

Installation guide

Web presentation checker

Jindřich Máca

15. prosince 2014

Obsah

1	Introduction	2
2	Sign up and Log in	2
2.1	Sign up	2
2.2	Log in	3
2.3	Log out	4
3	Checkup control	4
3.1	Create new checkup	4
3.2	Checkup options	5
3.2.1	Starting point (url)	5
3.2.2	Desired tests	5
3.2.3	Allowed domains (each on new line)	6
3.2.4	Crawling depth limit	6
3.2.5	Min request interval [ms]	6
3.2.6	Page limit	7
3.2.7	Additional HTTP headers (each on new line)	7
3.3	Stop the checkup	7
3.4	Checkup list	7
3.5	Repeat the checkup	8
4	Checkup results	8
4.1	Message results	9
4.2	Graph results	10
4.2.1	Text graph result	11
4.2.2	Graphical graph result	11

1 Introduction

This document is provided to serve as user guide for Web presentation checker tool. This tool was designed to run checkups on whole websites a check their links, validity and CSS redundancy and also to give users complete textual and even graphical report about their structures. This is a simple guide, which shows to users, what is possible to do with this tool and how should they use it. It is understood, that you have already installed the tool or you use it as an external service via its web interface. If you have not installed it yet, you are welcome to use our provided Installation guide. Either way, you should start this guide with prepared web address, where the tool is running and its ready for use.

2 Sign up and Log in

2.1 Sign up

First of all, when you use Web presentation checker for the first time, you have to sign up to it. This step is very important, because **you can't run this tool without signing up!** But don't be afraid, because signing up is there actually for you advantage and with it, you can run your checkups asynchronously. So, when you're ready and you have Welcome page of Web presentation checker tool before you, click on the Sign up button in the top menu, which is highlighted in the picture bellow. Now, you should see a registration form to which please fill your user account information such as your email, password, name and surname and click on Sign up button under the registration form. When everything went right, you should be able to Log in now. If you are not, please contact service administrator or once more look closely to Installation guide and try solve the problem by proper setting of the tool.

WWW presentations checker

You're not logged in Log in Sign up

Email:

Password:

Verify your password:

Name:

Surname:

Sign up

Obrázek 1: Sign up

2.2 Log in

If you have already Sign up, you should be able to log in. To do so, click on the Log in button in the top menu, as you can see in the picture bellow. Now fill your user account information, which you filled on signing up, into the log in form and finally click on the Log in button under the log in form. Then you should find yourself again on the Welcome page, but now there should be a welcome message in the left part of the top menu. If there is, you are now able to create your checkup.

WWW presentations checker

You're not logged in Log in Sign up

Email:

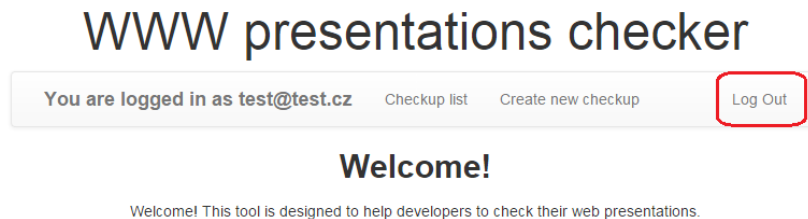
Password:

Log In

Obrázek 2: Log in

2.3 Log out

When you are done with the tool for the day, you should **always log out**. You can do it by Log out button situated in right part of the top menu. Just click on it and you should see a message, that you was successfully log out. You can do this anytime you want.



Obrázek 3: Log out

3 Checkup control

3.1 Create new checkup

For creating a new checkup you must first Log in. Then you click on Create new checkup button in the top menu. Now you should see a checkup form same as in the picture bellow. Now you fill you desired checkup options, click on the Start validation button under the checkup form and if everything is set properly, you should see the message, that your checkup has been created successfully. It usually take some time to finish the checkup, but, as was said in Sign up, it is running asynchronously. That means, that you can easily Log out, close entire page and return for the Checkup results later.

WWW presentations checker

You are logged in as test@test.cz
Checkpoint list
Create new checkpoint
Log Out

Starting point (url):

Desired tests: ☐ HTML validation ☐ CSS redundancy checker ☐ check links ☐ CSS validation

Allowed domains (each on new line):

Crawling depth limit:

Min request interval (ms):

Page limit:

Additional HTTP headers (each on new line):

Obrázek 4: Create new checkpoint

3.2 Checkup options

There are many option and several very important rules for them, so lets go through it in order as it is in picture above. Basic rule is, that if your checkup is not working, in the most cases its caused by wrong checkup options, so please check them carefully.

3.2.1 Starting point (url)

Starting point is URL address of the page, you want to run your checkup on. This is the starting point of the whole checkup, so put there the most basic URL address you want to start the checkup from. It also must be always entered in its full formatting. That means mainly you to **not forget about http:// or https:// protocols** as well, because they must always be there. It is also possible to put a port in there, if your selected website runs on different port then standard 80. You can read more about URL address formatting on http://en.wikipedia.org/wiki/Uniform_resource_locator.

3.2.2 Desired tests

This is the main part, where you check tests, that you are interested in and you want them to be proceed by the checkup. There follows the description of single tests.

- **HTML validation** - If this test is checked, every HTML page, that goes through the checkup is send for HTML validation to W3C validation service. You can read more about this service here <http://validator.w3.org/about.html>

- **CSS redundancy checker** - If this test is checked, the tool controls if CSS styles with their attributes are redundant in selected web presentation.
- **check links** - If this test is checked, the tool checks validity of all URL links, that appear in selected web presentation.
- **CSS validation** - If this test is checked, every CSS file, that goes through the checkup is send for CSS validation to W3C validation service. You can read more about this service here <http://jigsaw.w3.org/css-validator/manual.html>

3.2.3 Allowed domains (each on new line)

It is possible, that one web page runs on multiple domains or you simply want to run the checkup on multiple websites connected together. This is done by Allowed domains, which, logically by its name, specifies list of domains, on which is this checkup allowed. You can simply put there list of allowed domains on which will the Web presentation checker tool continues running the checkup. But there is one basic rule, in which the most users made mistakes, so read carefully. These domains unlike Starting point (url) are **put there only by their names**. So no protocols and no ports, just a simple domain name as you can see in the picture bellow. You don't need to put there domain from Starting point (url), this is done automatically.

3.2.4 Crawling depth limit

This number represents the depth, to which the checkup goes. That is decided by length of its URL address path. It is accepted in interval from 1 to 20. But the last layer of the checkup is also validated for example for invalid links, but the pages bellow are no longer downloaded and proceed for example by HTML or CSS validation, even there are HTML page or CSS files.

3.2.5 Min request interval [ms]

This number represents number of milliseconds, that the checkup waits before it asks the one concrete domain for the next page. Allowed interval is from 500 to 10 000 ms. It is place there, because if you will 'bomb' certain domain with too many request in short time, it might place you on some sort of black list and block your access to it completely. Of course the bigger the number is, longer will the checkup take, but it is recommended **not** to place there lower value than the default one.

3.2.6 Page limit

This is the number, that represents number of pages, which the checkup runs fully through. Allowed interval here is from 1 to 100 000 pages. The bigger the number is, longer will the checkup take. There is no option to set this number to infinite, because that might cause, that the checkup would never ended, but there can be set a pretty big number.

3.2.7 Additional HTTP headers (each on new line)

There can more experience users place their alternative HTTP headers, that are use for each request sent during the checkup. Read more about HTTP header here http://en.wikipedia.org/wiki/List_of_HTTP_header_fields.

3.3 Stop the checkup

You can of course stop any running checkup. For that, go to Checkup results and there, if the checkup is still running, what you can see from its State 'Running', you will see the red Stop button on the right side of it as you can also see in the picture bellow. Just click it and checkup will automatically stop and you can Checkup results, but remember it will be just partial.

WWW presentations checker

You are logged in as test@test.cz Checkup list Create new checkup Log Out

#	Start URL	State	Date created	Desired tests	Max depth	Page limit	Domains allowed	Actions
29	https://www.seznam.cz/	Running	17.11.2014 06:01	CSS redundancy checker, CSS validator, HTML validator, check links	5	1	www.seznam.cz	Stop
28	https://www.seznam.cz/	Finished	17.11.2014 05:25	CSS redundancy checker, CSS validator, HTML validator, check links	5	1	www.seznam.cz	Show results Show graph Retry
27	https://www.seznam.cz/	Finished	16.11.2014 23:09	CSS redundancy checker, CSS validator, HTML validator, check links	5	1	www.seznam.cz	Show results Show graph Retry

Obrázek 5: Stop the checkup

3.4 Checkup list

You also have a complete list of all your checkups. If you wish to see them, click on Checkup list button in the top menu as you can see in the picture bellow. This site is divided to pages, that you can navigate through by page navigation on the top and on the bottom as you can also see on the picture bellow. The checkup list has its specific format. First column is unique ID of the control, then all well known Checkup options with extra column State,

which show the status of the checkup and in the last column are Actions, which you can do with this checkup.

WWW presentations checker

You are logged in as test@test.cz
Checkup list
Create new checkup
Log Out

1 2 3

#	Start URL	State	Date created	Desired tests	Max depth	Page limit	Domains allowed	Actions
28	https://www.seznam.cz/	Finished	17.11.2014 05:25	CSS redundancy checker, CSS validator, HTML validator, check links	5	1	www.seznam.cz	Show results Show graph Retry
27	https://www.seznam.cz/	Finished	16.11.2014 23:09	CSS redundancy checker, CSS validator, HTML validator, check links	5	1	www.seznam.cz	Show results Show graph Retry
26	http://localhosttest/	Finished	15.11.2014 19:19	CSS redundancy checker, CSS validator, HTML validator, check links	1	10	localhost	Show results Show graph Retry

Obrázek 6: Checkup list

3.5 Repeat the checkup

Also if you want to repeat the same checkup, that you already did, you can just go to Checkup results, where you click on the Retry button on the right side of checkup you want to repeat. You will then see typical Create checkup form from Create new checkup, but with prefilled values from that previous selected checkup. You can also change some or all of them, if you want to. And then there is also an alternative way. When you to , then the first button with label Retry this checkup in the middle of the page on the top has the same effect like the Retry button from .

4 Checkup results

If you want to see the checkup result for specific checkup, that checkup have to be already finished. To check that go to and look at the column State from checkup, that you want to see results from and if there is state 'Finished', you are good to go, else you must wait a little bit longer for the checkup to finish its job. You will also recognize this, by buttons on the right side, because if there is only Stop button, the checkup still running, otherwise, you are also good to go. Then you have a choice from two types of results Message results, which is list of all important messages from all types of tests you checked in options or Graph results, where you can see graph of crawling through the website in its text or graphical form. If you want do display Message results, click on the Show results button on the right side of the checkup, you want to see the results from. For the Graph results click on Show graph button right next to it.

WWW presentations checker

You are logged in as test@test.cz

[Checkup list](#)
[Create new checkup](#)
[Log Out](#)

« 1 2 3 »

#	Start URL	State	Date created	Desired tests	Max depth	Page limit	Domains allowed	Actions
28	https://www.seznam.cz/	Finished	17.11.2014 05:25	CSS redundancy checker, CSS validator, HTML validator, check links	5	1	www.seznam.cz	Show results Show graph History
27	https://www.seznam.cz/	Finished	16.11.2014 23:09	CSS redundancy checker, CSS validator, HTML validator, check links	5	1	www.seznam.cz	Show results Show graph History
26	http://localhost/test/	Finished	15.11.2014 19:19	CSS redundancy checker, CSS validator, HTML validator, check links	1	10	localhost	Show results Show graph History

Obrázek 7: See the checkup results

4.1 Message results

Message results represents the list of all important messages from all types of checkup, you checked in options, for this specific checkup. For example you can find there messages for invalid links, invalid HTML pages or CSS files etc. Its format is specific and you can see it in the picture bellow. First is URL address to which is this message related, then Type of the message, then content of the Message and in the end Location of the message in the file, which it comes from, if it can be specified, otherwise there is 'N/A' expression. These messages can be simply filtered by form on the top of the page. You can choose there results from individual tests and you can also filter them by theirs type, which also determines their priority. Just check the results you want to display and click on Show button under it. For message types specifications see the list bellow.

- **Info** - Message with lowest weight, its purpose is to inform you about something.
- **Warning** - Message with a little bigger weight, this is harmless, but it could have been done better.
- **Error** - Serious error, which should be fixed.
- **Debug** - Not an error. This is used for debugging the application, should not appear in production.

WWW presentations checker

You are logged in as test@test.cz
[Checkup list](#)
[Create new checkup](#)
[Log Out](#)

[Retry this checkup](#)

Show only results by:
☐ HTML Validator
☐ Web crawler
☐ CSS redundancy checker
☐ CSS Validator

Filter results by priority:
☐ Warning
☐ Error
☐ Info
☐ Debug

[Show](#)

URL	Type	Message	Location
https://www.seznam.cz/st/3.34.18/css/0-homepage.css?3.34.18	Error	Value Error : margin (nullbox.html#propdef-margin) Unknown dimension	Line 1, column 0
https://www.seznam.cz/st/3.34.18/css/0-homepage.css?3.34.18	Error	Property zoom doesn't exist :	Line 1, column 0
https://www.seznam.cz/st/3.34.18/css/0-homepage.css?3.34.18	Error	Unknown pseudo-element or pseudo-class .none	Line 1, column 0
https://www.seznam.cz/st/3.34.18/css/0-homepage.css?3.34.18	Error	Unknown pseudo-element or pseudo-class .left	Line 1, column 0
https://www.seznam.cz/st/img/2011/overlyBg.jpg	InvalidLinkMsg	Invalid link	N/A
https://www.seznam.cz/st/3.34.18/css/0-homepage.css?3.34.18	Error	Value Error : cursor (nullui.html#propdef-cursor) hand is not a cursor value :	Line 1, column 0

Obrázek 8: Message results

4.2 Graph results

There are two main types of Graph results. Both representing crawling through the selected website in this specific checkup, but one Text graph result is in text represented form and the other Graphical graph result is in graphical represented form. If you follow Checkup results, you should now find yourself on page for displaying Graph results as you can see in the picture bellow. You can now select the type of graph you want to display from the select box in the middle of the page on top and then render it by clicking on Render button or downloaded it to you computer by clicking on the Download button bellow.

WWW presentations checker

You are logged in as test@test.cz
[Checkup list](#)
[Create new checkup](#)
[Log Out](#)

Select graph to be rendered
Traversal graph tag tree

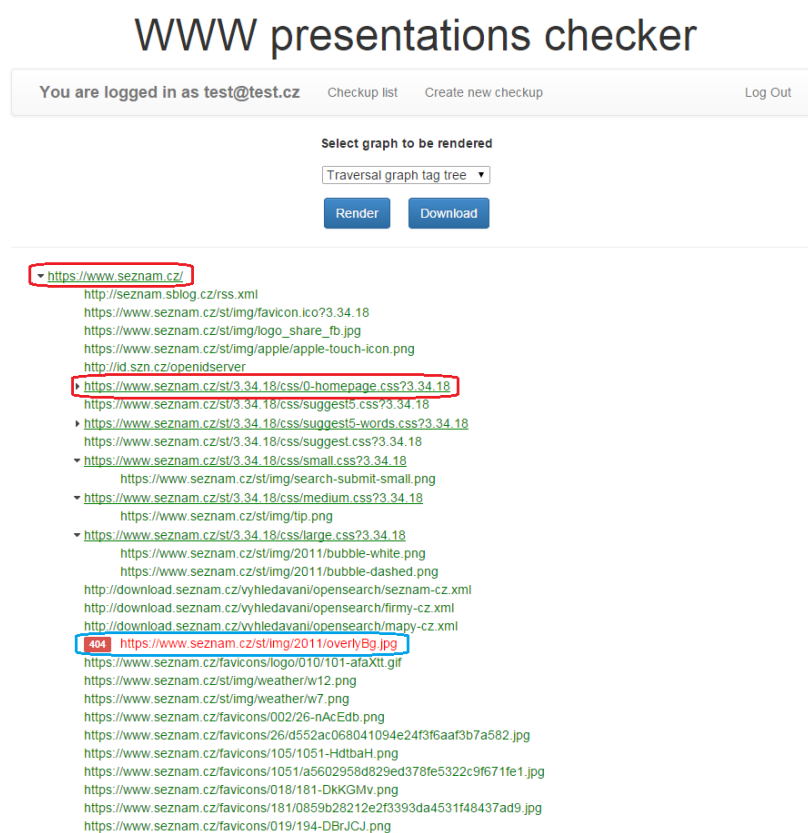
[Render](#)
[Download](#)

[Back to Checkup list](#)

Obrázek 9: Graph results

4.2.1 Text graph result

This graph represents crawling tree with starting url as root. Text graph result is the structured list of URL addresses crawled in this specific checkup. In the beginning it is collapsed, but by clicking on underlined URL addresses with little arrow on their left side, you can easily click yourself through it and see how checked web presentation is organized and also where are invalid links, because they are here highlighted by red color and have their HTTP error code by left side. All this you can again see in the picture below.

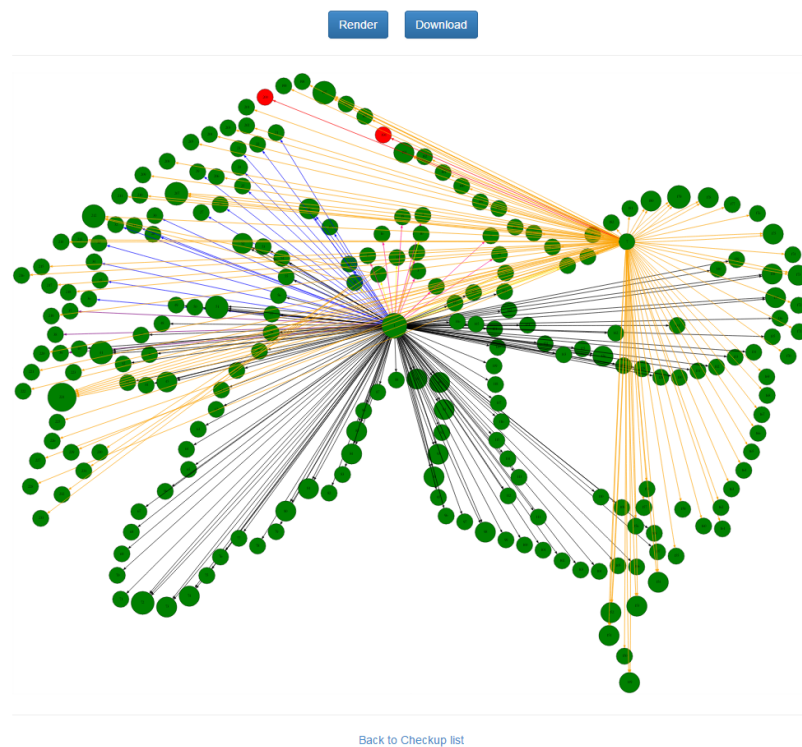


Obrázek 10: Text graph result

4.2.2 Graphical graph result

This is the graphical representation of URL addresses crawled in this specific checkup and its presented by oriented graph. More about these graph you can read here http://en.wikipedia.org/wiki/Directed_graph a example of the one is also here below.

Full graph All the nodes represent URL addresses. Nodes are clickable, and by their colors and size you can recognize, which of them are valid (green) or invalid (red) and how important (size) the node is. The edges then represents the specific way of the connection between two URL addresses displayed by nodes, tooltip shows you informations about the link between two nodes and color represents the source type.



Obrázek 11: Graphical graph result

Reduced graph This graph is same as Full graph, but it has fewer edges, because multiedges are transformed into one edge with number, which says, how many edges the reduced edge represents. All edges have the same color (black), because they possibly can have more source types.