

Web Downloader

for Intranets.

This program was designed for our Intranet to help our users to have fast and comfortable access to some often used information in the Internet. The main goal of **Web Downloader** is to copy some HTML pages with images and links to our Web server. Because of Intranet and Firewall here, I had to implement access through Proxy server.

This program is **FreeWare**. There are no special limitations in program's functionality. If you have any ideas or suggestions to improve it, You Are Welcome. Bug reports are also would be accepted with pleasure.

My E-mail address: arny@saransk.sitek.net.

If you are interested in this program, please E-mail to me and I will send you next versions announcements.

Development Time - 1 week (for 1.00.20 version).

Current Version - 2.0 Build 32

Operation system - Windows 95 or Windows NT 4.0. Used DLLs - wsock32.dll.

Author - Oleg 'Arny' Chernavin. (Russia)

Home page - <http://www.saransk.sitek.net/pages/arny/>

[Getting Started](#)

[How-To-Use Instructions](#)

[Program History](#)

Getting started

To install **Web Downloader** simply start WebDown.exe file. It will automatically put default parameters to Registry (HKEY_LOCAL_MACHINE\ SOFTWARE\ Army O. Chernavin\ WebDownloader1). All information about items for download is stored in current directory in **WebDown.dat** file.

The first step is to open **Options** dialog (File | Options menu item). You will see three tabs:

1. Download

Directory means the path where downloaded pages would be stored. You must enter fully qualified path here, including drive letter, path and the “\” slash at the end. This path could be local or networked. Default - “c:\temp\”.

URL Prefix. This parameter takes effect only if you will store downloaded pages on your Web server. In this case you must enter virtual directory name. For example, you may create the **servers** virtual directory on your Web server. Set this parameter to “/servers/”. The “/” slashes at the beginning and the end required.

Default document determines how to name files when there is no filename for downloaded page. For example, URL like “http://www.server.com/” and “http://www.server.com/News/” contain no file name. **Web Downloader** will create files with name specified here. Default - “default.htm”.

Log Connections. Check this option if you want to record download progress in a box at the bottom.

2. Internet

Socket Timeout determines time (in seconds) to wait remote server response while downloading. If server stalls for more then this time, connection breaks. Default - **300** seconds (5 minutes).

Number of Connections. **Web Downloader** can load several files (from 2 to 100) simultaneously. Default - **5** files at once.

Retry if server unavailable. With this option checked **Web Downloader** will retry if connection to remote Web server fails or downloading process breaks unexpectedly. This could be useful on unreliable Internet connections.

3. Proxy

Proxy type determines whether you connect to Internet via Proxy server or not. If you have proxy, check what type of it do you have. HTTP proxy usually works slightly faster then SOCKS4 proxy.

Proxy address and **Proxy Port.** Set these values according to your Proxy server settings.

After setting up all necessary options you may start using this program. See [How-To-Use Instructions](#) now.

How-To-Use Instructions

Everything here is simple.

You will see **folders section** on the top left and **Items section** on the top right on program's window. This section contains list of items for downloading. It divides in several sections: URL, Image load specifier (Yes/No), Download Level, Load another servers specifier (Yes/No), Download status and Number of files in current download queue.

Below is the **Connections status list**. Here you can see download progress. Number of lines in this list corresponds to **Number of Connections** value in **Options** dialog. Each line has prefix **HTTPx:** where x begins from 0.

If you have checked **Log Connections** box in **Options** dialog you will see **Connections Log** on the bottom. Contents of each line is the same as in previous section. The only notable exception is - new lines does not substitute old ones, but appends to the end of list.

There are splitters between sections. Press left mouse button and drag splitter up or down to resize each section.

To add new URL for download, select desired Folder in **Folders section** and then **Items/Add** menu item or press **Ins** button or right-click mouse button on **Items section** and select **Add**. See **Adding/Modify Dialog** for details.

To download desired item, select it in the **Items section** and **Execute** it (**Items/Execute** menu item or **F5** button). After this, status will change to **Executing...** After download it will be set to **Done**. In **Files in Queue** you will see how many files wait for download. This number could decrease or increase depending on remote Web site structure and **Download Level** you specified.

While downloading, **WebDownloader** builds a copy of remote Web server(s) in **Download Directory**. For example, you want to get URL: `http://www.server.com/page.html`. Program will automatically create all needed directories, starting with **www.server.com**. Links of type `http://www.server.com....` and `/images/pict.gif` it will translate to `/servers/www.server.com/...` and `/servers/www.server.com/images/pict.gif` (with **Translate links for Web server** option). **Web Downloader** will translate the ":" signs (remote Web server port redirector in URL) to "+" sign. URLs ending with "/" (default document) will be `...../default.htm` (you may change it in **Default document** value in **Options** dialog). Files "**descr.dat**" and with ".**primary**" extension are important - don't touch them.

While downloading you will see all channels with progress messages. These messages look like **Netscape Navigator's** messages.

To add or remove folders use **Items/Add Folder** or **Items/Remove Folder** menu items or right-click mouse button on **Folders section** and select needed item from popup menu. You may drag and drop Items from **Items section** onto different folders or another items to move them.

You may automatically begin downloading on program startup. To do this specify numbers (starting from **0**) of desired URLs in command line. Numbers should be separated by spaces. For example, if you have to download first, third and fourth items from URLs list in **Items section**, then you may start program with:

WebDown.exe 0 2 3

If you have several folders, **Web Downloader** will start automatic download from the first folder. To download Item from another folder, put **Folder/Number** to command line. For example to download second item from “**Auto**” folder, start program with: **WebDown.exe Auto/1**

Note: **Web Downloader** will NOT quit after automatic download.

Add Modify Dialog

Here are several parameters divided in three groups:

1. Parameters

URL means location in the Internet of desired Web page.

Check **Load images** if you want to download images, animations, ActiveX components, Java classes, etc.

Do not check **Load another servers** if you want to download files only on desired Web server.

If **Download only this way and down** option checked **Web Downloader** will load pages only below path you've specified in **URL**.

Level means how many Web pages will be downloaded.

Set **Translate Links for** to **Web server** if you plan to publish downloaded pages on your Web server or to **Offline browsing** if you will browse them from your local hard disk.

2. Scheduling

Use these options to start automatic download at specified time. **Manual** means no scheduling.

3. Authentication

Put username and password here to download password protected Web pages. Only the **Basic** authentication scheme is currently supported.

Seek For Permanent Job

I seek for permanent job outside Russia.

I'm 24 years old, married, have 8 months old daughter. I've finished Mordovian State University in Russia. I have worked with IBM PC - compatible computers for 7 years. Borland Pascal / Borland Delphi developer. Currently work at Saransk Power Station as programmer and system administrator of our network (Windows NT, Linux). Have experience in upgrading and testing computers, TCP/IP, client-server programming, know HTML, CGI, ISAPI, studying Java, JavaScript, VBScript now.

Oleg Chernavin.

E-mail: arny@saransk.sitek.net

22 April 1997.

Enter URL for downloading here.

Examples:

<http://www.infoart.ru/it/news/index.htm>

<http://www.somehost.com/>

Check this box if you want to load images on this page automatically.

Defines how many HTML pages would be downloaded.

With **Level = 0** program will download only page specified by URL field.

With **Level = 1** - the specified page and all pages this page refers to.

With **Level = 2** - everything with **Level 1** plus every page linked to downloaded ones.

And so on...

If this option checked, **Web Downloader** will load all URLs, even if they located on another servers.
Uncheck it, if you want to load **only** specified server.

With this option checked **Web Downloader** will load pages only below path you've specified in **URL**.

You can select whether download Web pages for Web server or offline browsing. This affects on links translating.

For Web server:

http://www.server.com/... /servers/www.server.com/...;

/images/image.gif /servers/www.server.com/images/image.gif

For Offline browsing:

http://www.server.com/... file:///c:\temp\www.server.com/...;

/images/image.gif file:///c:\temp/www.server.com/images/image.gif

So you can open this file location in your Web browser and enjoy WWW serfing from your hard disk.

Put username and password here to download password protected Web pages. Only the **Basic** authentication scheme is currently supported.

Manual means no scheduling.

Daily mode start download every day at specified time.

Weekly start download one time every week on **Weekday** at specified time.

Monthly and **Yearly** do the same one time every month or year.

Specify directory, where downloaded files would be saved.

Please, enter backward slash “\” at the end.

Specifies virtual directory on your Web server for downloaded files.

Every link in downloaded HTMLs with "http://www.server...." at the beginning would be converted to "**/servers**/www.server....", where "**/servers**" is contents of this line.

Slashes "/" at the beginning and end of this line are recommended.

Default document determines how to name files when there is no filename for downloaded page. For example, URL like “http://www.server.com/” and “http://www.server.com/News/” contain no file name. **Web Downloader** will create files with name specified here. Default - “**default.htm**”.

To log every HTTP transaction while downloading, check this box. To clear this **EXPANDING** log, use File/Clear Log in main menu.

With this option checked **Web Downloader** will retry if connection to remote Web server fails or downloading process breaks unexpectedly. This could be useful on unreliable Internet connections.

Select appropriate **Proxy type** for your proxy server. If you choose HTTP or SOCKS4 proxy you must fill **Proxy Address** and **Proxy Port** fields also.

Currently **Web Downloader** supports only HTTP and SOCKS v.4 Proxies. This option was tested with **WinGate 1.3.17**. Also HTTP Proxy is implemented in **Netscape Proxy Server**.

After waiting **Socket TimeOut** seconds with no response, **WebDownloader** breaks connection and moves current URL to the end of download queue. Default value is **300** (5 minutes).

You can open more than one connection at a time to Internet Web Servers. More connections mean more simultaneous files, but can also slow down the speed of each individual connection. Default value - **5**.
Range - from **2** to **100** connections at a time. Changes take effect after program reload.

Program History

4 Feb 1997 - Version 1.00.20

Initial release.

12 Feb 1997 - Version 1.00.21

Fixed URL Port Redirection (':' symbol).

Fixed URL downloading with '?' symbol.

When loading frames all of them treated as the same download level.

Added **Stop Download** to Main menu.

Added **Number of Connections** Value to **Options** Dialog.

Improved User Interface.

6 Mar 1997 - Version 1.00.22

Two fixes of URL Port Redirection (':' symbol)

Added **Load Another Servers** option to **Add/Modify Item** Dialog.

14 Mar 1997 - Version 1.00.23

Added **Command Line** options.

Added **Files in Queue** indicator to the upper list.

Socket Timeout value in **Options** Dialog changed to **seconds**.

Calls to **GetHostByName** and **Connect** are now non-blocking.

HTML translation is now faster.

Program creates **.primary** files only for HTMLs.

18 Mar 1997 - Version 1.00.24

Added **Translate links for...** option to **Add/Modify Item** Dialog.

Now remembers main window size and position (stored in Registry).

4 Apr 1997 - Version 2.0 Build 27

Added **Retry if server unavailable** to **Options** Dialog.

Added Outline view.

Added **Default Document** to **Options** Dialog.

9 Apr 1997 - Version 2.0 Build 28

Added **Scheduling** options to **Add/Modify Item** Dialog.

Added **SOCKS4 Proxy** support.

11 Apr 1997 - Version 2.0 Build 30

Added **Username/password** authentication for protected Web pages (in **Add/Modify Item** Dialog).

Added **Download only this way and down** option to **Add/Modify Item** Dialog.

Fixed strange behavior when **Load Images** option unchecked.

23 Apr 1997 - Version 2.0 Build 31

Bug fixes.

Help File rewritten.

Added support for another **9** HTML tags, including ActiveX and Java classes.

Added **302 Object moved** status support.

30 Apr 1997 - Version 2.0 Build 32

Bug fixes (Authentication and Files retrieving on some Unix Web servers).

