

■PHPの正規表現:PCREの主な関数比較

作成:2015.7.27 yoshi

【PCREの構文】

関数	機能	返り値	パラメータ				
			パターン	subject/callback/input	matches/replacement/ subject/limit/flags	flags/limit	offset/count
preg_match	正規表現によるマッチングを行う	pattern が指定した subject にマッチした場合に 1 を返し、マッチしなかった場合は 0、エラーが発生した場合は FALSE を返します。	検索のパターン文字列	[subject] 入力文字列	[matches] matches を指定した場合、検索結果が代入されます。	[flags] PREG_OFFSET_CAPTURE	[offset] 通常、検索は対象文字列の先頭から開始されます。オプションのパラメータ offset を使用して 検索の開始位置を (バイト単位で) 指定することも可能です。
preg_match_all	繰り返し正規表現検索を行う	パターンがマッチした総数を返します (ゼロとなる可能性もあります)。または、エラーが発生した場合に FALSE を返します。	検索のパターン文字列	[subject] 入力文字列	[matches] matches を指定した場合、検索結果が代入されます。	[flags] PREG_PATTERN_ORDER PREG_SET_ORDER PREG_OFFSET_CAPTURE	[offset] 通常、検索は対象文字列の先頭から開始されます。オプションのパラメータ offset を使用して 検索の開始位置を (バイト単位で) 指定することも可能です。
preg_replace preg_filter	正規表現検索および置換を行う 注): preg_filter は、マッチした結果 (を交換したもの) のみを返します。	subject 引数が配列の場合は配列を、その他の場合は文字列を返し、パターンがマッチした場合、[置換が行われた]新しい subject を返します。マッチしなかった場合、subject をそのまま返します。エラーが発生した場合、NULL を返します。	検索のパターン文字列又は配列	[subject] 検索・置換対象となる文字列もしくは文字列の配列 配列の場合、検索と置換は subject の各要素に対して行われ、返り値も配列となります。	[replacement] 置換を行う文字列もしくは文字列の配列	[limit] subject 文字列において、各パターンによる 置換を行う最大回数。デフォルトは -1 (制限無し)。	[count] この引数が指定されると、置換回数が渡されます。
preg_replace_callback	正規表現検索を行い、コールバック関数を使用して置換を行う	subject が配列の場合には配列を、それ以外の場合は文字列を返します。エラー時の返り値は NULL となります。マッチするものが見つかった場合は新しい subject を返し、それ以外の場合はもとの subject をそのまま返します。	検索のパターン文字列又は配列	[callback] コールバック関数は、検索対象文字列でマッチした要素の配列が指定されて コールされます。(返り値が、文字列であること)	[subject] 文字列あるいは文字列の配列で、検索および置換の対象となる文字列を指定します。	[limit] subject 文字列において、各パターンによる 置換を行う最大回数。デフォルトは -1 (制限無し)。	[count] この引数が指定されると、置換回数が渡されます。
preg_split	正規表現で文字列を分割する。指定した文字列を、正規表現で分割します。	pattern にマッチした境界で分割した subject の部分文字列の配列を返します。	検索のパターン文字列	[subject] 入力文字列	[limit] 最大 limit 個の部分文字列を返します。残りの文字列は、最後の部分文字列に含めて返されます。limit が -1、0 あるいは NULL の場合は「制限が無い」ことを意味します。制限を指定せずに flags パラメータを指定したい場合などに NULL を使用します。	[flags] PREG_SPLIT_NO_EMPTY PREG_SPLIT_DELIM_CAPTURE PREG_SPLIT_OFFSET_CAPTURE	
preg_grep	パターンにマッチする配列の要素を返す。input 配列の要素のうち、指定した pattern にマッチするものを要素とする配列を返します。	input 配列のキーを使用した配列を返します。	検索のパターン文字列	[input] 入力の配列	[flags] PREG_GREP_INVERT		

【matchesの詳細内容】

match	\$matches[0] にはパターン全体にマッチしたテキストが代入され、\$matches[1] には 1 番目のキャプチャ用サブパターンにマッチした文字列が代入され、といったようになります。
match_all	\$matches[0] はパターン全体にマッチした文字列の配列、\$matches[1] は第 1 のキャプチャ用サブパターンにマッチした文字列の配列、といった順番となります。

【各関数のflags】

match	PREG_OFFSET_CAPTURE	各マッチに対応する文字列のオフセットも返されます。これにより matches の値は配列となり、配列の要素 0 はマッチした文字列、要素 1 は対象文字列中におけるマッチした文字列のオフセット値 となることに注意してください。
match_all	PREG_PATTERN_ORDER	\$matches[0] はパターン全体にマッチした文字列の配列、\$matches[1] は第 1 のキャプチャ用サブパターンにマッチした文字列の配列、といった順番となります。
	PREG_SET_ORDER	\$matches[0] は 1 回目のマッチングでキャプチャした値の配列、\$matches[1] は 2 回目のマッチングでキャプチャした値の配列、といった順序となります。
split	PREG_OFFSET_CAPTURE	各マッチに対応する文字列のオフセットも返されます。これにより、返り値は配列となり、配列の要素 0 はマッチした文字列、要素 1 は subject におけるマッチした文字列のオフセット値となることに注意してください。
	なし	PREG_PATTERN_ORDER 指定
	PREG_SPLIT_NO_EMPTY	空文字列でないものだけが preg_split() により返されます。
	PREG_SPLIT_DELIM_CAPTURE	文字列分割用のパターン中の カッコによるサブパターンでキャプチャされた値も同時に返されます。
grep	PREG_SPLIT_OFFSET_CAPTURE	各マッチに対応する文字列のオフセットも返されます。これにより、返り値は配列となり、配列の要素 0 はマッチした文字列、要素 1 は subject におけるマッチした文字列のオフセット値となることに注意してください。
	PREG_GREP_INVERT	与えた pattern にマッチ しない 要素を返します。