

## TYPO3 Core - Feature #85127

Epic # 85006 (New): Reduce falsely reported broken links

### linkvalidator: Add possibility to exclude specific external URLs / domains or patterns

2018-05-31 12:09 - Sybille Peters

<b>Status:</b> New	<b>Start date:</b> 2018-05-31
<b>Priority:</b> Should have	<b>Due date:</b>
<b>Assignee:</b> Sybille Peters	<b>% Done:</b> 0%
<b>Category:</b> Linkvalidator	<b>Estimated time:</b> 0.00 hour
<b>Target version:</b>	
<b>PHP Version:</b>	<b>Complexity:</b>
<b>Tags:</b>	<b>Sprint Focus:</b>
<b>Description</b>	
Sometimes there may be problems with external sites that linkvalidator / Guzzle doesn't handle correctly or which can't be handled correctly at all.	
To make linkvalidator still usable, it would be good to be able to exclude certain domains / urls using regular expressions. This should be supplied in a way that an admin user can edit it.	
Prerequisite: Provide general configuration for extension. This does not need to be made configurable per page tree.	
<b>Examples</b>	
URLs that work in browser, but check will fail:	
<ul style="list-style-type: none"><li><a href="https://www.linkedin.com/school/kieluni/?trk=similar-pages_result-card_full-click">https://www.linkedin.com/school/kieluni/?trk=similar-pages_result-card_full-click</a></li><li><a href="https://www.drive2day.de/">https://www.drive2day.de/</a></li></ul>	
<b>Related issues:</b>	
Related to TYPO3 Core - Bug #86918: Linkvalidator stops working on specific l...	<b>Closed</b> <b>2018-11-13</b>
Related to TYPO3 Core - Feature #89457: Add possibility to mark as error spe...	<b>Needs Feedback</b> <b>2019-10-18</b>

## History

### #1 - 2019-09-26 11:04 - Sybille Peters

- Subject changed from Add configuration option: file with regex patterns to exclude for external link checking to Add possibility to exclude specific external URLs / domains or patterns

## Problem

In some rare cases, the checking of external URLs fails via linkvalidator even if they are not broken. See also [#86918](#). This leads to false negatives (URLs reported as broken which are not).

One of the most annoying things for editors when working through the list of broken links are false negatives which keep coming up, clutter up the list, cannot be removed and make it really tedious to actually work through the list and fix broken links.

## Proposed Solution

It would be ideal to optimize the crawl process so it always correctly reports broken links but this may not be entirely possible.

So as alternative we could add a mechanism to exclude specific URLs or URL patterns, e.g.

- exact URL
- URL starting with ... (or domain)
- regular expression

This could be done by just adding optional files or could be done in the GUI with an "ignore" button.

### #2 - 2019-09-26 11:05 - Sybille Peters

- Related to Bug #86918: Linkvalidator stops working on specific links (external URLs) added

**#3 - 2019-10-18 20:36 - Lina Ourima**

- Related to Feature #89457: Add possibility to mark as error specific external URLs / domains or patterns added

**#4 - 2020-03-24 21:02 - Sybille Peters**

- Subject changed from Add possibility to exclude specific external URLs / domains or patterns to linkvalidator: Add possibility to exclude specific external URLs / domains or patterns

**#5 - 2020-09-07 17:02 - Sybille Peters**

- Description updated