

# Регулярные выражения

**Регулярное выражение** - это шаблон задающий ограничение на вводимые пользователем значения. Оно может состоять как из обычных, так и из специальных символов и их комбинаций.

Обычные символы интерпретируются буквально, например регулярное выражение **абв** интерпретируется как строгая последовательность соответствующих букв в этом порядке.

Помимо этого, для составления более сложных регулярных выражений используются специальные символы. Их написание и значение подробнее расписаны ниже.

## Специальные символы

Символ	Описание	Пример
Общие символы и символы количества		
\	Для <b>обычных символов</b> , означает что следующий символ будет интерпретироваться как специальный. Для <b>специальных символов</b> , означает что следующий символ будет интерпретироваться буквально.	<b>\a</b> - символ <b>a</b> становится специальным символом, означающим границу слова. <b>\*</b> - специальный символ <b>*</b> интерпретируется как обычный символ "звездочка".
^	Следующий символ должен присутствовать в начале строки.	<b>^A</b> - символ <b>A</b> должен быть первым в строке.
\$	Предыдущий символ должен присутствовать в конце строки.	<b>a\$</b> - символ <b>a</b> должен быть последним в строке.
*	Предыдущий символ может быть повторен 0 или больше раз.	<b>a*</b> - символ <b>a</b> может не быть в строке, а может повториться множество раз.
+	Предыдущий символ должен быть повторен 1 или больше раз.	<b>a+</b> - символ <b>a</b> может повториться один раз в строке, а может повториться множество раз.
?	Предыдущий символ должен быть повторен 0 или 1 раз.	<b>a?</b> - символ <b>a</b> может не быть в строке, а может быть один раз.

.	Любой символ кроме переноса строки.	<b>a.c</b> равно <b>abc</b> , <b>a1c</b> и тд.
<b>x y</b>	В строке должен присутствовать либо x либо y.	<b>abc xyz</b> - в строке может быть <b>abc</b> или <b>xyz</b> .
<b>{n}</b>	Предыдущий символ должен быть повторен n раз.	<b>a{2}</b> - <b>a</b> должна быть повторена в строке дважды.
<b>{n,m}</b>	Предыдущий символ должен быть повторен не меньше n раз, но и не больше m раз.	<b>a{2,5}</b> - <b>a</b> должна быть повторена в строке хотя бы дважды и не больше пяти раз.
Скобочные группы		
<b>(x)</b>	Объединение символов в единый элемент (с запоминанием).	<b>abc*</b> - * относится только к <b>c</b> .
<b>(?:x)</b>	Объединение символов в единый элемент (без запоминания).	<b>(abc)*</b> - * относится ко всему выражению в скобках.
<b>x(?=y)</b>	Соответствует x только если за ним идет y.	<b>ab(?=c)</b> - после <b>ab</b> должна быть <b>c</b> .
<b>x(?!y)</b>	Соответствует x только если за ним не идет y.	<b>ab(?!c)</b> - после <b>ab</b> не должно быть <b>c</b> .
Соответствия набору символов		
<b>[xyz]</b>	Соответствует любому из символов в скобках.	<b>[abcd]</b> или <b>[a-d]</b> - строка должна состоять из любых символов от <b>a</b> до <b>d</b> .
<b>[^xyz]</b>	Соответствует любому из символов не в скобках.	<b>[^abcd]</b> или <b>[^a-d]</b> - строка должна исключать символы от <b>a</b> до <b>d</b> .
<b>[\b]</b>	Бэкспейс.	<b>[\b]</b> - находит символ бэкспейса в строке.
<b>\b</b>	Граница слова.	<b>\b</b> - перед или после предыдущего символа должен быть пробел или другая граница слова.
<b>\d</b>	Цифровой символ.	<b>\d</b> - любой из символов от 0 до 9.
<b>\D</b>	Нецифровой символ.	<b>\D</b> - любой из символов кроме символов от 0 до 9.
<b>\s</b>	Пробел.	<b>\s</b> - находит пробел, табуляцию, перенос строки.

<code>\S</code>	Не пробел.	<code>\S</code> - любой из символов кроме пробела, табуляции, переноса строки.
<code>\w</code>	Цифробуквенные символы и нижнее подчеркивание.	<code>\w</code> - любой из символов от 0 до 9, от А до Я и <code>_</code> .
<code>\W</code>	Нецифробуквенные символы.	<code>\W</code> - любой из символов кроме символов от 0 до 9, от А до Я и <code>_</code> .
<code>\0</code>	NULL	<code>\0</code> - находит пустое множество.

# Примеры регулярных выражений и их разбор

## Дата формата ДД.ММ.ГГГГ

**Выражение:** `(0?[1-9]|[12][0-9]|3[01])\.(0?[1-9]|1[012])\.(19|20)\d\d`

<code>(0?[1-9] [12][0-9] 3[01])</code>	
<code>0?[1-9]</code>	0 может быть повторен 1 или 0 раз, а за ним идет любая цифра кроме 0.
<code>[12][0-9]</code>	1 или 2, а за ними любая цифра.
<code>3[01]</code>	3 и 0 или 1.
<code> </code>	или.
<code>(0?[1-9] 1[012])</code>	
<code>1[012]</code>	1 и 0 или 1 или 2.
<code>\.(19 20)\d</code>	
<code>\.</code>	обычный символ точки.
<code>19 20</code>	19 или 20.
<code>\d</code>	любая цифра.

**Возможная строка:** 11.03.2005

## Email

**Выражение:** [^@\\s]+@[^@\\s]+\\.([^@\\s]+)

[^\s]+	
^[^\s]	любой из символов кроме @ и пробела.
+	предыдущий символ может быть повторен 1 или более раз.
@[^\s]+	
@	символ @.
\.[^\s]+	
\.	обычный символ точки.

**Возможная строка:** email@pp.com

## Другие регулярные выражения

## Email без латинских букв

(?:[a-z0-9!#\$%&'\*+,-./:;=?@^\_`{|}~]+(?:\.[a-z0-9!#\$%&'\*+,-./:;=?@^\_`{|}~]+)\*)"(?:[x01-x08x0bx0cx0e-x1fx21x23-x5bx5d-x7f]|\.[x01-x09x0bx0cx0e-x7f])")@(?:([a-z0-9]([a-z0-9]\*[a-z0-9])?.)+[a-z0-9]([a-z0-9]\*[a-z0-9])?)|([[:digit:]]([0-4]([0-9])|1[0-9]([0-9])|1[9]?[0-9])\.)}{3}([[:digit:]]([0-4]([0-9])|1[0-9]([0-9])|1[9]?[0-9])|([a-z0-9]\*[a-z0-9]([a-z0-9]([x01-x08x0bx0cx0e-x1fx21-x5a|x53-x7f]|\.[x01-x09x0bx0cx0e-x7f])\.)+))\.)

Для валидации номера телефона без рамок кодов страны

`[\+\-]?[\d| ]+[\.\,]?[\d| ]*`

**Возможная строка:** +7 900 700 30 10 / 79007003010 / 8900 700 30 10 / + 37 900 700 30 10

## Валидация номера с кодом определенных стран и возможностью выйти

$$(\wedge(\backslash+?7|\backslash+?37)\backslash d\{10\}$|отмена|назад|выход|стоп)$$

Возможная строка: +79007003010 / +379007003010 / +77007003010

## Имя или слово с заглавной

[А-Я]{1}[а-я]{2,20}

Возможная строка: Метабот

---

Версия #2  
Юрий Гарашко создал 12 June 2023 10:59:04  
Ирина Петрова обновил 5 February 2024 11:45:17