

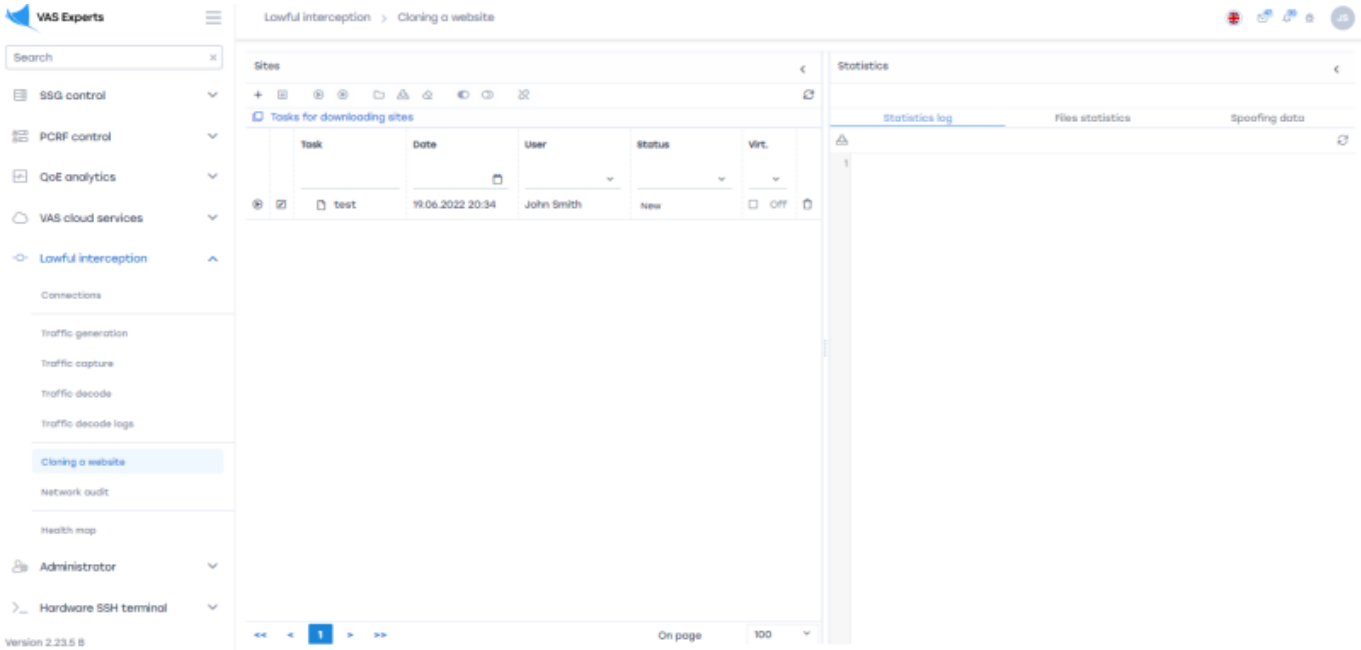
# Содержание

<b>5 Cloning a website</b>	3
<b>Sites</b>	3
Add a subtask	3
Add a crawl	8
Run scan	8
Stop scanning	8
Get acquainted with the properties of data folders	8
Download archive	8
Clear Folder	8
Enable Virtual Host	9
Turn off the virtual host	9
Site Spoofing	9
Update the list	9
Change List Item	9
Delete a list item	9
<b>Statistics</b>	9



# 5 Cloning a website

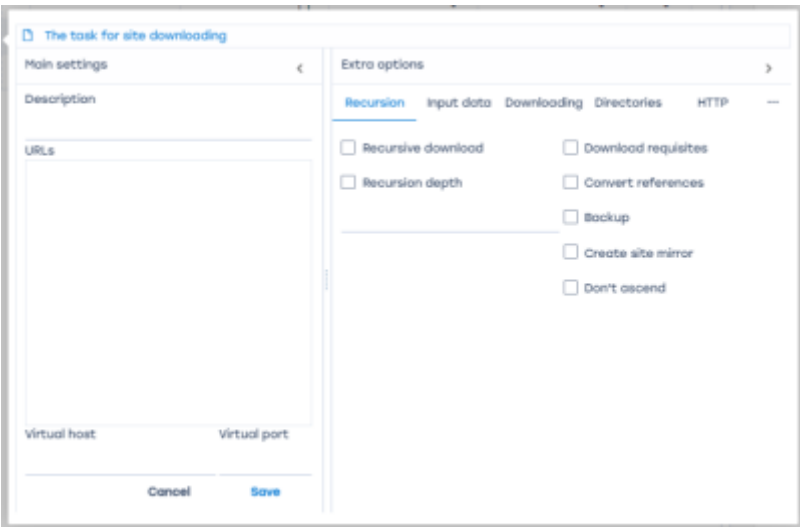
To go to the section, click the **LAWFUL INTERCEPTION** menu item, then click the **CLONING A WEBSITE** menu item.



## Sites

### Add a subtask

To add a subtask to the website download "**Add a subtask**". In the window that opens, enter the parameters.

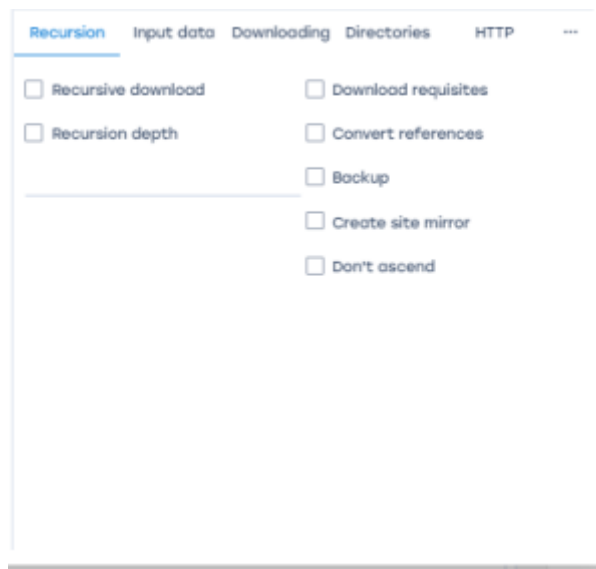


Basic parameters:

1. **Description.** This field contains the name or description of the profile
2. **URL(a).** Url(a) or a list of URLs of files that the user needs to download.
3. **Virtual host address.** The address of the virtual host on which the cloned site will be deployed (URL links).
4. **Virtual host address.** The address of the virtual port on which the cloned site will be deployed (URL links).

**Additional parameters:**

## Recursion



The screenshot shows a web interface with a tabbed menu at the top. The tabs are 'Recursion', 'Input data', 'Downloading', 'Directories', and 'HTTP'. The 'Recursion' tab is selected and highlighted in blue. Below the tabs, there are two columns of checkboxes. The left column contains 'Recursive download' and 'Recursion depth'. The right column contains 'Download requisites', 'Convert references', 'Backup', 'Create site mirror', and 'Don't ascend'. All checkboxes are currently unchecked.

- **Recursive loading.** If the user needs all the links contained on the cloned site, then it is necessary to check the box in this cell.
- **Recursion depth**
- **Upload banking details.** If the user needs the images needed to load the html site, then it is necessary to check the box in this cell.
- **Convert links.** Make links local in uploaded HTML or CSS.
- **Make a backup.** Creating a copy of a file with расширением.оргд . The ORIG format is used to indicate the original version of a file that is saved as a backup after editing or updating.
- **Create a site mirror.** If the user wants to create a duplicate (copy) of the site, then it is necessary to check the box in this cell. The mirror is a duplicate, a copy of the site.
- **Do not climb.** Do not go up to the parent directory.

## Input data

Recursion **Input data** Downloading Directories HTTP ...

☐ Continue downloading

☐ Force HTML

☐ URL prefix

Extra options

- **Continue downloading.** If the operation was interrupted while downloading files, then you can resume downloading from the place of interruption by checking the continue downloading checkbox.
- **Force HTML**
- **The URL prefix** indicating the protocol used to access the location, i.e. the data transfer protocol.
- **Additional parameters.**

## Download

Recursion Input data **Downloading** Directories HTTP ...

☐ Pause ☐ Tries

☐ Random wait ☐ Retry even if refused

☐ Speed (bps) ☐ Time-out

☐ Server response ☐ Don't re-retrieve

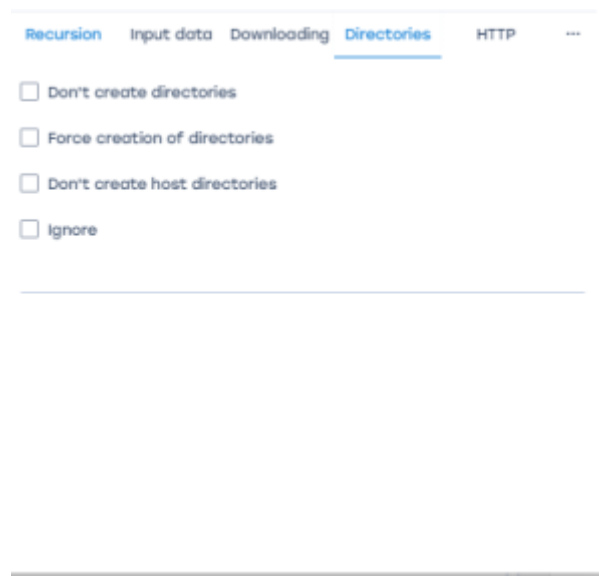
☐ Skip downloads

- **Pause.** The user can set a pause before each file download. The duration of the pause is specified by the user independently in seconds.
- **Attempts.** The user can configure the number of repeated attempts to download the file in case it is interrupted. The number of attempts is entered manually by the user.
- **Arbitrary pause.** The user can choose an arbitrary pause before downloading each file. The pause duration is determined randomly in the range from 0.5 to 1.5 seconds.
- **Speed (bits/sec).** If necessary, the download speed can be limited. As a result, this operation will not occupy the entire available data channel and will not affect other processes associated

with the network. You can do this by specifying a speed limit expressed in bytes (as a regular number) per second.

- **Repeat in case of refusal** . Repeat the download even if the connection is denied.
- **Timeout**. File download standby mode. Entered manually by the user in seconds.
- **Response from the server**. Conclusions of the response from the server.
- **Skip downloads**. If a user wants to copy an entire website, but does not need files of a certain type, you can disable their download using the parameter.
- **Do not re-upload**. Do not re-upload files only if they are not newer than the local ones.

## Directories



The screenshot shows a software interface with a tabbed menu at the top. The tabs are 'Recursion', 'Input data', 'Downloading', 'Directories' (which is selected and underlined), 'HTTP', and '...'. Below the tabs, there are four unchecked checkboxes with the following labels: 'Don't create directories', 'Force creation of directories', 'Don't create host directories', and 'Ignore'. A horizontal line is positioned below the 'Ignore' checkbox.

- **Do not create directories.**
- **Force the creation of directories.**
- **Do not create host directories.**
- **Ignore the number of components of the remote directory.**

## FTP

Recursion   Input data   Downloading   Directories   **FTP**   ...

☐ Don't remove '.listing' files

☐ Upload files by reference

☐ Turn off masks

☐ Disable passive mode

- **Do not delete files.listing**
- **Upload files by links, not directories**
- **Turn off masks**
- **Disable passive mode**

## HTTP

Recursion   Input data   Downloading   Directories   **HTTP**   ...

☐ Http user                      ☐ Http password

---

☐ Proxy user                      ☐ Proxy password

---

☐ POST data                      ☐ Referer header

---

☐ Don't use cookies              ☐ Agent

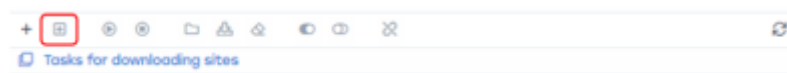
☐ Do not use keep alive

☐ Do not use cache

- **HTTP user.** Login input.
- **HTTP password.** Enter the password.
- **Proxy user.** Login input.
- **Proxy password.** Enter the password.
- **POST-lines.** Sending a string as data.
- **Referer header.** The HTTP referer header is transmitted from the browser to the server to tell which page the reader was on before switching to the current one.
- **Do not use cookies.**
- **Do not use keep alive.**
- **Do not use the cache.** Reject data cached by the server.
- **Agent.** Identify yourself as an agent instead of Wget

## Add a crawl

To add a subtask to the website download, click on the button **"Add subtask"**, which is located in the toolbar and do the tasks described above.



## Run scan

To start scanning a profile, select a profile from the list and click on the **"Start scanning"** button.



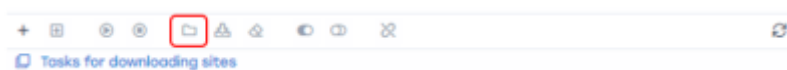
## Stop scanning

To stop the running profile scan, click on the **"Stop scanning"** button.



## Get acquainted with the properties of data folders

To view the properties of the data folder, click on the "Data Folder Properties" button, after which a window with the folder data will be displayed.



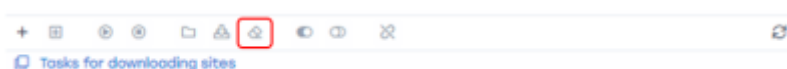
## Download archive

To download the zip archive, click on the "Download" button.



## Clear Folder

To clear the folder with subtasks, click on the "Clear Folder" button.





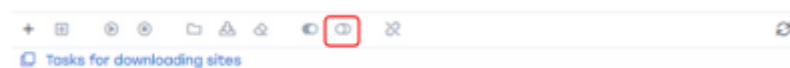
## Enable Virtual Host

To enable the virtual host of the selected task or subtask, click on the "Enable virtual host" button.



## Turn off the virtual host

To turn off the virtual host of the selected task or subtask, click on the "Turn off virtual host" button.

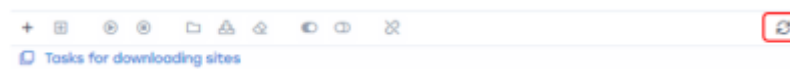


## Site Spoofing



## Update the list

To update the list of profiles, click on the "**Update**" button.



## Change List Item

To change the profile or subprofile parameters, click on the "**Edit**", which is located to the left of each list item.



## Delete a list item

To delete an item from the list, click on the "**Delete**", which is located to the right of each list item.



## Statistics

This subsection contains information about logical statistics, file statistics - information about the

download of each file, spoofing — information about the final download.