). The bottom section shows a detailed view of a selected news item.

Дата	Заголовок	Источник
2010-07-22 11:06:18	Руководство EADS: В отношении американских оборонных...	РБК
2010-07-22 11:06:18	В ряде городов Китая власти планируют ввести налог на не...	РБК
2010-07-22 11:06:18	В г.Бийске Алтайского края могут быть отменены прямые ...	РБК
2010-07-22 11:06:18	Командир флотилии НАТО по борьбе с пиратами: Морские...	РБК
2010-07-22 11:06:18	Крупнейшие нефтяные компании вложат 1 млрд долл. в созда...	РБК
2010-07-22 11:06:18	Температура воздуха в Москве днем 22 июля составит плю...	РБК
2010-07-22 11:06:18	Увелич сотовым Забайкальского края, сбивший 17 июля	РБК

**Заголовок:** Крупнейшие нефтяные компании вложат 1 млрд долл. в создание системы ликвидации утечек нефти

**Ссылка:** <http://www.rbc.ru/rbcfrenews/20100722065405.shtml>

**Дата:** 2010-07-22 11:06:18

### Крупнейшие нефтяные компании вложат 1 млрд долл. в создание системы ликвидации утечек нефти.

22.07.2010, Лондон 06:54:05 Крупнейшие нефтяные компании, среди которых ExxonMobil Corp., Chevron и Royal Dutch Shell, вложат 1 млрд долл. в создание новой системы ликвидации утечек нефти, передает Би-би-си. Как отмечается, новая система будет способна собирать на глубине до 3 тыс. м до 100 тыс. барр. нефти в день. Компании объединились для создания такой системы после аварии на нефтяной платформе BP в Мексиканском заливе.

Напомним, нефтяная катастрофа у побережья США началась 20 апреля 2010г. со взрыва метана на платформе Deepwater Horizon, находившейся под управлением британской нефтяной компании BP. Спустя два дня платформа затонула, а из поврежденной скважины на глубине 1,5 км начала вытекать нефть. За время аварии в воды Мексиканского залива вылилось несколько миллионов литров нефти, что сделало эту аварию крупнейшей нефтяной катастрофой в истории США.

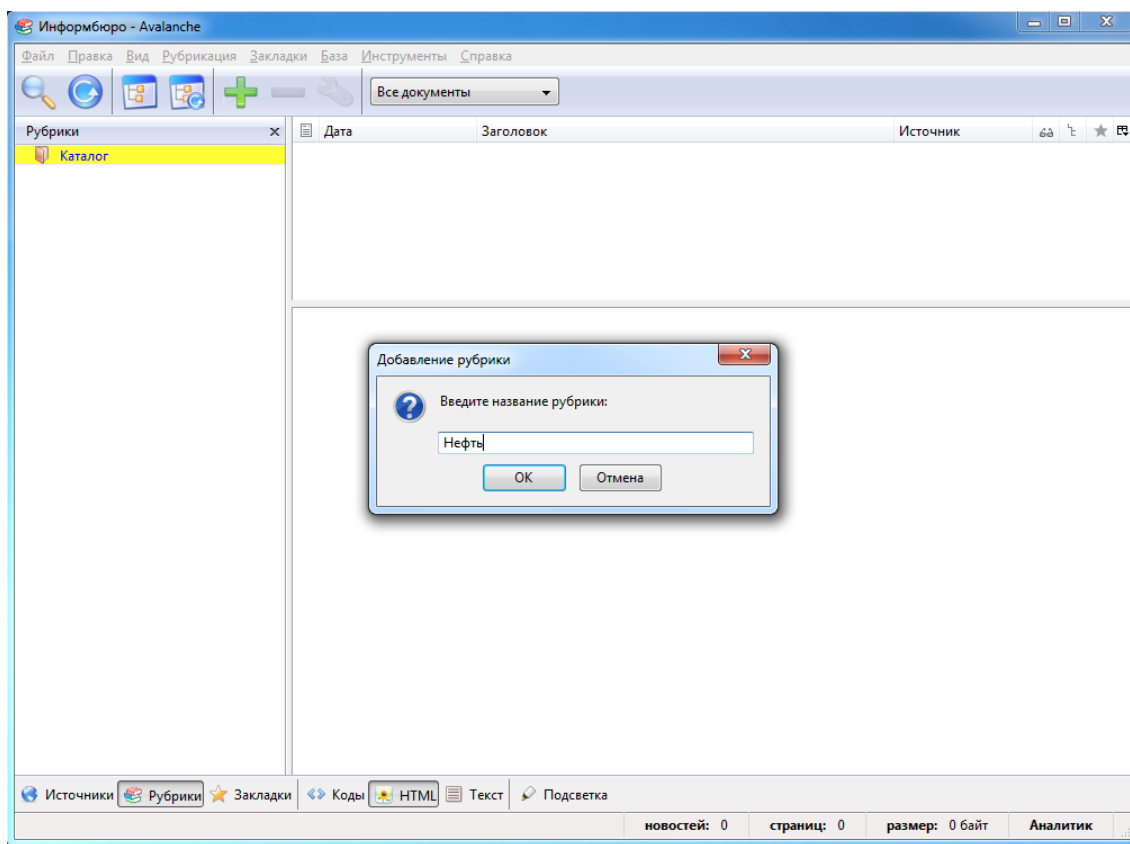
Авария в Мексиканском заливе нанесла огромный удар по имиджу и финансовому состоянию BP. С

At the bottom of the window, there is a status bar with the following information: 'новостей: 30', 'страниц: 1', 'размер: 165 КБ', and 'Аналитик'.

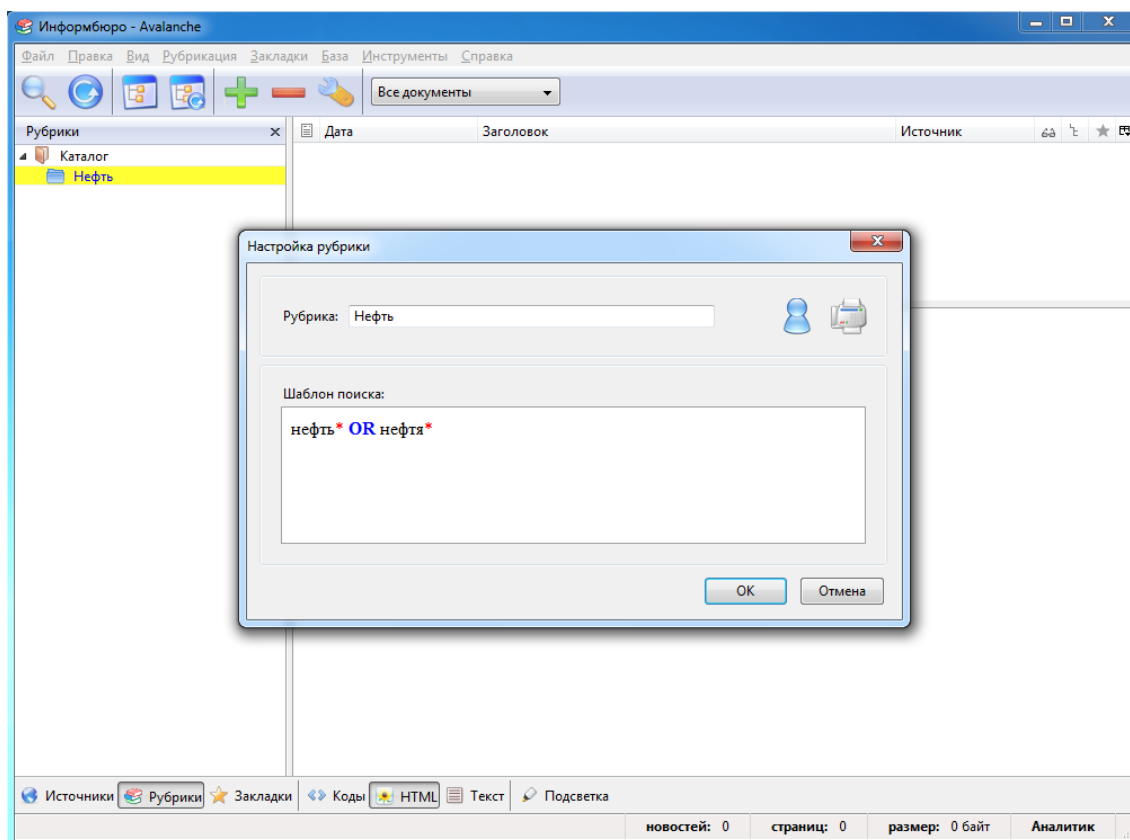
В следующей главе мы рассмотрим процесс создания и настройки новых рубрик.

## 5 Настройка новых рубрик

На закладке с рубриками выберите корневой Каталог и нажмите на кнопку с зеленым плюсиком. И в диалоге добавления рубрике введите ее имя, например рубрика будет называться нефть.

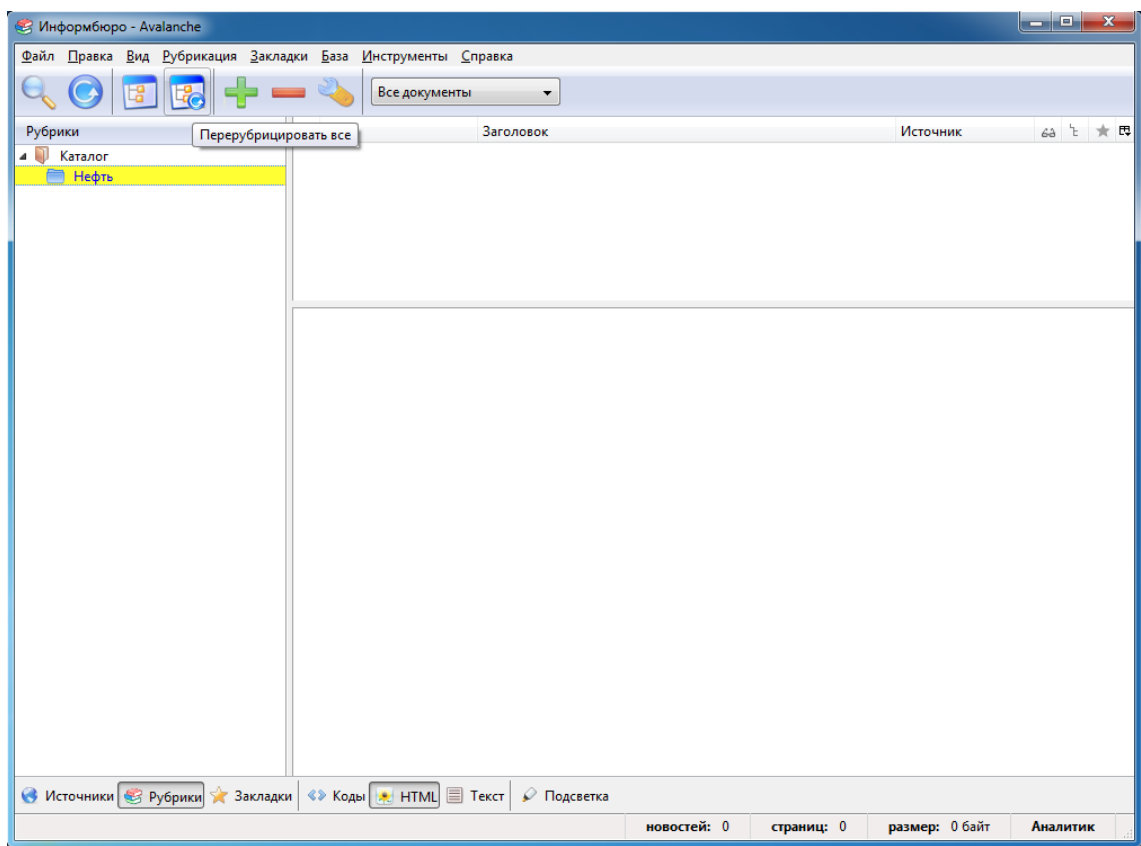


Теперь выберите эту новую рубрику и нажмите на кнопку с изображением гаечного ключа для ее настройки. Задайте шаблон поиска по которому документы будут попадать в эту рубрику.

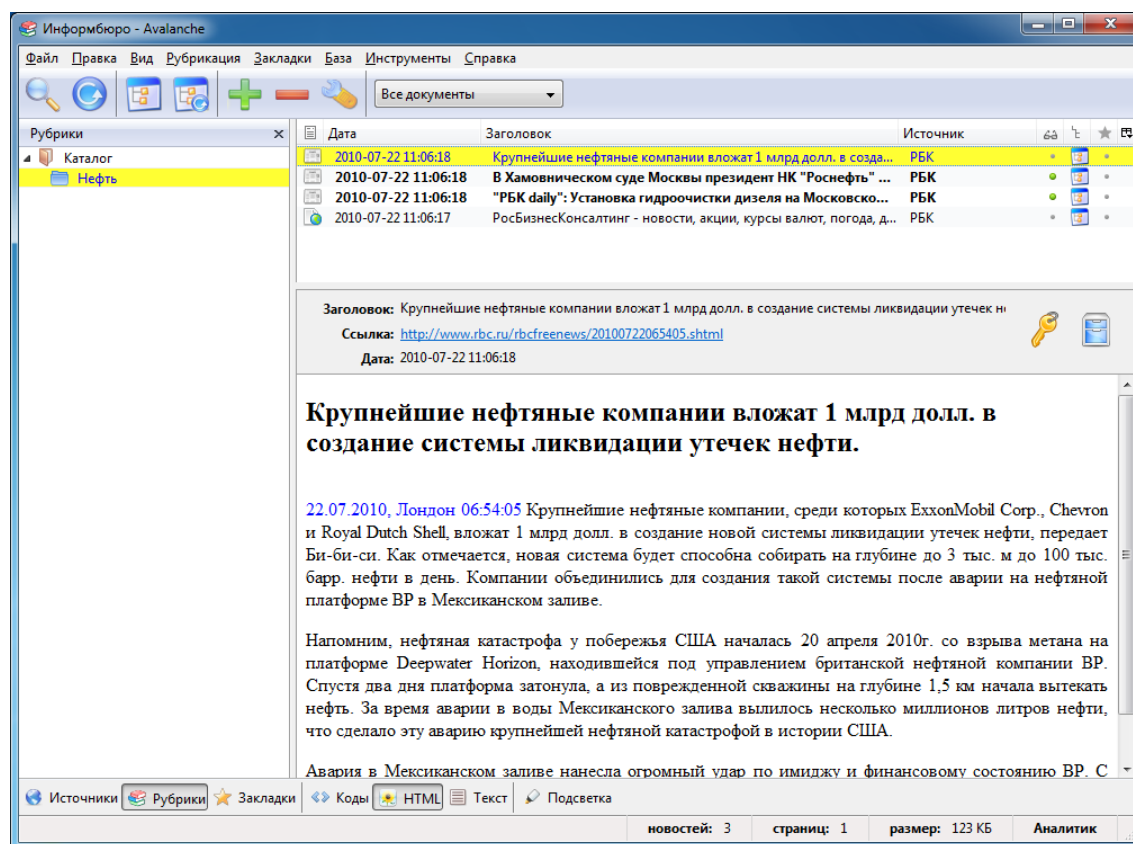


Звездочка в конце слова отражает любое возможное окончание и нет необходимости в перечислении вариантов.

Чтобы заново перерубрицировать уже загруженную базу документов нажмите на кнопку Перерубрицировать все. А вот будущие загрузки будут уже рубрицироваться с учетом этой новой рубрики.



Ну и наконец посмотрим результат рубрикации, для этого выберем только, что созданную рубрику и посмотрим попавшие в нее документы.



Итак, мы закончили первое знакомство с персональной поисковой системой Avalanche 2.5. За более подробной информацией обратитесь к руководству пользователя, входящему в состав пакета приложения.

Также полезную информацию вы можете найти на сайте разработчиков.