



MyBox: Easy Tools Set **User Guide - Network Tools**

Author: Mara

Date: 2019-4-20

Contents

1 Resource Addresses.....	3
2 Documents.....	3
3 Menu of Network Tools.....	4
4 Web Editor.....	5
4.1 Edit page.....	5
4.2 Web Codes.....	6
4.3 Web Browser.....	7
4.4 Web Snapshot.....	8
5 Weibo Snapshot Tool.....	9
5.1 Why need this tool?.....	9
5.2 What it can do?.....	10
5.3 What it can not do?.....	10
5.4 Parameters.....	10
5.5 Principle of the Tool.....	12
5.5.1 Major problems need be solved.....	12
5.5.2 How to access the public data of WeiBo account?	12
5.5.3 The keys in weibo pages.....	12
5.5.4 Main Logic of the Tool.....	14
5.5.5 Implementation.....	14

1 Resource Addresses

This is GUI(Graphic User Interface) program developed in JavaFx, whose target is to provide simple and easy tools . It is free and open sources, and its main page is following:

<https://github.com/Mararsh/MyBox>

Source codes, compiled packages, and documents are under Releases directory:

<https://github.com/Mararsh/MyBox/releases>

Welcome to submit software requirements and problem reports online:

<https://github.com/Mararsh/MyBox/issues>

Cloud storage:

https://pan.baidu.com/s/1fWMRzym_jh075OCX0D8y8A#list/path=%2F

2 Documents

This document introduces Network Tools of MyBox. It can be download from following address:

<https://github.com/Mararsh/MyBox/releases/download/v5.0/MyBox-UserGuide-5.0-NetworkTools-en.pdf>

Other documents include:

“MyBox User Guide – Overview”

<https://github.com/Mararsh/MyBox/releases/download/v5.0/MyBox-UserGuide-5.0-Overview-en.pdf>

“MyBox User Guide – Image Tools”

<https://github.com/Mararsh/MyBox/releases/download/v5.0/MyBox-UserGuide-5.0-ImageTools-en.pdf>

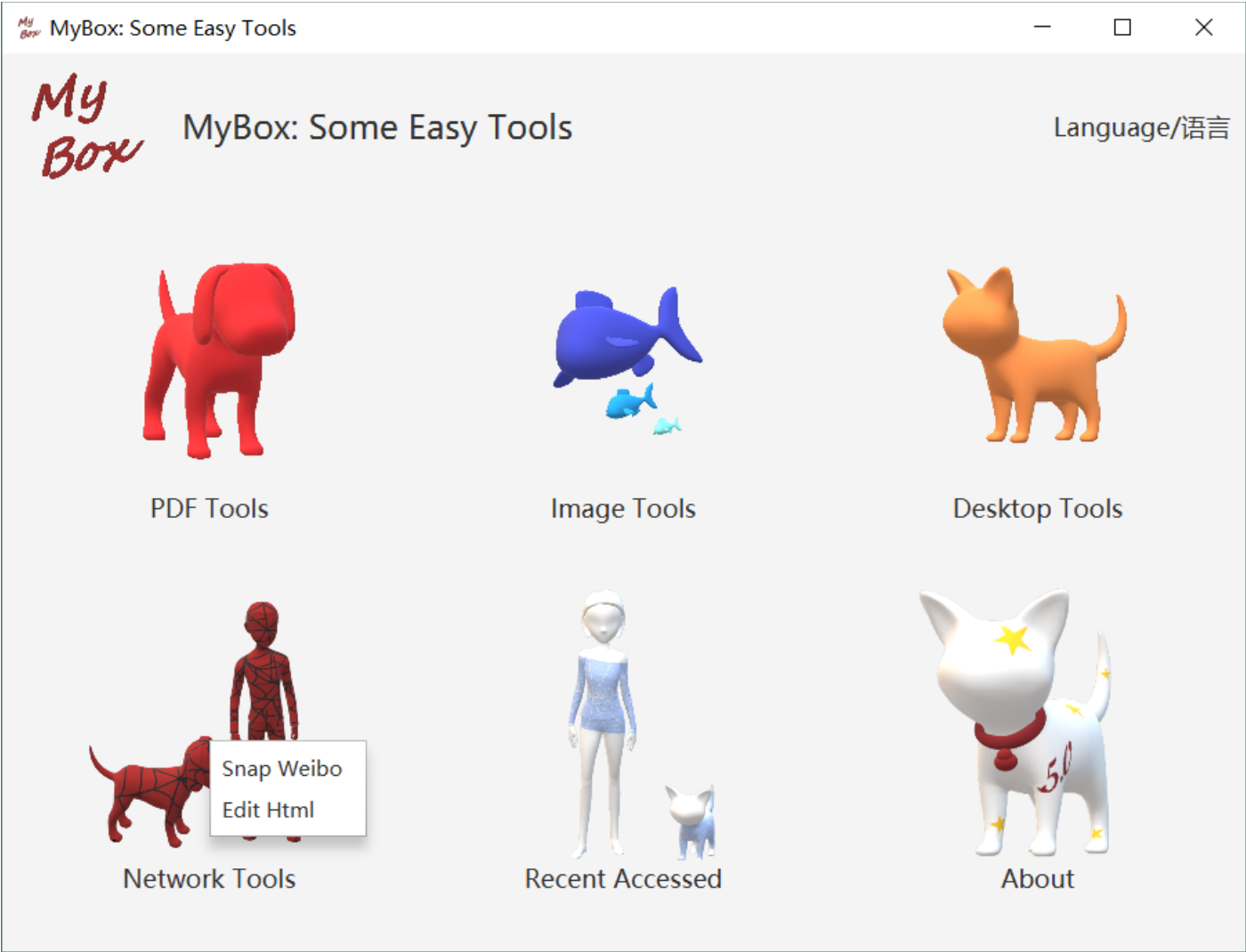
“MyBox User Guide – Desktop Tools”

<https://github.com/Mararsh/MyBox/releases/download/v5.0/MyBox-UserGuide-5.0-DesktopTools-en.pdf>

“MyBox User Guide – PDF Tools”

<https://github.com/Mararsh/MyBox/releases/download/v5.0/MyBox-UserGuide-5.0-PdfTools-en.pdf>

3 Menu of Network Tools



4 Web Editor

4.1 Edit page

- 1) Open, new, save, save as html file.
- 2) Edit the page file in rich text. Set color, font, paragraph with toolbar. What you see is what you get.
- 3) Synchronize the updates in page editor to page code.

MyBox Web Editor C:\Users\mara\mybox\ImageHelp.html*

Window PDF Tools Image Tools Desktop Tools Network Tools Settings Help

Web Editor Web Codes Web Browser

File Create Save Save As

段落 Blackadder ... 10 磅 B I U

MyBox Main Page

Supported Compression Types of Image Format "bmp"
From API of class
"com.github.jaiimageio.plugins.bmp.BMPImageWriteParam"

Type String	Description	Image Types
BI_RGB	Uncompressed RLE	<= 8-bits/sample
BI_RLE8	8-bit Run Length Encoding	<= 8-bits/sample
BI_RLE4	4-bit Run Length Encoding	<= 4-bits/sample
BI_BITFIELDS	Packed data	16 or 32 bits/sample
BI_JPEG	JPEG encoded	grayscale or RGB image

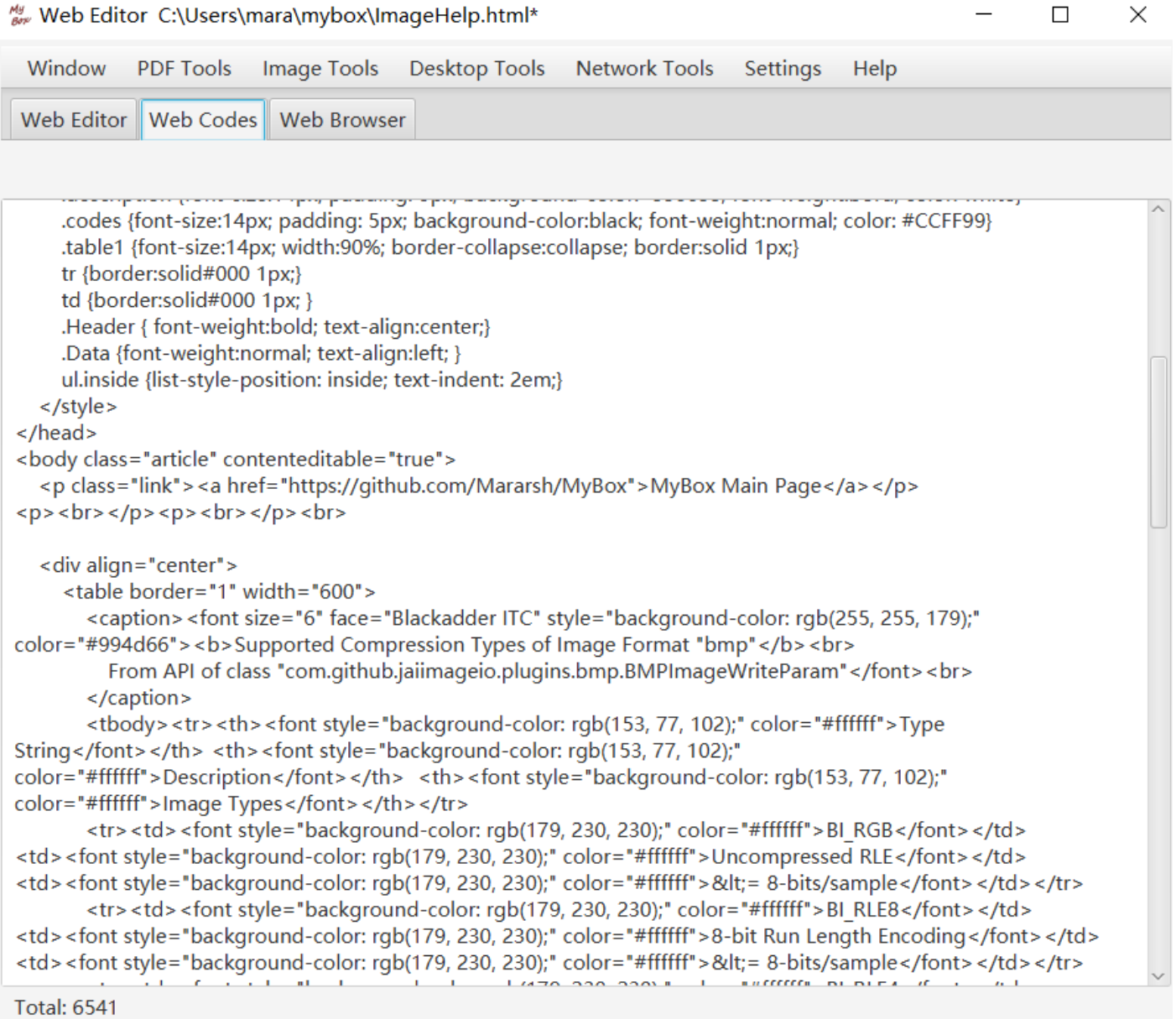
Supported Compression Types of Image Format "tif"/"tiff"
 From API of class "com.github.jaiimageio.plugins.tiff.TIFFImageWriteParam"

Compression Type	Description	Reference
CCITT RLE	Modified Huffman compression	TIFF 6.0 Specification, Section 10

Total: 5364

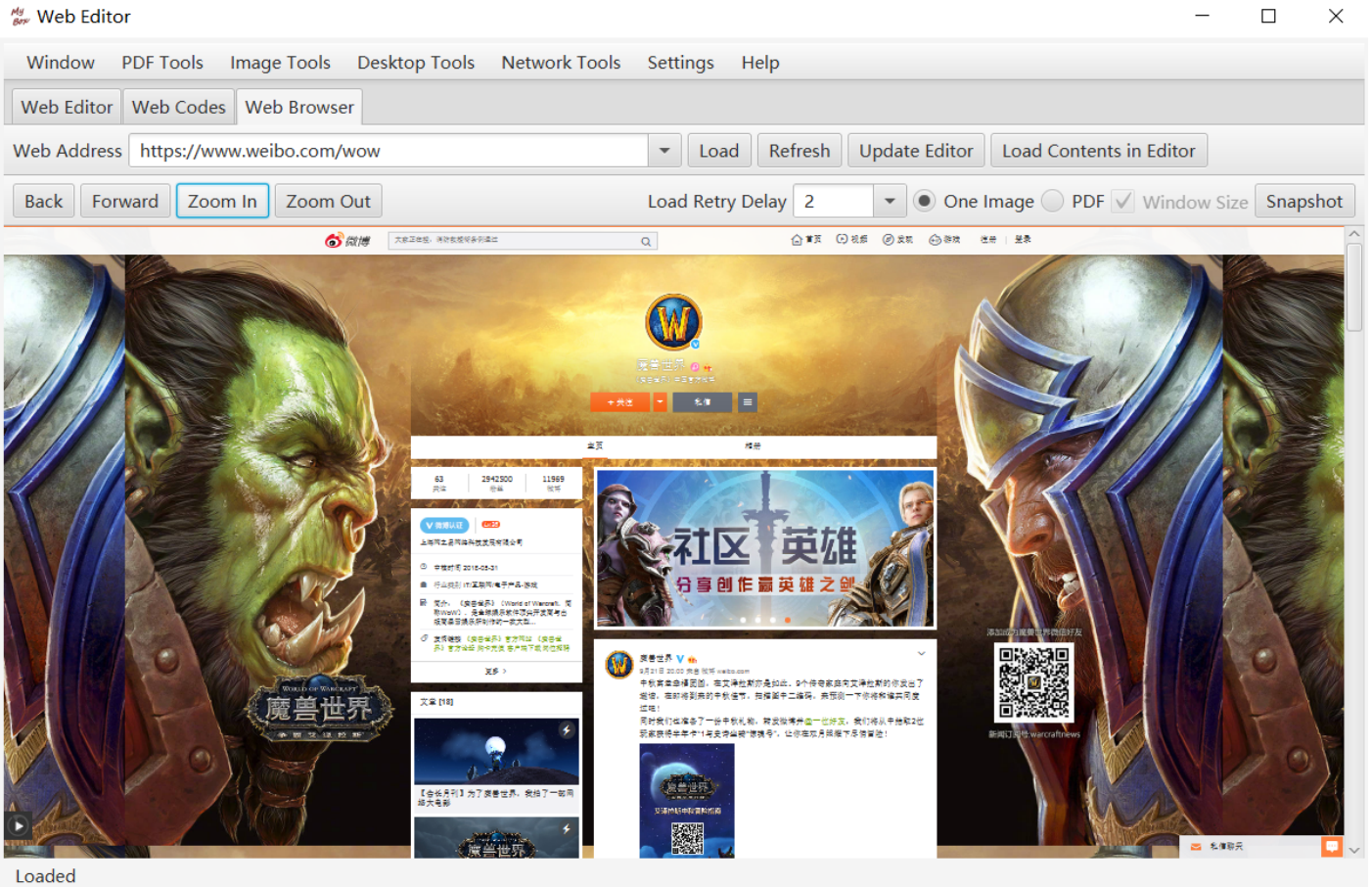
4.2 Web Codes

- 1) View and edit the codes of html page
- 2) Synchronize the updates between page editor and code editor.



4.3 Web Browser

- 1) Input URL address and load the html page.
- 2) Synchronize the browser's contents to html editor.
- 3) Load the contents of html editor in the browser.
- 4) Zoom in/out and back/forward the page.



4.4 Web Snapshot

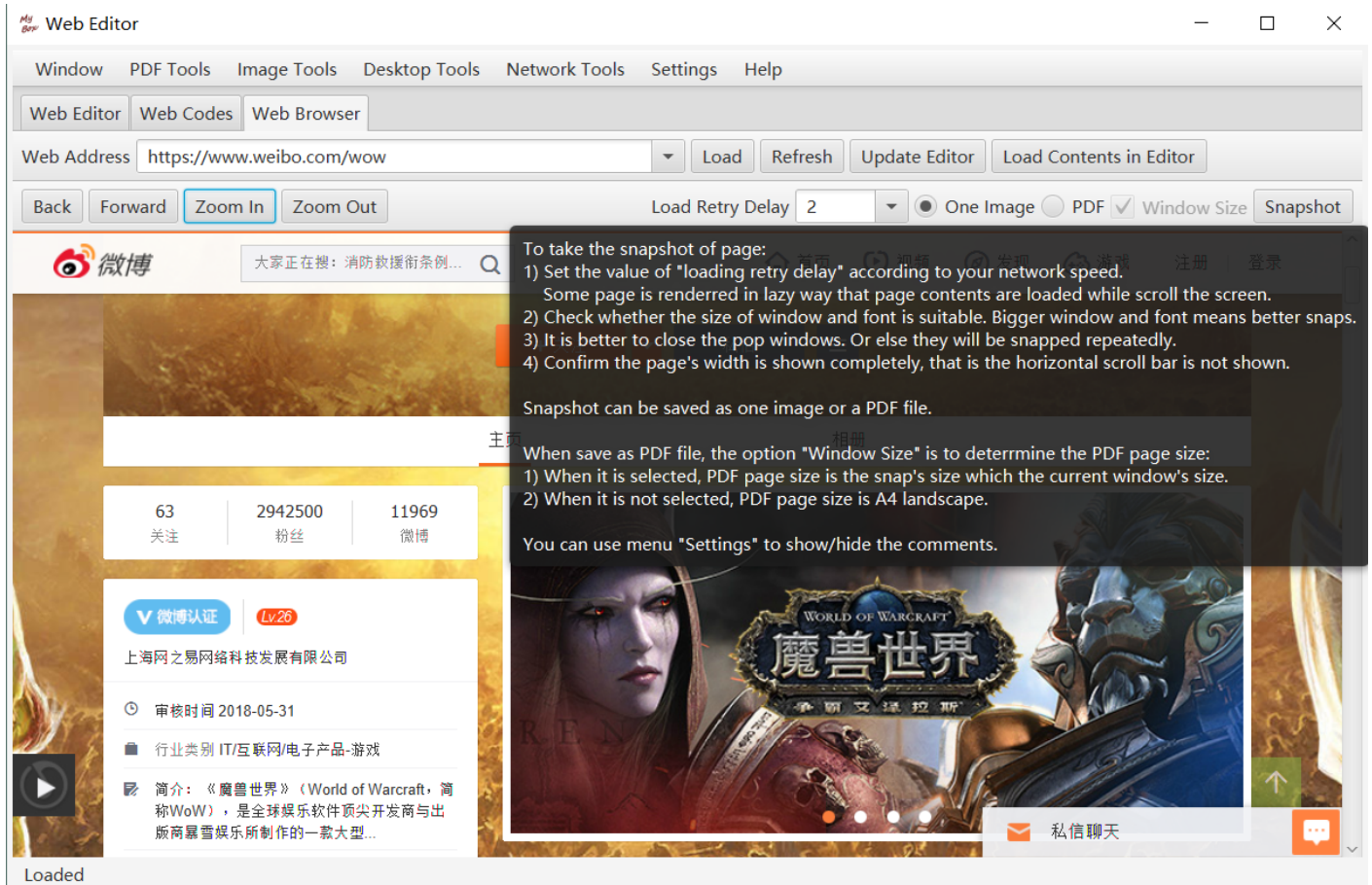
To take the snapshot of page:

- 1) Set the value of "loading retry delay" according to your network speed. Some page is rendered in lazy way that page contents are loaded while scroll the screen.
- 2) Check whether the size of window and font is suitable. Bigger window and font means better snaps.
- 3) It is better to close the pop windows. Or else they will be snapped repeatedly.
- 4) Confirm the page's width is shown completely, that is the horizontal scroll bar is not shown.

Snapshot can be saved as one image or a PDF file.

When save as PDF file, the option "Window Size" is to determine the PDF page size:

- 1) When it is selected, PDF page size is the snap's size which the current window's size.
- 2) When it is not selected, PDF page size is A4 landscape.



5 Weibo Snapshot Tool

5.1 Why need this tool?

Users of Weibo(<https://weibo.com>) may want to backup the contents of their accounts, or may want to save the contents of some accounts of others who they care. Save the pages manually one by one? Nonsense. The contents of Weibo pages are loaded dynamically and the pages can not be rendered correctly after they are saved locally. Currently it looks no official function for users to backup contents of Weibo accounts automatically.

This tool is very simple. It opens browser, load the pages, snap the contents, and save as files, all by codes automatically.

The screenshot displays the MyBox Weibo Snapshot Tool interface. The left pane shows a list of Weibo posts from 2012-1-18. The right pane shows a 'Snaping Page...' status window with various statistics and a list of saved files. The bottom pane shows a preview of a Weibo post with a game screenshot and a '私信聊天' button.

Snaping Page...

Weibo Main Page Address: <https://weibo.com/wow>
 Account: 魔兽世界的世界
 First Weibo Month: 2011-03
 Last Weibo Month: 2018-10
 Snap Duration: 2012-01 ~ 2012-01
 Number of Months To Be Snapped: 1
 Completed Months: 0
 Current Loading Month: 2012-01
 Pages Number of This Month: 2
 Current Loading Page: 1
 PDF files Saved: 0
 Html files saved: 1
 Pictures Saved: 39
 Pictures Number in This Page: 39
 Page Zoom Scale: 1.0
 Current Window Height: 943
 Current Window Width: 654
 Snaping for PDF file: D:\tmp\9\魔兽世界的世界\2012-pdf\魔兽世界的世界-2012-01-第1页.pdf
 Pictures Number in This Page: 39
 Snaping Start Time: 2018-10-26 20:06:33 (Elapsed Time: 1 minutes)
 Current Page Height: 68999
 Current Snapshot Number: 37

MyBox Available Memory:1801MB Required Memory:400MB(4.94%) Used Memory:229MB(2.83%)
 Windows 10 Total Physical Memory:8102MB Used Physical Memory:7002MB (86.42%)

Open Path MyBox Reload Current Page Stop

5.2 What it can do?

- 1) Save contents of any months of any Weibo accounts automatically.
- 2) Set the months range.
- 3) Make sure contents of pages have been loaded completely,
- 4) Expand all comments and pictures in the pages.
- 5) Save the pages as local html files which are for extracting texts but not for viewing the pages due to contents loaded dynamically.
- 6) Snap the pages as PDF files, with options like page size, margins, author, and image format.
- 7) Save original size pictures in the pages.
- 8) Display the progress in time.
- 9) Stop the progress at any time. The interrupted month will be record and inputted as start month for next execution.
- 10) Set retry times in case of broken or bad network.

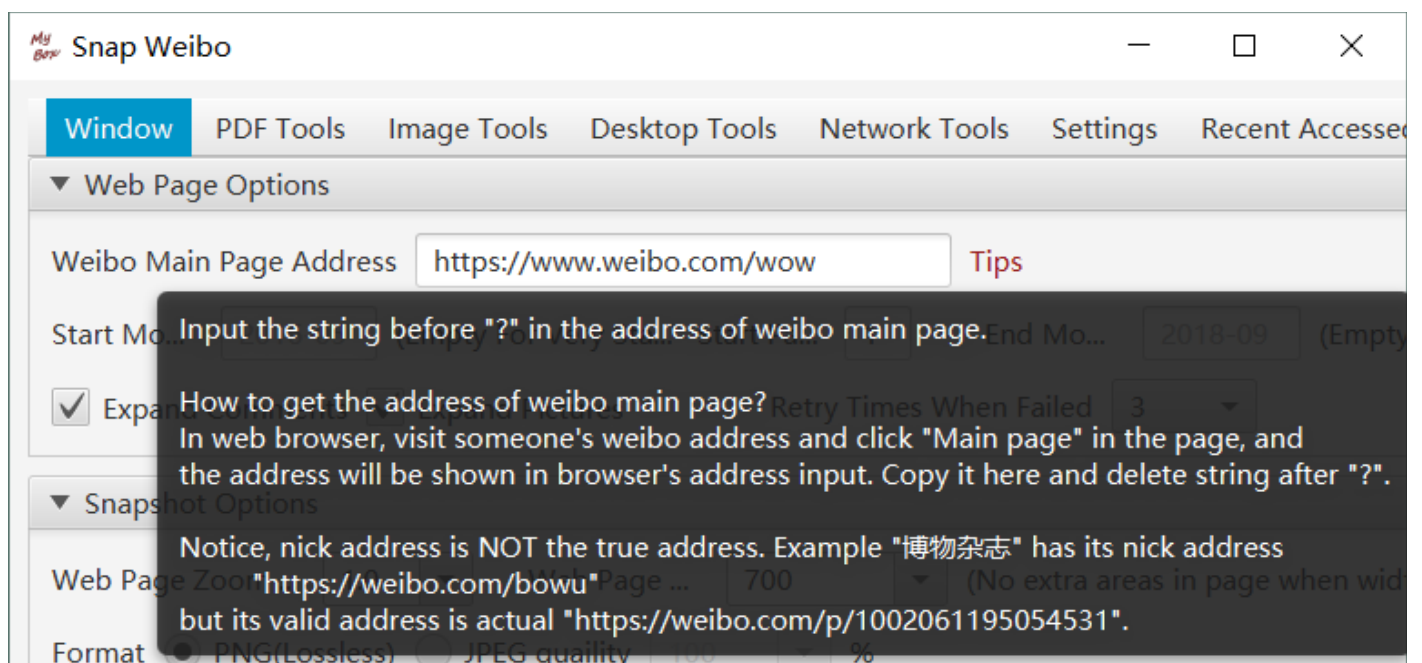
5.3 What it can not do?

Full contents of long message can not expanded since such operation need login.

5.4 Parameters

Except for “Main address” and “Target files path”, other parameters are fine with default settings.

Move mouse upon the red label “Tips”, and information about “Weibo address” will be popped.:



Snap Weibo

Window

PDF Tools
Image Tools
Desktop Tools
Network Tools
Settings
Recent Accessed
Help

▼ Web Page Options

Weibo Main Page Address
Tips

Start Month
(Empty For Very Start)
Start Page
End Month
(Empty For Now)

☒ Expand Comments
☒ Expand Pictures
Retry Times When Failed

▼ Snapshot Options

Web Page Zoom Scale
Web Page Width
(No extra areas in page when width is suitable)

Format
☒ PNG(Lossless)
☐ JPEG quality
 %

☐ Black White CCITT group 4 (T6) Threshold
(Empty for Default)
☐ Dithering

▼ Pdf Options

Page Size
☒ Images Size
☐ Standard Size

☐ Custom (Pixels)

Page Margin
Author
Default Display Scale
 %

Maximum Main Memory for PDF
☒ 500M
☐ 1G
☐ 2G
☐ Unlimit

▼ Target Files

Files Path
Select...
Category Files
☒ In Months Paths
☐ In One Path

☒ Save HTML Files
☒ Save Pictures
☒ Save Snapshots in PDF files
☒ Keep Each Page's PDF

☒ meow
☒ Open Path When Stop
☒ Close This Window After Started
Suggested Settings
Wow As Example
Start

Prompt from GuaiGuai when completed or failed

5.5 Principle of the Tool

5.5.1 Major problems need be solved

- 1) How to access Weibo pages without login?
- 2) How to setup SSL certificate of Weibo site in browser on Windows/Linux/Max?
- 3) How to determine the contents of page have been completely loaded?
- 4) How to determin all pictures in the page have been expanded? How to skip borken addresses?
- 5) How to avoid losing contents due to popup windows on the page?
- 6) How to find the correct addresses of original size pictures in the page?

5.5.2 How to access the public data of WeiBo account?

- 1) The addresses of each month histories of a weibo account are in the right vertical bar of its main page. They can be visited directly without login.
- 2) Following is example to visit the contents of first page of account “月球车玉兔” in 2013 December:

https://weibo.com/u/3926428816?is_all=1&stat_date=201312&page=1

5.5.3 The keys in weibo pages

- 1) Search “stat_date” in main page, and the list of all valid months can be found. Like:

```
<a href="javascript:void(0);" class="S" action-type="login" action-data="is_all=1&amp;stat_date=201807"
suda-uatrack="key=Profile_V6_Timeline&amp;value=month"><em class="bor_t"></em><em
class="S_dot"></em><em class="bor_b"></em><span>7 月</span></a>
```

- 2) The status of page:

- (1) If page includes “还没有发过微博” , this months has not valid data.
- (2) If page includes “查看更早微博” , this page has been loaded completely and it is the last page of this month.
- (3) If page includes “currentPage=” and “countPage=”, this page has been loaded completely and this is information about current page of current month and total page number of this month.

- 3) The link of comments of a message is like:

```
<a href="javascript:void(0);" class="S_txt2" action-type="fl_comment" action-
data="ouid=1444865141&amp;location=profile&amp;comment_type=0" suda-
uatrack="key=profile_feed&amp;value=comment:923874989"><span class="pos"><span class="line
S_line1" node-type="comment_btn_text"><span><em class="W_ficon ficon_repeat
S_ficon"> </em><em>149</em></span></span></span></a>
```

4) There are 3 types of picture links:

```
<li class="WB_pic S_bg2 bigcursor" suda-
uatrack="key=comment_pic_click&value=cmt_thumbnail_click" action-type="comment_media_img"
action-data="pid=0061XAWoly1furz0kqwrj30sp0xrnw&cid=4278804458837161"><i class="W_loading"
style="margin: 74px 74px 74px -90px; display: none;"></i></li>
```

```
<li class="WB_pic li_1 S_bg1 S_line2 bigcursor li_n_mix_w" action-
data="isPrivate=0&relation=0&pid=561ee475gy1fursmpp6pbj20zk0qid&object_ids=10420
18%3A76c124f20aaf2b908244de97cbdc37ce&photo_tag_pids=&uid=1444865141&mid=42
78749829157046&pic_ids=561ee475gy1fursmpp6pbj20zk0qid&pic_objects=" action-
type="feed_list_media_img" suda-
uatrack="key=tblog_newimage_feed&value=image_feed_unfold:4278749829157046:561ee475gy1fur
smpp6pbj20zk0qid:1444865141:0"> <i class="W_loading"
style="display: none;"></i>
```

```
<li class="WB_pic li_1 S_bg1 S_line2 bigcursor " action-
data="isPrivate=0&relation=0&pic_id=8ccdd811gy1fvmsp4wf77j218s0q2k09" action-
type="fl_pics" suda-
uatrack="key=tblog_newimage_feed&value=image_feed_unfold:4288467662528718:8ccdd811gy1fv
msp4wf77j218s0q2k09:2362300433:0"></li>
```

5) Replace following strings as “large” and get the addresses of original size pictures:

```

```

```

```

```

```

5.5.4 Main Logic of the Tool

- 1) Main logic of this tool is in file WeiboSnapRunController.java.
- 2) loadMain(): Find the account name and its valid months range.
- 3) loadPage(): Determine whether the contents of page and the pictures have been loaded.
- 4) snapPage(): Scroll screen to snap page and save images as PDF files.

5.5.5 Implementation

- 1) Embed browser Webview of JavaFx
- 2) Code to interactive between WebView and page:
`webEngine.executeScript("window.scrollTo(0, " + snapHeight + ");");`
- 3) Function to make snapshot:
`Image snapshot = webView.snapshot(snapPara, null);`
- 4) PDF API:
<https://pdfbox.apache.org/>

<End of Document>