Aggregate Anonymized Survey

Start of Block: Informed Consent

Q123 We provided information regarding consent to the reader and ask if they are a professional software developer that has used regex. This has been removed to maintain anonymity.

1.2
CAPTCHA

End of Block: Informed Consent

Start of Block: Regex Re-use

2.1 Page **1** of **4**

This section is about your experience **re-using** (i.e., copy-paste or copy-paste and modify) regular expressions.

2.2

Think about your typical practices when putting regular expressions into your code.
What percentage of the time would you estimate that these regular expressions are **re-used from another source (i.e., copy-paste or copy-paste and modify)** such as other source code or Stack Overflow, instead of written yourself "from scratch"?

- ○ 0% (I never re-use; I always write from scratch)
- ○ About 25% (I sometimes re-use)
- ○ About 50% (I re-use about half the time)
- ○ About 75% (I re-use most of the time)
- ○ 100% (I always re-use; I never write from scratch)

2.3

Why have you in the past **decided to re-use regular expressions** as opposed to writing them yourself from scratch? Please select all that apply

- [ ] ... because I needed a regular expression for a very common purpose
- [ ] ... because I knew of a trusted source where I could probably find a regular expression to re-use
- [ ] ...because I believed that a re-used regular expression would be of higher quality than what I would write
- [ ] ... because I believed that a re-used regular expression would be better tested than my own testing
- [ ] ... because I trusted my abilities to validate the re-used regular expression
- [ ] ... because I had many inputs that I could use to validate the re-used regular expression
- [ ] Other factor #1 _____
- [ ] Other factor #2 _____
- [ ] Other factor #3 _____
- [ ] Other factor #4 _____
- [ ] Other factor #5 _____

2.4
In the past why have you selected a **particular** regular expression as the right one to reuse?

- [ ] ... because it was easy to understand.
- [ ] ... because it was well documented; I understood clearly how the regular expression worked from its documentation.
- [ ] ...because it was short.
- [ ] ... because it was designed for my use case, e.g. "a regex for email".
- [ ] ... because it did not need to be modified at all.
- [ ] ...because modifying it for my use case seemed easy.
- [ ] ...because it was from a source that I trust.
- [ ] Other factor #1 _____
- [ ] Other factor #2 _____
- [ ] Other factor #3 _____
- [ ] Other factor #4 _____
- [ ] Other factor #5 _____

2.5 Can you tell us more about your thought process when you **decide to re-use a regular expression** vs **writing it from scratch**?

_____

_____

_____

_____

_____

2.6 Where do you **re-use** regular expressions from? Please **rank** based on how frequently you use the source, 1 being the most frequent source.

	Most Frequent	2	3	4	5	6 (Least Frequent)	I do not use this source
Stack Overflow	○	○	○	○	○	○	○
RegExLib	○	○	○	○	○	○	○
Other code I have written	○	○	○	○	○	○	○
Other code someone at my company has written	○	○	○	○	○	○	○
Other code, e.g. open-source code	○	○	○	○	○	○	○
Other(s)	○	○	○	○	○	○	○

2.7 Is there anything else you would like to say about how you **re-use** regular expressions?

_____

_____

_____

_____

_____

End of Block: Regex Re-use

Start of Block: Regex Concerns

3.1 Page **2** of **4**

This section is about **concerns and validation** strategies associated with using regular expressions.

3.2
When you re-use (copy-paste and possibly modify) a regular expression, estimate the percentage of the time you **know** the regular expression was written **in the same language** in which you will use it.

- ○ 0% (I never know the source language of the regular expression, or they are never the same language)
- ○ About 25% (I sometimes know that the source language and destination language are the same)
- ○ About 50% (Around half of the time I know that the source language and destination language are the same)
- ○ About 75% (Most of the time I know that the source language and destination language are the same)
- ○ 100% (I always know that the source language and destination language are the same)

3.3 Which of the following have you ***actually*** **worried about or encountered** in the past when copy-pasting a regular expression? (check all that apply)

	Worries	Problems
	I have worried about this	I have actually encountered this problem

Syntactic Differences For example, have you worried that something you copied from stack overflow would not "**compile**" (unsupported syntax/features) in your code?	☐	☐
Semantic Differences For example, have you worried that a regular expression you copied from Stack Overflow would **behave differently** in your code then it was described on Stack Overflow?	☐	☐
Performance Impact Have you ever worried that a regular expression you copy pasted would slow down your code?	☐	☐
Other(s)	☐	☐

3.4 For each one of the concerns that you selected above, how did you address them?

_____

_____

_____

_____

_____

3.5 How do you **validate** a regular expression, i.e. how do you make sure that it does what you want it to do? (select all that apply)

- [ ] I write unit tests
- [ ] I test with sample input by hand
- [ ] I ask a colleague
- [ ] I use other tools (please list) _____
- [ ] Other: _____

3.6 Does your validation strategy change when **re-using** a regular expression instead of writing it from scratch?

- ○ Yes, please explain how: _____
- ○ No

3.7 In your experience, what makes regular expression validation difficult (e.g., lack of tools to assist in validation)?

_____

_____

_____

_____

_____

3.8 Is there anything else that you would like to say about your **concerns and validation** strategies associated with using regular expressions?

_____

_____

_____

_____

_____

End of Block: Regex Concerns

Start of Block: Final questions

4.1
Page **3** of **4**

Some quick questions...

4.2
Are you familiar with the concept of Regular Expression Denial of Service (REDoS)?
(REDOS is also known as "catastrophic backtracking" or "super-linear regular expression behavior")

- ○ Yes
- ○ No

4.3 Have you ever modified a regular expression to improve its runtime performance?

- ○ Yes
- ○ No

4.4 Have you ever deployed an application that was bottlenecked by regular expression evaluations? (e.g. scientific computing, big data analysis, etc.)

- ○ Yes
- ○ No

4.5 Is there a person on your team who is the "regular expression wizard" -- the go-to person for all regular expression questions?

- ○ Yes
- ○ No

4.6 Are you the "regular expression wizard" on your team?

- ○ Yes
- ○ No

4.7 Do you feel as if code review impacted how you use regular expressions?

- ○ Yes, please explain how: _____
- ○ No
- ○ My team does not do code review
- ○ I don't work on a team

4.8
When you design a regular expression, is your design influenced by the programming language in which it will be executed?

○ No -- I think that regular expressions are handled the same way in different programming languages.

○ No -- I simply design them using the syntax that I know, even if it is possible that different programming languages may execute them differently

○ No, other: _____

○ Yes, I normally design regular expressions by consulting the specific syntax of the programming language in which they will be executed

○ Yes, other: _____

4.9 Is there anything else that you would like to add about any of the topics that we have discussed?

_____

_____

_____

_____

_____

End of Block: Final questions

Start of Block: Background

5.1 Page **4** of **4**

This section covers **background** with regular expressions and some wrap up logistics.

Display This Question:

If Which of the following have you actually worried about or encountered in the past when copy-pasti... : Worries != Performance Impact Have you ever worried that a regular expression you copy pasted would slow down your code? [I have worried about this]

And Have you ever modified a regular expression to improve its runtime performance? = Yes

5.2 You previously noted that you do not worry about **performance impact** when copy pasting regular expressions. You also stated that you have modified a regular expression to improve its runtime **performance**.

Can you please expand on why you do not worry about **performance** when copy pasting even if you have needed to make **performance improvements** to regular expressions in the past.

_____

_____

_____

_____

_____

Display This Question:

If Which of the following have you actually worried about or encountered in the past when copy-pasti... : Worries = Semantic DifferencesFor example, have you worried that a regular expression you copied from Stack Overflow would behave differently in your code then it was described on Stack Overflow? [I have worried about this]

And If

When you design a regular expression, is your design influenced by the programming language in wh... = No -- I think that regular expressions are handled the same way in different programming languages.

Or When you design a regular expression, is your design influenced by the programming language in wh... = No -- I simply design them using the syntax that I know, even if it is possible that different programming languages may execute them differently

Or When you design a regular expression, is your design influenced by the programming language in wh... = No, other:

5.3 You previously noted that you do worry about **semantic differences** when copy pasting regular expression. You also stated that your design of regular expressions **is not** influenced by the programming language in which it will be executed.

Can you please expand on why you worry about **s**emantic differences when copy pasting even

if you think that regular expressions are handled the same way in different programming languages.

_____

_____

_____

_____

_____

5.4 How long have you worked as a professional software developer?

- ○ Less than one year
- ○ 1-2 years
- ○ 3-5 years
- ○ 6-10 years
- ○ more than 10 years

5.5
Indicate your experience with programming languages, and with regular expressions in those languages. Leave blank any languages in which you have no experience.

	How long have you programmed in this language?						How long have you used regular expressions in this language?					
	Less than one year	1-2 years	3-5 years	6-10 years	More than 10 years	None	Less than one year	1-2 years	3-5 years	6-10 years	More than 10 years	None

C/C++	○	○	○	○	○	○	○	○	○	○	○	○
C#	○	○	○	○	○	○	○	○	○	○	○	○
Java	○	○	○	○	○	○	○	○	○	○	○	○
Perl	○	○	○	○	○	○	○	○	○	○	○	○
PHP	○	○	○	○	○	○	○	○	○	○	○	○
Python	○	○	○	○	○	○	○	○	○	○	○	○
Ruby	○	○	○	○	○	○	○	○	○	○	○	○
JavaScript	○	○	○	○	○	○	○	○	○	○	○	○
TypeScript	○	○	○	○	○	○	○	○	○	○	○	○
Go	○	○	○	○	○	○	○	○	○	○	○	○
Rust	○	○	○	○	○	○	○	○	○	○	○	○
Shell e.g. Bash	○	○	○	○	○	○	○	○	○	○	○	○
Text Editor/IDE	○	○	○	○	○	○	○	○	○	○	○	○
Other #1	○	○	○	○	○	○	○	○	○	○	○	○
Other #2	○	○	○	○	○	○	○	○	○	○	○	○

5.6 Where did you learn about regular expressions from? (check all that apply)

- [ ] Online Tutorial or Article
- [ ] Reference Book (e.g. O'Reilly)
- [ ] On the Job Experience Through Co-Workers
- [ ] Technical Training Such as a Coding Bootcamp
- [ ] Undergraduate Course in Computer Science
- [ ] Graduate Level Course in Computer Science
- [ ] Undergraduate Course in Another Field
- [ ] Graduate Level Course in Another Field
- [ ] Other: _____

5.7 Estimate your regular expression expertise level:

- ○ **Novice**. For example, you know what repetition operators like * and + do.
- ○ **Intermediate**. For example, you have used more sophisticated features like non-greedy quantifiers */a+?/* and character classes */ \d | \w | [abc] | [^\d]/.*
- ○ **Expert**. For example, you have used features like backreferences */(a+) \1/* and look-ahead/behind assertions */(?<=abc)def/*.
- ○ **Master**. You have written a regular expression engine.

5.8 Estimate the last time you used regular expressions in a professional context:

- ○ In the last week
- ○ In the last month
- ○ In the last year
- ○ More than a year ago

5.9 What is the size of the company you most recently worked at as a professional developer?

- ○ Small (100 employees or less)
- ○ Medium (101 - 999 employees)
- ○ Large (1000 employees or more)

5.10 Is there anything else you would like to say about your **background and general use** of regular expressions such as **how or why you use regular expressions**?

_____

_____

_____

_____

_____

5.11 If you are interested in compensation, provide your email. We will only use this to distribute compensation.

_____

5.12 If you are interested in reading the results of our study, provide your email. We will only use this to distribute results.

_____

5.13 Optional: If you would be willing to participate in a follow-up interview over Zoom (~30 minutes), list the best way to contact you. Interview participants will be compensated an additional $15

_____

End of Block: Background