

## ANEXO UD3. REGULAR EXPRESSIONS

### 1. INTRODUCCIÓN

Una expresión regular, a menudo llamada también regex, es una secuencia de caracteres que forma un patrón de investigación, principalmente utilizada para la búsqueda de patrones de cadenas de caracteres u operaciones de sustituciones.

#### 1.1. ¿Dónde se utilizan las RegEx?

A continuación se muestra una lista de lenguajes de programación y herramientas que soportan expresiones regulares:

- The .NET Framework
- ActionScript
- ColdFusion
- Java
- JavaScript
- Perl
- PHP
- Python
- Ruby
- Tcl
- Apache HTTP Server
- Google Analytics
- Google Code Search
- Yahoo! Pipes
- Microsoft Exchange Server
- Windows PowerShell
- OpenOffice.org
- Oracle Database
- Cualquier editor de programación o IDE
- Y muchos más!

### 2. RESUMEN DE SÍMBOLOS

Símbolo	Descripción	Ejemplo	Cadenas válidas de ejemplo
	Separador de alternativas	blue   red	blue, red
+	Cuantificador de 1 a infinito	a+	a, aa, aaa, aaaa, ...
?	Cuantificador de 0 o 1	otr?o	oto, otro
*	Cuantificador de 0 a infinito	a*	cadena buida, a, aa, aaa, aaaa. ...
()	Agrupamiento de caracteres	(abc)? Nov(embre iemb re ember)?	cadena buida, abc Nov, Novembre, Noviembre, November
[]	Agrupamiento de caracteres per escoger uno del conjunto	[xyz] [xyz]+	x, y, z xy, xxx, xxz, zyx, zzzzzzyyyyyzx, ...
.	Un carácter cualquiera	a.b	azb, acb, adb, a_b, ...
\	Protector de caracteres reservados	\.	.
-	Rango de caracteres	0-9 A-Z a-z	0, 1, 2, 3, 4, 5, 6, 7, 8, 9 A, B, C, D, E, ..., X, Y, Z a, b, c, d, e, ..., x, y, z
{}	Definición del largo de una cadena	[0-9]{3} [0-4]{3,5} [0-4]{3,}	001, 323, 711, 222, ... (largo 3) 404, 4040, 40404, ... (largo entre 3 i 5) 404, 4040, 40404, 404040, 4040404... (largo de 3 o más)

^	Doble significado: inicio de la línea o negación	^bar [^xyz]	barcelona, barcelonés, barceloneta Cualquier carácter excepto x, y, z
\$	Fin de la línea	[a-z]nil*	pernil, senil, pueril, ...

Simbología específica per a conjuntos de caracteres:

Símbolo	Descripción	Ejemplo	Cadenas válidas de ejemplo
\d	Dígito del 0 al 9	\d+	234, 1111, 00909, ...
\D	Cualquier carácter no dígito	\D*	casa, lugar, centre, ...
\s	Cualquier espacio en blanco	ab\s\sc	ab c
\S	Cualquier carácter no espacio en blanco	\Sa	la, ma, sa...
\w	Cualquier carácter alfanumérico o guion bajo	\w*	mi_casa
\W	Cualquier carácter no alfanumérico o guion bajo	\W	?¿
\b	Límite de palabra	\bmeta\b	meta
\n	Nueva línea	\n	
\r	Retorno de carro	\r	
\t	Tabulación	\t	

### 3. EJEMPLOS

Sample Patterns	
([A-Za-z0-9-]+)	Letters, numbers and hyphens
(\d{1,2}\d{1,2}\d{4})	Date (e.g. 21/3/2006)
([^\s]+(?:\.(jpg gif png)))\2)	jpg, gif or png image
(^[1-9]{1}\$ ^[1-4]{1}[0-9]{1}\$ ^50\$)	Any number from 1 to 50 inclusive
(#?([A-Fa-f0-9]){3}((([A-Fa-f0-9]){3})?))	Valid hexadecimal colour code
((?=[*\d])(?=[*a-z])(?=[*A-Z]).{8,15})	8 to 15 character string with at least one upper case letter, one lower case letter, and one digit (useful for passwords).
(\w+@[a-zA-Z_]+?\.[a-zA-Z]{2,6})	Email addresses
(\<(/?[^\>]+)\>)	HTML Tags

### 4. JUEGO ONLINE SOBRE REGEX

<http://regexcrossword.com/>

## 5. TESTER ONLINE

<http://regex101.com/>

## 6. RECURSOS ADICIONALES

<http://www.rexegg.com/regex-quickstart.html>  
<https://cs.uwaterloo.ca/~dtompkin/teaching/08a/lab7/>