

Primer #04 on Using Regular Expressions

Test Programs and Test Data – page 1

For things here to make sense, I'm assuming you're already familiar with the RegEx concepts and conventions presented in Primer #01 and Primer #02 and Primer #03.

Release 4.17

NOTE: The goal with Primers 01 and 02 is to help the user develop a minimal Knowledge of Regular Expressions. Primer 03 should enable users populate the spreadsheet application with relatively simple RegEx that will be generate compound and complex RegEx that will really help us keep up with the spammers. Primer 04 provides information about selected RegEx Test Programs and shows some comparisons.

Below is an example of starting with a simple RegEx and then over time, making it more complex. This simplicity to complexity example comes as a result of experiencing how spammers transform the data they are putting email addresses, subject lines and body text:

Example:

- **1st iteration: (2 separate items)**
 - wife wont stop flirting
 - girlfriend won't stop flirting
- **2nd iteration: (combining the two and allowing for both wont and won't)**
 - (girlfriend|wife) won'?t stop flirting
- **3rd iteration: (adding other females & can't or cant)**
 - (girl friend|wife|woman|sister) (won'?t|can'?t) stop flirting
- **4th iteration: (adding cannot)**
 - (girl friend|wife|woman|sister) (won'?t|can'?t|can not) stop flirting
- **5th iteration: (adding other verbs)**
 - (girl friend|wife|woman|sister|daughter) (won'?t|can'?t|(can|will) not|refuses? to) (stop|cease|end) flirting
- **6th iteration: (adding males and more verbs)**
 - (girl friend|wife|woman|sister|daughter|mother|spouse|husband|brother|boy friend|man|father|son) (won'?t|can'?t|(can|will) not|refuse(s?|d) to) (stop|cease|end|quit|resist) flirting

As you develop your RegEx, you should test them. Each of the TABS in the spreadsheet application has a table in which you can enter test data. The test data is transformed into a lower case, ordered list. The application allows the user (with the click of a button) to

1. Copy the RegEx into the clipboard where it can be pasted into the test program's RegEx window
2. Copy test data into the clipboard where it can be pasted into the test program's test data window

Primer #04 on Using Regular Expressions

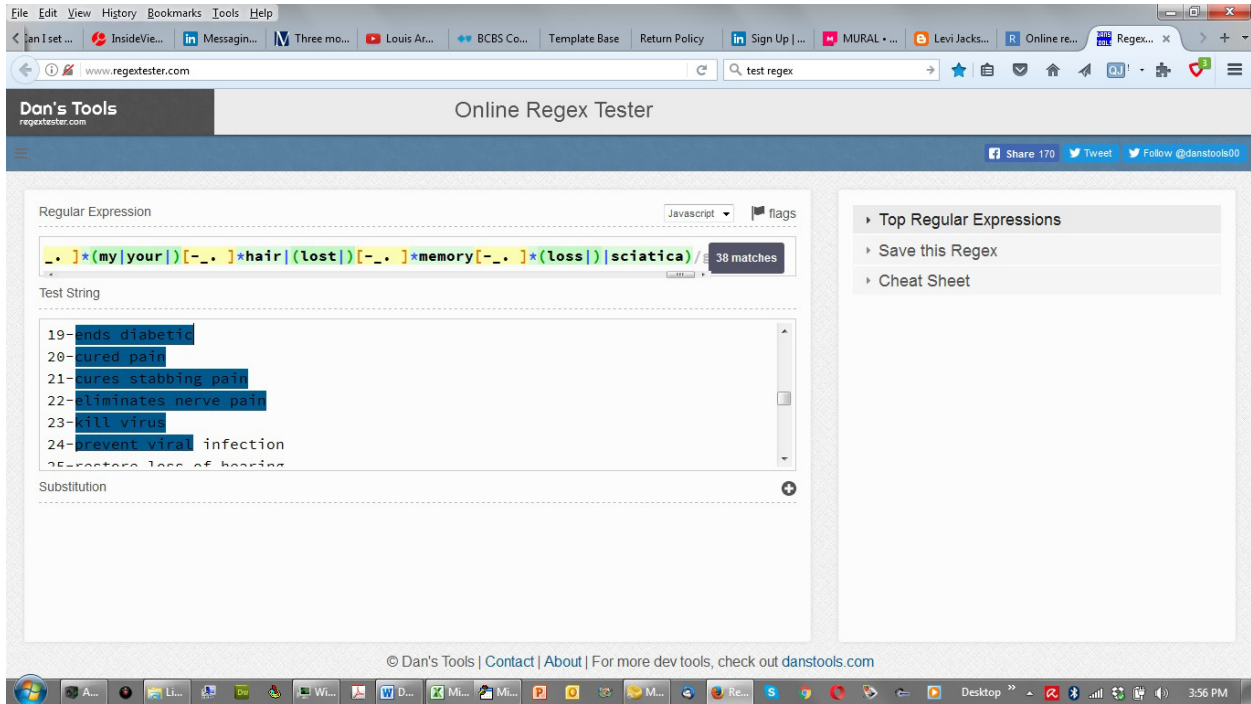
Test Programs and Test Data – page 2

1. **Recommendation:** Test your regular expressions as you develop them...
 - There are many resources for testing Regular Expressions (RegEx). I have highlighted three that I have used.
 - Rather than recommending a specific program, I will share a comparison of the performance of these three programs and let you decide which to use.
 - Each of these three programs I have used and tested have their strengths.
 - On occasion I do all three in order to check whether the RegEx is really working or in order to debug a challenging RegEx or when trying something new.
 - I've created a table in the application's control TAB that can accommodate up to five RegEx teste programs. The application allows users to select the preferred test program. The TestRegEx button on each TAB launches what has been selected as the preferred test program.
 - <http://www.regextester.com> – online testing at this website
 - <https://www.regex101.com> – online testing at this website
 - <http://regexpstudio.com/TRegExpr/TRegExpr.html> - downloadable .exe file
 - This .exe may no longer be on this page but the pascal code for a later version is available
 - I have found that their performance across all three programs is about 80% similar
 - There are some execution and results differences highlighted in the pages that follow.
 - **Here is a link to a very informative and useful website for almost anything you'd want to know about Regular Expressions**
 - <http://www.regular-expressions.info>

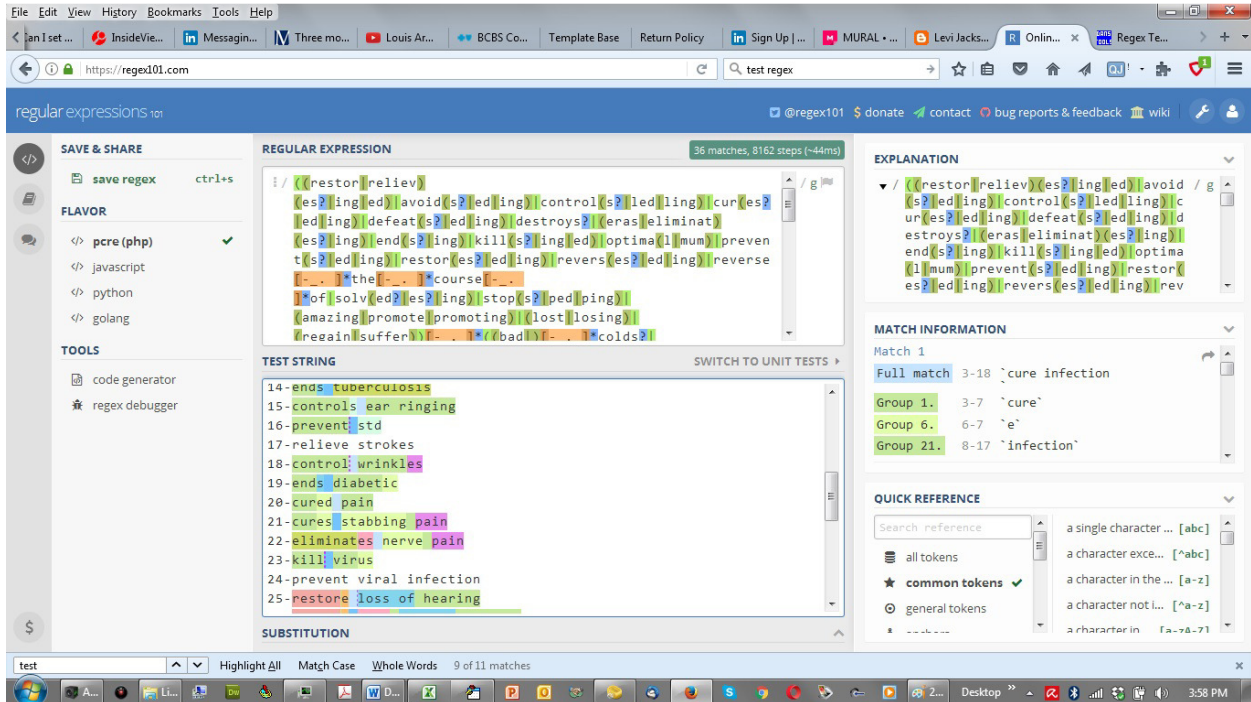
Primer #04 on Using Regular Expressions

Test Programs and Test Data – page 3

Differences between www.regextester.com and www.regex101.com when testing the RegEx against the same set of test data.



Regex101.com (below) has 36 matches whereas regextester.com (above) has 38 matches. not sure why



Primer #04 on Using Regular Expressions
Test Programs and Test Data – page 4

Primer #04 on Using Regular Expressions
Test Programs and Test Data – page 5

Test Data Used

01-cure infection
02-kill bad colds
03-prevents dementia
04-stop kidney disease
05-cure kidney disease
06-eliminate cellulite
07-avoid toenail fungus
08-solve insomnia
09-relieve infection
10-prevent herpes
11-end blood disease
12-restored hearing loss
13-ends tb
14-ends tuberculosis
15-controls ear ringing
16-prevent std
17-relieve strokes
18-control wrinkles
19-ends diabetic
20-cured pain
21-cures stabbing pain
22-eliminates nerve pain
23-kill virus
24-prevent viral infection
25-restore loss of hearing
26-restore your loss of hearing
27-restore your hearing loss
28-restore your hearing
29-restore lost hearing
30-prevent loss of hearing
31-prevent your loss of hearing
32-prevent your hearing loss
33-prevent hearing loss
34-restore hearing loss
35-restores hearing
36-restoring hearing
37-stopping knee pain
38-eliminating knee pain

regex101.com - not found

regex101.com - not found