

# SYSTEMY BAZ DANYCH

## SQL



**TEMAT 21-07:** Wyrażenia regularne w języku SQL.

Autor dokumentu: Wojciech Galiński

poniedziałek, 21 listopada 2016 r.

351203 Technik informatyk

ŹRÓDŁA WIEDZY: <http://dev.mysql.com/doc/refman/5.1/en/regexp.html>.

## Zagadnienia obowiązkowe

1. **Wyrażenia regularne** – to wzorce, które opisują łańcuchy znaków. Mogą one określać zbiór pasujących łańcuchów, mogą również wyszczególniać istotne części łańcucha (definicja opracowana na podstawie Wikipedii).
2. **Operatory „[NOT] REGEXP”, „[NOT] RLIKE”** – służą do tworzenia warunków z wykorzystaniem wyrażeń regularnych w języku SQL. Oto podstawowe wzorce:

```
SELECT * FROM tabela WHERE pole_K [NOT] REGEXP 'wyrazenie_regularne';
```

```
SELECT * FROM tabela WHERE pole_K [NOT] RLIKE 'wyrazenie_regularne';
```

3. **Metaznaki wyrażeń regularnych dla języka SQL (MySQL)** – opisuje je poniższa tabela:

ZNAK	ZNACZENIE	PRZYKŁAD RLIKE
<b>^</b>	Początek ciągu	SELECT * FROM uczen WHERE wzrost RLIKE '^2';
<b>\$</b>	Koniec ciągu	SELECT * FROM uczen WHERE wzrost RLIKE '0\$';
<b>.</b>	Dokładnie 1 dowolny znak	SELECT * FROM uczen WHERE wzrost RLIKE '..5';
<b>a b</b>	Wystąpienie znaków „a” albo „b” albo obydwu naraz	SELECT * FROM uczen WHERE stypendium REGEXP '85 430';
<b>[]</b>	Dowolny znak ze zbioru wartości (kolejność nieistotna). Znak „-” oznacza przedział wartości. Patrz: punkty 4-7.	SELECT * FROM uczen WHERE imie REGEXP '[az]\$';
<b>[^]</b>	Dowolny znak spoza zbioru wartości	SELECT imie,nazwisko FROM uczen WHERE nazwisko REGEXP '[^azk]\$';
<b>{N}</b>	Liczba wystąpień (obok siebie) wynosi dokładnie N	SELECT stypendium FROM uczen WHERE stypendium REGEXP '3{2}';
<b>{M,N}</b>	Liczba wystąpień (obok siebie) wynosi pomiędzy M i N	SELECT stypendium FROM uczen WHERE stypendium REGEXP '3{2,3}';
<b>{N,}</b>	Liczba wystąpień (obok siebie) wynosi co najmniej N	SELECT stypendium FROM uczen WHERE stypendium REGEXP '3{1,}';
<b>a?</b>	Liczba wystąpień znaku „a” wynosi 1 lub mniej; to samo, co <b>a{0,1}</b>	SELECT imie FROM uczen WHERE imie REGEXP 'a?';
<b>a+</b>	Liczba wystąpień znaku „a” wynosi 1 lub więcej; to samo, co <b>a{1,}</b>	SELECT * FROM uczen WHERE wzrost REGEXP '0+';
<b>()</b>	Sekwencje znaków (istotna kolejność)	SELECT data_urodzenia FROM uczen WHERE data_urodzenia REGEXP '(1.*){4}';
<b>a*</b>	Liczba wystąpień znaku „a” wynosi 0 lub więcej; to samo, co <b>a{0,}</b>	

Stosowanie znaków specjalnych (m. in. + - \* ^ . \) w wyrażeniach regularnych wymaga zastosowania dodatkowego backslasha poprzedzającego (\).

WIĘCEJ PRZYKŁADÓW:

```
SELECT nazwisko, nazwisko REGEXP '^.$' FROM uczen; -- LIKE '_';
```

```
SELECT nazwisko FROM uczen WHERE nazwisko REGEXP '^K'; -- LIKE 'K%'
```

```
SELECT email FROM nauczyciel WHERE email REGEXP '^[a-z0-9]+@[1]{1}[\.a-z0-9]*$'; -- adres e-mail ?
```

4. **Kody znaków w wyrażeniach regularnych** – reprezentowane są za pomocą wzorca:

```
[ [.nazwa_kodu. ] ]
```

W miejsce „nazwa\_kodu” wstawiamy jedną z pozycji z poniższej tabeli.

5. Lista wybranych kodów znaków w wyrażeniach regularnych – przedstawione są w poniższej tabeli:

newline	'\n'		HT, tab	'\t'	011	IS2, RS		036	seven	'7'	067
carriage-return	'\r'		LF		012	IS1, US		037	eight	'8'	070
apostrophe	'\''		VT, vertical-tab	'\v'	013	space	' '	040	nine	'9'	071
quotation-mark	'\"'		FF, form-feed	'\f'	014	exclamation-mark	'!'		colon	':'	
plus-sign	'+'		CR		015	number-sign	'#'		semicolon	';'	
tilde	'~'		S0		016	percent-sign	'%'		less-than-sign	'<'	
backslash, reverse-solidus	'\\'		SI		017	ampersand	'&'		equals-sign	'='	
asterisk	'*'		DLE		020	left-parenthesis	'('		greater-than-sign	'>'	
circumflex, circumflex-accent	'^'		DC1		021	right-parenthesis	)'		question-mark	'?'	
dollar-sign	'\$'		DC2		022	comma	','		commercial-at	'@'	
			DC3		023	hyphen, hyphen-minus	'_'		left-square-bracket	'['	
NUL		000	DC4		024				right-square-bracket	']'	
SOH		001	NAK		025	period, full-stop	'.'		underscore, low-line	'_'	
STX		002	SYN		026	slash, solidus	'/'				
ETX		003	ETB		027	zero	'0'	060	grave-accent	'`'	
EOT		004	CAN		030	one	'1'	061	left-brace, left-curly-bracket	'{'	
ENQ		005	EM		031	two	'2'	062	vertical-line	' '	
ACK		006	SUB		032	three	'3'	063	right-brace, right-curly-bracket	'}'	
BEL, alert		007	ESC		033	four	'4'	064			
BS, backspace	'\b'	010	IS4, FS		034	five	'5'	065			
			IS3, GS		035	six	'6'	066	DEL		0177

PRZYKŁADY:

```
SELECT wzrost FROM uczen WHERE wzrost REGEXP '[[.zero.]]'; -- LIKE '%0%'
-- LIKE '%\\%-%' -- to polecenie LIKE nie działa prawidłowo nas serwerze MySQL 5.5
SELECT zdjecie FROM uczen WHERE zdjecie regexp '\\\\.\\.*\\~';
SELECT zdjecie FROM uczen WHERE zdjecie regexp '[[.\\.\\.]].*[[.~.]]';
SELECT zdjecie FROM uczen WHERE zdjecie regexp '[[.backslash.]].*[[.tilde.]]';
```

6. Klasy znaków w wyrażeniach regularnych – reprezentowane są za pomocą wzorca:

**[[ :nazwa\_klasy: ]]**

W miejsce „nazwa\_klasy” wstawiamy jedną z pozycji z poniższej tabeli.

7. Lista klas znaków w wyrażeniach regularnych – przedstawiona jest w poniższej tabeli:

<b>alnum</b>	znaki alfanumeryczne	<b>graph</b>	znaki graficzne	<b>space</b>	spacja, tabulacja, znak nowej linii oraz powrotu karetki
<b>alpha</b>	litery alfabetu	<b>lower</b>	małe litery alfabetu		
<b>blank</b>	białe znaki	<b>print</b>	znaki graficzne lub spacje	<b>upper</b>	duże litery alfabetu
<b>cntrl</b>	znaki kontrolne	<b>punct</b>	znaki interpunkcyjne	<b>xdigit</b>	znaki heksadecymalne
<b>digit</b>	cyfry				

PRZYKŁADY:

```
SELECT zdjecie, zdjecie REGEXP '[[ :punct: ]]' FROM uczen; -- zawiera znaki interpunkcyjne
SELECT nazwisko FROM uczen WHERE nazwisko REGEXP '^[[ :upper: ]]$'; -- tylko duże litery
```

## Pytania kontrolne

1. Czym są wyrażenia regularne i do czego stosujemy je w języku SQL?
2. Jakie operatory języka SQL wykorzystują wyrażenia regularne?
3. Czym są metaznaki wyrażen regularnych? Jak je stosować w języku SQL?
4. Do czego służą kody znaków oraz klasy znaków w wyrażeniach regularnych?

## Zadania

1. Przeanalizuj wyniki zwracane przez przykłady zawarte w powyższym temacie (baza danych „i\_dzennik”).
2. Przeanalizuj wyniki zwracane przez przykłady zawarte pod adresem: <http://dev.mysql.com/doc/refman/5.1/en/regexp.html>.
3. Poszukaj w Internecie więcej przykładów wyrażen regularnych i przeanalizuj je.
4. Znajdź budowę adresu e-mail i wykonaj wyrażenie regularne wyświetlające prawidłowe adresy e-mail z tabeli „nauczyciel”.