

RegEx

By El-3aboud web school

What is RegEx?

- RegEx is a sequence of characters that specifies a search pattern in text.

More

Regular Expressions by Computerphile: <https://t.ly/wyqp>

What is that mean?

- /my name is budi/

literal Characters

literal characters is a sequence of literal characters like (cow, rainbow, dog, it's, 222-426, etx...)

Meta Characters

Meta characters are special characters in regex refer to a specific types of characters and they are:

- ``.` for any characters
- ``\d`` for digits characters
- ``\w`` for word characters (a-z and 0-9)
- ``\s`` for spaces characters like whitespace, tab, and new line.

Using “Meta Characters” with uppercase will search for the opposite characters for ex. ``\D`` will search for any character that is not (a-z and 0-9)

See example #1

Quantifiers

Quantifiers specify how many instances of a character, group, or character class must be present in the input for a match to be found.

- ``?`` for 0 or 1
- ``*`` for 0 or more
- ``+`` for 1 or more
- ``{length}`` for a specific length
- ``{min,max}`` for a range of length
- ``.*`` for greedy
- ``.*?`` not greedy

See example #2

Positions

We use position for characters to search for a specific pattern in a certain position

- `^`` for the start of the string
- `$`` for the end of the string
- `\b`` for a word barrier

See example #3

Characters Classes

Character classes (Characters Set) is the stuff that wrapped inside ``[]``. It means find any character that match any of the character set. For example:

- `[a-zA-Z]` set for Latin characters from ``a`` to ``z`` lower and upper case
- `[1-4]` set for digits characters from ``1`` to ``4``

We can mix them up like

- `[a-d1-6]` Latin characters from ``a`` to ``d`` with digits from ``1`` to ``6``
- `[3-1ي-ء]` all Arabic characters from ``ء`` to ``ي`` with digits from ``1`` to ``3``

Also we can use meta characters inside it like:

- `[\w\da-g1-3]+`

Special Characters Inside Characters Classes

- ``.` is a normal dot
- ``^`` it's mean `not` if it comes at the first of the character class for example: `[^a-f]` means find any character that's not `a` to `f`
- ``-`` If the dash is not the first character then it will not be a literal character ``-``. It will mean `between` for example: `[a-h]` means from literal character `a` to the literal character `h` or `[1-5]` means from literal digit `1` to the literal digit `5`
- ``\`` skip character (works outside the character class too)

See example #4

Alternation (Grouping)

Alternation or grouping is a way of wrapping a set of characters with ``()`` and refer to them later in the regex. For example:

- `(com)` means search for a set of literal characters that represent the word ``com``
- `(com|net)` find the set of literal characters that represent the word ``com`` or `-` (which refer to by ``|``) - net

You can refer to each group inside regex by ``$`` followed by the group order (1, 2, 3, etc...)

See example #5

Flags

We can use special characters as a flag for specific rule.

- ``i`` it means ignore case sensitive
- ``g`` means general or all the cases that match the regex

The flag should comes after the regex and not inside for example: `\w+\s?\d+/i` or `\w+\s?\d+/g` or `\w+\s?\d+/ig`

Thank you for your time <3

El-3aboud web school