Losing the fear of

^[Reg]ular[Ex]pressions$

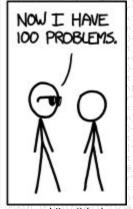^Jonatha (?:Daguerre)?Vasconcelos$

For this talk,
the idea is to go from here...

https://xkcd.com
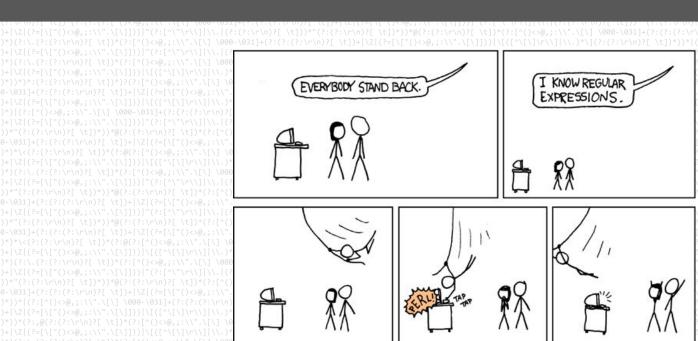
… to here

https://xkcd.com

… or at least, from here

Αυτή είναι μια πρόταση μπορείτε να διαβάσετε για σίγουρος. Είτε κατανόηση νόημά της ή όχι είναι ένα άλλο ζήτημα.

… to here

This is a sentence you can read for sure. Whether you're understanding the whole meaning or not is another question.

What is Regular Expression?

"A sequence of characters that define a search pattern."

"A way to describe a set of strings."

Why should I learn it?

- You can quickly find, replace, change or validate text using lots of program languages, text editors or command line tools.

- The possibilities increase even more when you understand how to combine them.

To understand RegEx, first remember...
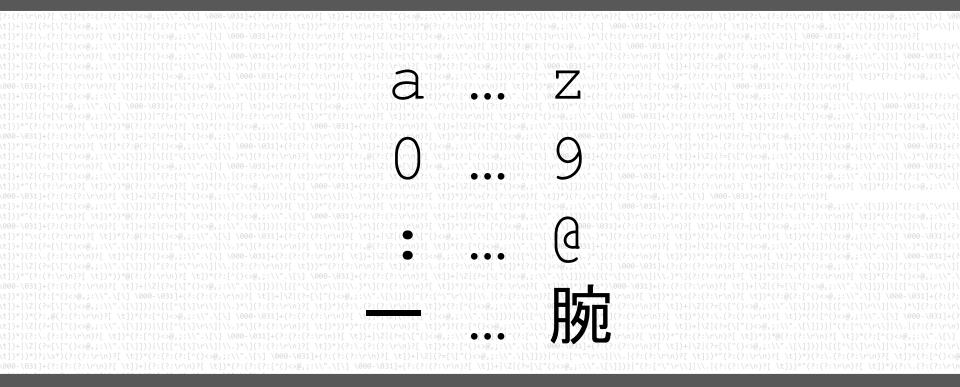
- There is a single idea, but there are many implementations.

  - BRE (POSIX Basic RegEx)
    - grep, ed, sed

  - ERE (GNU Extended RegEx)
    - egrep, gawk, emacs

  - PCRE (Perl 5 Compatible RegEx)
    - Apache, PHP, Java, JavaScript, PCRE Library, Python, Ruby, .NET

To understand RegEx, think about these statements...

- "RegEx are imperative specification of actions to be taken on Strings."

- "They are subroutines."

- "They are an (ugly) program language."

- "Start thinking of each character as an instruction."
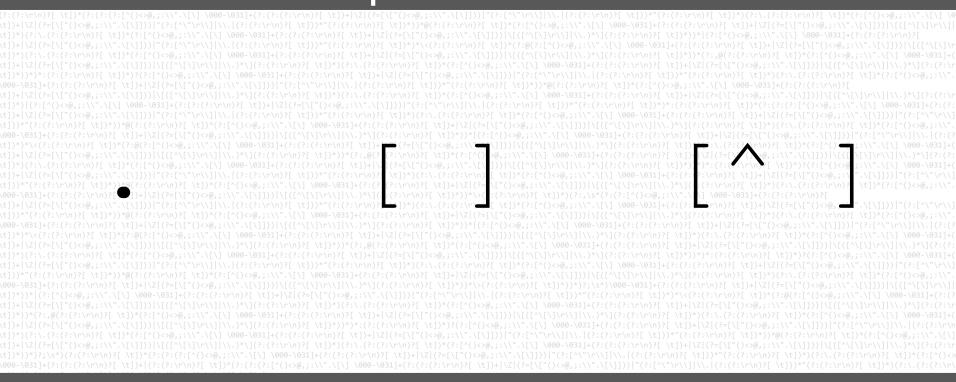
-Damian Conway

Metacharacters
Literals

a ... z

0 ... 9

: ... @

一 ... 腕
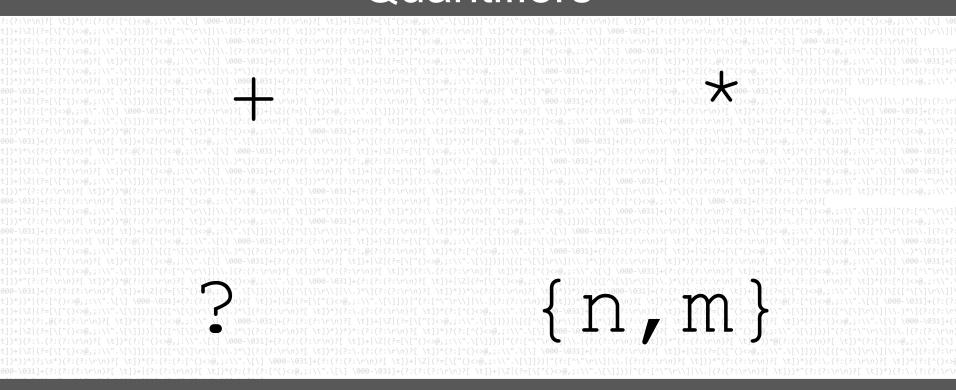
Metacharacters
Anchors
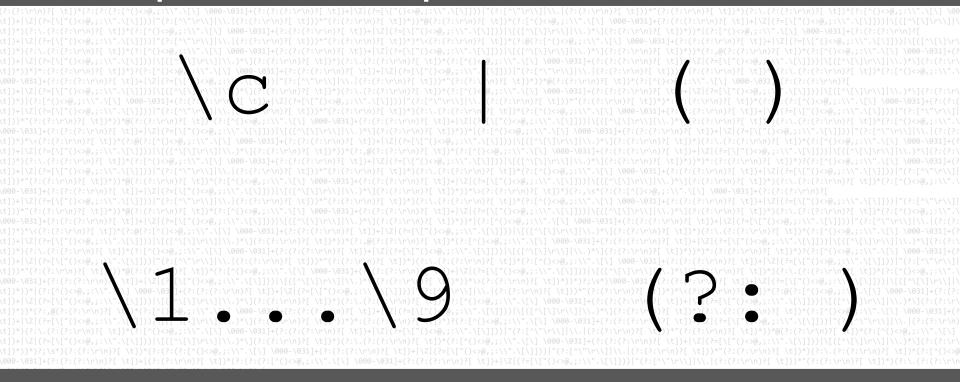
$\wedge$ $\$$ $\backslash b$

Metacharacters
Representatives

• [ ] [^]

Metacharacters

Quantifiers

+

*

?

{n,m}

Metacharacters
Escape, Or, Groups and Backreferences

\c
|
( )
\1...\9
(?:)

Metacharacters

\ letter

$$\backslash w = [a-zA-Z0-9]$$

$$\backslash W = [\char`^a-zA-Z0-9]$$

$$\backslash d = [0-9]$$

$$\backslash D = [\char`^0-9]$$

$$\backslash s = [\ \backslash t\backslash n\backslash r\backslash f\backslash v]$$

Tips to use RegEx better

- Be as specific as possible
- Do not complicate
- Avoid .*
- Avoid [^]
- For accented letters try [À-Üà-ü]
- Prefer non-capturing groups
- When available and applicable, prefer ignore case options
- Sometimes worth to prefer some false-positives instead of creating huge RegExes to exact match

More about RegExes

Lea Verou
Best of Fluent 2012: /Reg(exp){2}lained/:
Demystifying Regular Expressions

https://www.youtube.com/watch?v=EkluES9Rvak

Demian Conway
Understanding and Using Regular Expressions

https://www.infoq.com/presentations/regex
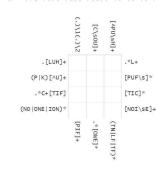
More about RegExes

Aurelio Marinho Jargas

Expressões Regulares - Uma abordagem divertida

http://piazinho.com.br
http://aurelio.net/regex/

Regex Cross-word

Crossword puzzle game, where the crossword clues are defined using regular expressions.

https://regexcrossword.com

Questions\?