### **Browser Support**



If you see the navigation tree on the left and the information panel on the right, then that means that your browser supports all functionality this data repository provides.

Browsers (that we know of) that do work include:

- Google Chrome
- Firefox
- Opera
- Microsoft Edge

If you do not see the navigation tree on the left and the information panel on the right, then that means that your browser does not support all functionality this data repository provides.

Browsers (that we know of) that do not work include: ➤ Internet Explorer

### Downloading Data

- Since the data can be retrieved using HTTP protocol, using something like wget will be the main avenue of data retrieval.
- The following tools will be looked at in this document:
  - Command line wget for Linux and Windows
  - VisualWget for Windows

The following pages will focus on showing how to download an entire directory.

**IMPORTANT:** Scripts that run too aggressively may result in the server banning the originating host for 10 minutes (you know this happened if the website seems to be "down"). You can alleviate this by ensuring to space HTTP requests such that you do not exceed more than 3 or 4 requests per second.

### Command Line wget for Windows

To get a version of GNU Wget that works with Windows you can go to: <a href="http://gnuwin32.sourceforge.net/packages/wget.htm">http://gnuwin32.sourceforge.net/packages/wget.htm</a>

You want to download the "Complete package, except sources" file.

Description	Download	Size	Last change	Md5sum
• Complete package, except sources	<u>Setup</u>	3012464	31 December 2008	b4679ac6f7757b35435ec711c6c8d912
Sources	<u>Setup</u>	1270850	31 December 2008	25cb2d164024b3e478c4ab14b89585a9
	2.5			
Binaries	Zip	850448	31 December 2008	254b95bd96564eb6db590f2b51f8fd8b
Documentation	Zip	1244187	31 December 2008	acdf67b1d524526199e51a6224c4071f
Sources	Zip	2189442	31 December 2008	ef29b3c8b5708825ca008a3938adb8fd
Dependencies	Zip	1443871	4 September 2008	b23ce98bc2d03f852f8516183c619d42
Original source	http://ftp.gm	1.org/gnu/v	vget/wget-1.11.4.tar.g	12

You can also download the files from the GnuWin32 files page. New releases of the port of this package can be monitored.

### Downloading Data: Command Line wget

#### \$ wget -w 0.5 -l 0 -r -nH -np -I /models/\*,/models -R \*.html,\*.htm files.ntsg.umt.edu/models

The above command will download everything within the "models" folder to "./models".

These are arguments for wget that we recommend to use:

Argument	Description
-w 0.5	<b>OPTIONAL:</b> Wait 0.5 seconds between each retrieval. This lessens the load for the server, and decreases the chance that the server will ban the computer that ran this command. However, this is not strictly necessary especially if the files that are being downloaded are large and thus the time between file retrievals is already spaced out.
-10-г	Retrieve items recursively and specify the maximum recursion depth to be infinity.
-nH	Disable the generation of host-prefixed directories. (e.g/files.ntsg.umt.edu).
-np	Do not ever ascend to the parent directory when retrieving recursively. This guarantees that only files below a certain hierarchy will be downloaded.
-I /models/*,/models	This guarantees that all directories within models and all files within models will be included in the retrieval. The /models/* argument specifies to include all directories within /models (e.g. /models/otterbgc), and the /models argument specifies to include all non-directory files within /models (e.g. /models/README.HELP).
-R *.htm, *.html	Do not download html files. These are only useful for web browsers, and does not contain the data we need.
files.ntsg.umt.edu/models	The URL of the website that will be traversed. Note that you need to include the full URL of the directory that you wish to download.

### Downloading Data: Command Line wget

### Downloading single folder from MOD16 example

 $\$  wget -l 0 -r -nH -np -I  $\setminus$ 

> /data/NTSG\_Products/MOD16/MOD16A2.105\_MERRAGMA0/Y2001/\*,/data/NTSG\_Products/MOD16/MOD16A2.105\_MERRAGMA0/Y2001 \
> -R \*.html,\*.htm files.ntsg.umt.edu/data/NTSG\_Products/MOD16/MOD16A2.105\_MERRAGMA0

Here we are downloading data from the Y2001 from the /data/NTSG\_Products/MOD16/MOD16A2.105\_MERRAGMAO/ URI. (This is all one line, and the \ at the end allows a user in bash to press ENTER without executing the command)

### Downloading multiple folders from MOD16 example

 $\$  wget -l  $\odot$  -r -nH -np -I  $\setminus$ 

> /data/NTSG\_Products/MOD16/MOD16A2.105\_MERRAGMAO/Y2001/\*,/data/NTSG\_Products/MOD16/MOD16A2.105\_MERRAGMAO/Y2001, \

> /data/NTSG\_Products/MOD16/MOD16A2.105\_MERRAGMAO/Y2002/\*,/data/NTSG\_Products/MOD16/MOD16A2.105\_MERRAGMAO/Y2002 \

> -R \*.html,\*.htm files.ntsg.umt.edu/data/NTSG\_Products/MOD16/MOD16A2.105\_MERRAGMAO

Here we are downloading data from the Y2001, Y2002 from the /data/NTSG\_Products/MOD16/MOD16A2.105\_MERRAGMAO/URI.

(This is all one line, and the  $\$  at the end allows a user in bash to press ENTER without executing the command. Although there is a line break after Y2001 for illustration purposes, all of the folder listings in -I should be on a single line.)

View the wget docs here: <a href="https://www.gnu.org/software/wget/manual/wget.html">https://www.gnu.org/software/wget/manual/wget.html</a>

There is a special utility for Windows called VisualWget that allows you to schedule wget jobs via a graphical user interface. To download this, you can go to https://sites.google.com/site/visualwget/a-download-manager-gui-based-on-wget-for-windows.

You want to download the "vwget-2.5a2-wget-1.11.4-bin.zip" file.

vwget-2.3-lang.zip (	6k)	Khomsan Phongphisansakun, May 8, 2012, 5:57 A	v.1	+
vwget-2.4-lang.zip (	4k)	Khomsan Phongphisansakun, Jan 11, 2013, 10:31	v.4	ŧ
vwget-2.4-src.zip (18	59k)	Khomsan Phongphisansakun, May 8, 2012, 5:55 A	v.1	+
vwget-2.4-wget-1.10	.2-bin.zip (958k)	Khomsan Phongphisansakun, May 8, 2012, 5:54 A	v.1	+
vwget-2.4-wget-1.11	.4-bin.zip (962k)	Khomsan Phongphisansakun, May 8, 2012, 5:54 A	v.1	+
vwget-2.5a2-src.zip	1862k)	Khomsan Phongphisansakun, May 8, 2012, 5:56 A	v.1	+
vwget-2.5a2-wget-1	10.2-bin.zip (961k)	Khomsan Phongphisansakun, May 8, 2012, 5:56 A	v.1	ŧ
vwget-2.5a2-wget-1	11.4-bin.zip (965k)	Khomsan Phongphisansakun, May 8, 2012, 5:55 A	v.1	Ŧ

- Extract VisualWget (In this example, the file vwget-2.5a2-wget-1.11.4-bin.zip was extracted to Documents\VWget)
- Double-click on VisualWget.exe

📕   🛃 📑 🖛   VWget								>
File Home Share	View						~	/
$\leftarrow \rightarrow \checkmark \uparrow $ -> This P	C > Documents > VWget			~	Ō	Search VWget		۶
🔹 Quick access	Name	Date modified	Туре	Size				
	Lang	12/19/2016 5:17 PM	File folder					
a OneDrive	📄 cacert.pem	1/2/2008 11:12 AM	PEM File	501 KB				
This PC	DownloadList.cfg	12/19/2016 5:18 PM	CFG File	0 KB				
- mare	libeay32.dll	6/27/2010 7:31 PM	Application extens	1,202 KB				
i Network	libssl32.dll	6/27/2010 7:31 PM	Application extens	244 KB				
	License.txt	11/10/2008 6:47 AM	Text Document	3 KB				
	OptionPresets.cfg	9/2/2010 10:11 PM	CFG File	0 KB				
	UserAgents.cfg	7/26/2010 7:43 PM	CFG File	1 KB				
	VisualWget.cfg	12/19/2016 5:18 PM	CFG File	1 KB				
	🤎 VisualWget.exe	11/27/2010 10:13	Application	218 KB				
	🖓 VisualWget.exe.config	8/31/2010 5:27 PM	XML Configuratio	1 KB				
	wget.cfg	3/7/2009 5:41 PM	CFG File	1 KB				
	📧 wget.exe	8/22/2010 9:08 PM	Application	246 KB				
	📄 wget.opt	7/27/2010 5:54 PM	OPT File	11 KB				



In this example, we are going to download the files.ntsg.umt.edu/models directory.

- 1. Click on the icon with a plus sign in the upper left hand corner to start a new job.
- 2. In <u>General > URLs</u>, put files.ntsg.umt.edu/models
- 3. In <u>General > Save files to</u>, put a where you would like to save the retrieved files.
- 4. In <u>HTTP > User agent</u>, put Chrome. (Or whatever browser you want and have installed- this allows wget to emulate a browser connection)
- 5. In <u>Advanced > Directory</u>, ensure that --no-host-directories is checked.
- 6. In <u>Advanced > Recursive Retrieval</u>, ensure that:
  - --recursive is checked
  - --level is checked and 0 is entered for the NUMBER value
- 7. In <u>Advanced > Recursive Accept/Reject</u>, ensure that:
  - --reject is checked and \*.htm,\*.html is entered for the LIST value
  - --include-directories is checked and /models/\*,/models is entered for the LIST value
  - -- no-parent is checked
- 8. Press OK. If "Start this download when I click OK" is checked, then the job will start right away. If it's not, then you have to right click the job and press "Start" to start it.



General HTTP Advanced	URLs: files.ntsg.umt.edu/models Save files to: E-VMTSC Files	General HTTP Advanced V Load Empty Save As Delete			General HTTP Advanced	Presets  Load Empty Save As Delete				
	Resuming         Resume getting a partially-downloaded file         Don't re-retrieve files unless newer than local         Scheduler         Start automatically every         O \$\overline\$ days         Write documents to:	Be Lo Dy H H FT Re Re	Basic Startup .ogging and Input File Download Directory HTTP (SSL/TLS) TP Recursive Retrieval Recursive Accept/Reject	io-directories orce-directories io-host-directories irotocol-directories lirectory-prefix=PREFIX ut-dirs=NUMBER			Basic Startup Logging and Input File Download       ✓ -recur Image: -recursive start         Directory      conve         HTTP      back         HTTPS (SSL/TLS)      mirroi         FTP      page         Recursive Retrieval      strict	rsive I=NUMBER te-after vert-links up-converted or e-requisites t-comments		
	Note:	De	Don't create host directories.				Maximum recursion depth (inf or 0 for i	infinite).		
Start this download when I click OK OK Cancel		Start this download when	n I dick OK		OK Cancel	Start this download v	when I didk OK		ОК	Cancel

General HTTP Advanced	User agent: Chrome 5.0	General HTTP Advanced	Presets	General HTTP Advanced	Presets
	Referer:		Load Empty     Save As     Delete       Basic Startup Logging and Input File Download Directory    accept=LIST     Image:no-parent       Directory    closting =LIST       HTTP    exclude-domains=LIST       HTTPS (SSL/TLS)    follow-ftp       FTP    follow-tags=LIST       Recursive Retrieval    relative       Recursive Accept/Reject    exclude-directories=LIST      exclude-directories=LIST    exclude-directories=LIST		Load Empty     Save As     Delete       Basic Startup Logging and Input File Download    reject=LIST    no-parent       Directory    reject=LIST    no-parent       HTTP    comains=LIST    reject=LIST       HTTPS (SSL/TLS)    follow-ftp    follow-ftp       FTP    follow-ftp    span-hosts      relative    relative    relative       Image: Comparison of the state of th
			Comma-separated list of rejected extensions. LIST: *.htm,*.html		List of allowed directories. LIST: /models/*,/models
Start this download w	hen I dick OK OK Cancel	Start this download w	hen I dick OK Cancel	Start this download w	when I dick OK OK Cancel