

Overview

Newspiper is a "*news aggregator*" program specially designed to organize channel (or "news pipe") which delivers headlines from your favorites news internet sites (like cnn.com or abcnews.com) directly to your desktop.

You can setup NewsPiper to retrieve news headlines (as well as any other information) from any web page or RSS file and then it will show downloaded information in the special News Line - small always-on-top window which can be moved in any place of your screen.

For each shown headline you can open "full story" page in your default web browser simply by clicking on the headline.

You can add new site for scanning (channel) just by indicating the URL and special script for headline extraction (is not necessary in case of using RSS). You can also specify the language and code-page of your channel for non-English news sites.

In addition NewsPiper can check your connection to some Internet address (your can set your own one) by the schedule you establish (one time per minute, 5 times per hour, etc.) and changes the color of its icon in the system tray bar accordingly: green - your are online, red - offline.

Getting started

Below you'll find the answers to frequently asked questions about NewsPiper and step-by-step instructions for several most useful operations.

Q: Two kinds of channel types: HTML and RSS. What are they?

NewsPiper can work with two types of channels. HTML channels retrieve information from the usual HTML web pages using special [News extraction script](#) while RSS channels take all necessary information from the special XML files (called RSS files) which aggregate news and the content of news-like (and not only) sites. To learn more about RSS - visit Internet resources which describe this format and its using, for example ["What is RSS"](#) article on [xml.com](#) site.

It is more preferable to use RSS channels since they already contain only necessary information which is right structured. So extracting the information from RSS file usually occurs more quicker while in case of using of HTML channel it is necessary to download the whole web page and then process it by "extraction script".

However most news sites do not publish their information via RSS yet so in such case you can create your own HTML channel or install already existing one from [Online Channels Repository](#).

To install new channels from Online Channels Repository:

- Open main window by double-clicking on icon in the system tray bar;
- Select "Channels | Online repository" menu item;
- The list of available channels will appear;
- Select necessary channels and press Install/Update button;
- Press Close button to close the dialog.

To add new channel:

- Open main window by double-clicking on icon in the system tray bar;
- Press "Add HTML/RSS Channel" button, the Channel Parameters dialog will appear;
- Set the title of the channel, URL, language and code page of the site and [News extraction script](#) for HTML channels;
- Press OK button.

To set the schedule for news retrieving:

- Open "Options" dialog through menu Action | Options...;
- Go to "News Line" page;
- Select the period (e.g. 1 hour) and number of times NewsPiper will connect to news sites during indicated period;
- Press OK button to save the changes;

To disable News Line:

- Right-click on the NewsPiper's icon in the system tray;
- Unselect "Show News Line" item in the menu.

How to register

NewsPiper is distributed as shareware. It means you can freely use unregistered version of NewsPiper during 30-days trial period. To continue use it after the end of trial period you should to register NewsPiper. Registration costs: \$19.95

For more information see our licence agreement (file license.txt).

To register online using secure server go:

<http://www.korzh.com/order.shtml#newspiper> and select type of license you would like to purchase.

Contact information

For comments, suggestions, bug reports contact: newspiper@korzh.com

To get up-to-date information about the program visit Crammer homepage at:
<http://www.korzh.com/newspiper>

Localization

NewsPiper has an ability of interface localization (in this case, it's setting the interface language by user). You should choose the item "Languages" and then choose necessary language.

Additional languages you can be download from the NewsPiper's home page:

<http://www.korzh.com/newspiper>

If your language is not supported yet then you can translate NewsPiper manually using special Language Manager utility: <http://www.korzh.com/download/lmsetup.exe>

Instructions for translation:

1. Download, install and run Language Manager;
2. Select "File | Open project" menu item and open **NewsPiper.ntv.lng** file placed in the ... \NewsPiper\lang folder;
3. Create new language using "File | Create new language" command;
4. Translate all items and save your language file. It will be saved in the same folder with **NewsPiper.xxx.lng** file name where xxx is two or three-letter abbreviation of your language (EN – English, DE – German, FR – French and so on);
5. Run NewsPiper. New created language must appear in View | Languages menu item;
6. Select it and NewsPiper interface elements will appear in new language.

Main Window

By double clicking NewsPiper tray icon or choosing "*Channel list*" from the NewsPiper tray menu, you will get the main form which consists of three main parts:

- menu and toolbar;
- list of channels and categories;
- news headlines for selected channel or category (only for enabled channels);

Using the menu commands or toolbar buttons you can setup channels list, edit program settings, import/export channels etc.

See [Menu and toolbar commands](#) topic to get the description of all operations available through menu or toolbar.

Channels list contains list of all installed categories, channels and corresponding URLs. Each channel can be enabled/disabled by checking/unchecking corresponding checkbox control near the channel name. NewsPiper does not retrieve news and does not show headlines for disabled channels.

Below the channels list you can see "Current News" panel which contains list of news headlines for currently selected channel (if it is not disabled) or category. In the second case (category) all headlines are grouped by sub-categories and channels.

You can hide "Current News" panel by unchecking corresponding checkbox control at the top of the panel or corresponding menu item in View menu ("Show headlines panel").

Menu and toolbar commands

Below you will find the list of all menu commands (and corresponding them toolbar buttons):

menu "Action"

Show News Line menu item

Turns on/off the NewsLine - a small always-on-top window for scrolling news headlines.

Retrieve all news

Connect to all currently listed news sites (for enabled channels only) and refresh headlines list.

Stop all network jobs

Terminates all currently executing network job.

Edit program settings

Bring on "Program Settings" dialog where you customize program appearance, connection schedule and other options.

Languages

Using this menu item you can change current interface language of the program. Just select the necessary one from the list of available languages.

Exit

Closes the program

menu "View"

Horizontal/Vertical Layout

Switch between two types of main window layout: first - "Current News" panel at the bottom, second - "Current News" panel at right;

Show headlines panel

Turn on/off "Current News" panel;

Show URL column

Turn on/off URL column in channel/categories list;

menu "Channels"

Add HTML Channel

Creates new HTML channels and shows Channel Parameters Dialog where you can set all channel settings.

Add RSS Channel

Creates new RSS channels and shows Channel Parameters Dialog where you can set all channel settings.

Edit

Opens Channel Parameters Dialog where you can change any channel settings.

Delete

Deletes currently selected channel from the list. You will be asked for a confirmation of deleting.

Add Category

Creates new category as sub-category of the current one;

Edit Category

Show Category Parameters dialog where you can change category name;

Delete Category

Deletes currently selected category from the list. The category must be empty (no sub-categories, no channels).

You will be asked for a confirmation of deleting.

Online Repository

Opens Channels Repository dialog where you can see the list of channels and categories available in repository placed on www.korzh.com site and install/update selected channels.

Channel parameters dialog

"Channel parameters" dialog contains the following fields:

General page:

"*Title*" - represents the name of the channel which will be shown in main window.

"*Page URL*" or "*RSS file URL*" - address of the web-page or RSS file which contains news headlines (for example: <http://www.cnn.com>);

"*Site base URL*" - address of root web-page for current site. For example for the page <http://www.somesite.com/news.html> this parameter will be <http://www.somesite.com>;

"*Web page language*" - specifies the language of the news page;

"*Code page*" - news page encoding;

Script page (HTML channels only):

"*Script*" - specifies the rules how the program should extract headlines from the page. See more detailed description in [Extraction Script](#) topic.

Appearance page:

"*Headlines font and color*" panel - allows to specify font and color of the [News Line](#) for current channel.

"*Default*" button - allows to set the default color and font of News Line for current channel. The default color and font are set in the [Program Settings](#) dialog;

"*Channel tag*" check box - turn on/off tag of the channel - short (6 symbol maximum) label of the news site shown in the left side of the News Line;

"*Tag name*" text field - name of the tag;

"*Tag font and color*" panel - allows to specify font and color of the tag label;

"*Show Channel Icon*" check box - allows to turn on/off a small icon shown in News Line for each channel. The icon is downloaded directly from the site (favicon.ico).

Online channels repository dialog

Use this dialog to install/update your channels from online channels repository published on www.korzh.com/newspiper/channels.

When you open this dialog NewsPiper tries to connect to repository and get the list of channels.

On success it shows you the list of available channels including the title, URL and the language of each channel. The image in the second column indicates if any particular channel is already installed on your computer or it is new for you.

To install or update channels just check the checkboxes near the channels you select and press Install/Update button at the bottom of the dialog.

News Line window

News Line is a special "always-on-top" window which continuously shows news headlines.

It consists of the following two parts:

Tag panel - left side of the News Line window which shows the short name of the current channel;

Headline panel - the place where the headlines appears.

The color and font of the tag and headline panels can be changed separately for each channel. See [Channel parameters dialog](#) topic for details.

By clicking on the currently showing headline you can open associated "full story" page in default browser.

To move the News Line - point mouse cursor to the tag panel or window border (the cursor will be changed accordingly), press left mouse button and drag the window.

Right mouse click on the News Line window will raise the context menu where you can call [Program settings dialog](#) (News Line section), hide the NewsLine or open [main window](#) of the program (channels list).

News extraction script

Overview

NewsPiper has special script language to extract news headlines and other information from the web-pages. A program on this language contains a list of predefined operations separated by semicolon. Here is the example of such script program:

```
StartAt '<!-- BEGIN HEADLINES --->';

AssignTemplate '<a class="headline" href="($url)">($title)</a>' ;

FindTemplates;
```

Function **FindTemplate** makes search and extraction of the headlines which are suit to the template specified in **AssignTemplate** operator. Function **StartAt** is used to set start point of searching which allow to speed up extraction process and avoid including unnecessary parts of web-page. There is similar function **EndAt** which can specify the end point for headlines searching.

Regular Expressions

As you see to recognize necessary information on a web page each channel uses regular expressions which specified in **AssignTemplate** function. The format of the regular expressions is similar to the one used in PERL language.

The regular expression can contains the part of the HTML text from a web page and the following main variables (all variables begins from the '\$' sign):

(\$url) - substituted by URL to "full story" page for current headline;

(\$title) - substituted by headline title;

(\$anytext) - any text (including HTML tags and simple text). This variable is used only for beter readability of the script. It corresponds to ".*?" factor definition (without quotes) in PERL regular expressions.

NewsPiper scans indicated web-pages and searches for pieces of HTML code which fit to specified template. If such piece is found then it show the part of this code substituted by (\$title) variable and stores URL for headline using (\$url) variable.

Here is the example of real headline code and corresponding regular expression:

HTML code:

```
<a class="topHeadline" href="/money/2003/09/29/technology/fortune_women/index.htm">Fortune's
most powerful women</a>
```

Regular expression:

```
<a class="topHeadline" href="($url)">($title)</a>
```

as you see after scanning the avobe piece of web page NewsPiper will extract [Fortune's most powerful women as the headline title and /money/2003/09/29/technology/fortune_women/index.htm](/money/2003/09/29/technology/fortune_women/index.htm) as corresponding URL.

Old Script description

Code:

```
<div class="cnnMainNewT2">&#8226; <span class="cnnBodyText" style="font-weight:bold;">CNN/Money: </span><a href="/money/2003/09/29/technology/fortune_women/index.htm?cnn=yes">Fortune's most powerful women</a>
```

Template:

```
<div class="cnnMainNewT2">$ignore<a href="$url">$title</a>
```

In this example **\$title** variable represents the following headline: "Fortune's most powerful women" and the **\$url** variable contains the URL corresponds to this headline:

"/money/2003/09/29/technology/fortune_women/index.htm?cnn=yes".

This URL is relative. To get the full URL NewsPiper combine it with the base URL for this channel.

There are two additional variables: **\$stop** and **\$textwithlinks**. First is used to make headline recognition process more determined. In general case NewsPiper will search for some part of HTML source till the end of the page. If you do not need this and want to make it stop searching after some sub-string - place **\$stop** variable after it. Example:
we have the following HTML code:

```
<a href="somelink.html"> Link to other site </a>
```

.....

```
<a href="/news/page1.html" class="headline"> Headline1 </a>
```

and the following template:

```
<a href="$url" class="headline">$title</a>
```

In such case NewsPiper can recognize all text between first `` as long URL and place it into **\$url** variable. Of course we do not need that. To prevent such behaviour we should add **\$stop** variable after the first URL in our template:

```
<a href="$url"$stop class="headline">$title</a>
```

So the first link will be ignored as not suitable for our template and we will get the second one as headline.

Second additional variable **\$textwithlinks** is used when the headlines on pages do not contain links to "full story" pages but the links follow in the text after headline. Example:

```
<h3>Some Title</h3>
```

```
<p>
```

```
text
```

```
.....
```

```
text
```

```
<a href="/page1.html"> Read more </a>
```

```
</p>
```

For this case you can use the following template:

```
<h3>$title</h3> <p>$textwithlinks</p>
```

NewsPiper will scans text between `<p>` and `</p>`, extract all links from there and use the first found link as link to "full story" page.

Program settings

Application's options are brought up by choosing *Options.* from tray icon menu. Here you can customize options affecting application's performance.

NewsPiper can be run on Windows startup automatically or started manually; it can start in minimized state – these options are set by appropriate check boxes.

Another group of options control the connection testing parameters.

"*Address*" field indicates the address of internet site which NewsPiper will try to connect periodically to see if you on-line or off-line. This can be the address of your ISP site or any other host which can be the indicator of your connection status.

"Font" group box gives the user an opportunity to change font used for word list showing in the main window. Just press "Select..." button and choose name and size of the font.

About "Pro" edition

You can upgrade your copy of the NewsPiper to Pro edition. NewsPiper Pro has many additional features which can help you in learning words:

1. Reminder mode

NewsPiper can remind you the words at intervals such as twice an hour or five times a week. You can choose the kind of reminder (or Flash-Card) window: "flash", "message" or "question", increase or decrease the frequency of reminder and the time of showing for each word, move and hide the reminder window. For more info about the different parameters of word in Reminder mode see [Word Parameters](#) topic

2. Exam mode

you can examine how rightly you had remembered the words by two means: "variable"-"value" (e.g. word in English – its translation in German) and vice versa. If you didn't answer rightly NewsPiper shows the correct value. NewsPiper will gather the statistic of your answers which can be viewed then in "Grade" column (see the next section for more information).

3. Statistic for each word.

your mistakes and right answers during exam (or when you give answer in Flash-Card window in reminder mode) are stored into the dictionary and can be viewed in the "Grade" column as percent of right answers in all answers you made. It allows you to see degree of learning each variable visually and to edit the frequency of word showings accordingly.

4. Export/Import operations

there is the possibility to export/import dictionary to/from text file. It is comfortable if you need to print your dictionary like list of words, dates or formulas for creating summary or to read it when the PC is unavailable.

5. Flash-card customization

All parameters of the Flash-Card window are totally customizable. You can assign size, position and color of window, font, what parts of word article to show (the word itself, its translation, transcription, etc), and other parameters.

6. Transcription for words

NewsPiper Pro has a capability to store the phonetic transcription for each word. It is necessary you should memorize not only spelling of the word but its articulation as well. See topic for more information

You start trial period for the Pro edition of the NewsPiper in About window (Help\About menu item). Just open that window and press "Pro edition" button

Script Functions Reference

Here you find the description of those operations:

AssignTemplate(Template)

Specifies a template for extracting news headlines from web pages. Template parameter is a regular expression used to recognize necessary parts of web-page. The format of regular expression is described below.

