

תבניות טקסט

[מהי תבנית טקסט?](#)

[שימוש ב-Meta Characters](#)

[שימוש ב-Grouping Expression](#)

[שימוש ב-Quantifiers](#)

[תת תבניות](#)

[הפונקציה preg_match](#)

[הפונקציה preg_match_all](#)

[הפונקציה str_replace](#)

[הפונקציה preg_replace](#)

מהי תבנית טקסט?

תבנית טקסט (Regular Expression) היא מחרוזת תווים שמייצגת קבוצה של מחרוזות תווים. קבוצה זו יכולה לכלול מחרוזת תווים אחת או יותר.

כל תבנית טקסט מייצגת, למעשה, קבוצה של כללים שעל פיהם ניתן להתאים לתבנית הטקסט קבוצה של מחרוזות תווים.

תבניות טקסט שימושיות באותם מקרים שבהם איננו יודעים מהי מחרוזת התווים הספציפית שאנו מצפים למצוא. כך למשל, בהתמודדות עם SQL Injectinos ניתן להגדיר תבנית טקסט שכל מחרוזת התווים שמתאימות לה הן כל מחרוזות התווים החשודות כניסיון לחדור לבסיס הנתונים.

שפת התיכנות PHP תומכת בתבניות טקסט שמבוססות על כללי התחביר ב-Perl.

כל תבנית טקסט (Regular Expression) ב-PHP חייבת להתחיל ולהסתיים בתו './'.

אחת הפונקציות השימושיות בעבודה עם תבניות טקסט ב-PHP היא הפונקציה `preg_match` יכולה לקבל שני ארגומנטים ובודקת אם יש התאמה בין תבנית הטקסט (הארגומנט הראשון) ומחרוזת התווים שנבדקת (הארגומנט השני). התכנית הבאה מדגימה שימוש פשוט בפונקציה זו.

```
<?php
    $rgx = "/israel/";
    $str = "moshe israeli";
    if (preg_match($rgx,$str))
    {
        echo "match";
    }
    else
    {
        echo "no match";
    }
?>
```

הפלט שנקבל:

match

הפונקציה `preg_match_all` פועלת באופן דומה ומאפשרת לשלוח כארגומנט שלישי מערך דו מימדי אשר יתמלא בכל ההתאמות שנמצאו.

הדוגמא הבאה מציגה את אופן הפעלתה של פונקציה זו.

```
<?php
    $rgx = "/1\d\d\d/";
    $str = "1920_1830_2030";
    $matrix = array();
    preg_match_all($rgx,$str,$matrix);
    print_r($matrix);
?>
```

הפלט:

```
Array ( [0] => Array ( [0] => 1920 [1] => 1830 ) )
```

באמצעות הפונקציה `str_replace` ניתן להחליף כל התאמה לתבנית הטקסט בתת מחרוזת אחרת. הפעלת הפונקציה בצורה הפשוטה ביותר דורשת שליחה של 3 ארגומנטים: תבנית הטקסט, מחרוזת התווים שאמורה להחליף כל התאמה שנמצאת עם תת המחרוזת ומחרוזת התווים שבה יש לחפש את ההתאמות. באמצעות הפונקציה `preg_replace` ניתן לבצע פעולה דומה תוך שליחת ארגומנט ראשון שהוא תבנית טקסט כך שהחיפוש יתבצע אחרי תתי מחרוזות שמתאימות לה.

הדוגמא הבאה מציגה את אופן הפעלתה של הפונקציה `preg_replace`.

```
<?php
    $rgx = "/1\d\d\d/";
    $str_before = "1920_1830_2030";
    $vecA = array();
    $vecA[0] = $rgx;
    $vecB[0] = "yoyo";
    $str_after = preg_replace($rgx,"uouo",$str_before);
    echo $str_after;
?>
```

הפלט שנקבל הוא:

```
uouo_uouo_2030
```

שימוש ב-Meta Characters

כל Meta Character מייצג קבוצה (set) של characters. כל Meta Character שכלול בתוך תבנית טקסט מייצג התאמה ל character אחד במחרוזת שבינה ובין תבנית הטקסט יש התאמה. כל ה-characters שיכולים להתאים ל Meta Character מסויים מהווים את הקבוצה (set) שתבנית הטקסט מייצגת.

ה-Meta Characters הפופולריים ביותר הם:

.	התאמה לכל תו מכל סוג
^	התאמה לתחילתה של מחרוזת תווים
\$	התאמה לסופה של מחרוזת תווים
\s	התאמה לכל ריווח
\d	התאמה לכל סיפרה
\w	התאמה לכל תו שהוא סיפרה או אות או קו תחתי

הדוגמא הבאה מציגה שימוש בסיסי ב-Meta Characters.

```
<?php
$rgx = "/\d\d\d\d\w/";
$str = "moshe israeli 197_";
if(preg_match($rgx,$str))
{
    echo "match";
}
else
{
    echo "no match";
}
?>
```

הפלט שנקבל הוא:

match

הפונקציה preg_match בודקת אם ניתן למצוא במחרוזת התווים הנתונה תת מחרוזת אשר מתאימה לתבנית הטקסט הנתונה. אם הפונקציה מוצאת תת מחרוזת אחת (או יותר) היא מחזירה true.

שימוש ב-Grouping Expression

אוסף של characters אפשריים כהתאמה ל-character בודד נקרא Grouping Expression.

כל Grouping Expression מתוחם בתוך סוגריים מרובעות.

```
[a-d\d]
```

'a' or 'b' or 'c' or 'd' or a digit

```
[1-3\w]
```

'1' or '2' or '3' or any word character

הדוגמא הבאה מציגה שימוש בסיסי ב-Grouping Expression.

```
<?php
    $rgx = "/mo[s-z]he/";
    $str = "moshe israeli 197_";
    if(preg_match($rgx,$str))
    {
        echo "match";
    }
    else
    {
        echo "no match";
    }
?>
```

הפלט שנקבל:

match

שימוש ב-Quantifiers

ה-quantifier מייצג את מספר הפעמים ש-character מסויים (או Meta Character) יכול להופיע במחרוזת התווים שעימה נבדקת שאלת ההתאמה.

שפת התיכנות PHP תומכת ב-quantifiers הבאים:

*	ה-character יכול להופיע 0 (או יותר) פעמים.
+	ה-character יכול להופיע פעם אחת (או יותר).
?	ה-character יכול להופיע 0 או פעם אחת.
{n,m}	ה-character יכול להופיע לפחות n פעמים ולא יותר מ-m.
{n,}	ה-character יכול להופיע לפחות n פעמים
{,m}	ה-character יכול להופיע לכל היותר m פעמים

הדוגמא הבאה מציגה שימוש בסיסי ב-quantifiers:

```
<?php
    $rgx = "/\d{2,3}/";
    $str = "moshe israeli 197_";
    if (preg_match($rgx,$str))
    {
        echo "match";
    }
    else
    {
        echo "no match";
    }
?>
```

הפלט יהיה:

match

תת תבניות

תת תבנית (Sub Expression) היא תבנית טקסט (Regular Expression) שמוכלת בתוך תבנית טקסט (Regular Expression) אחרת. תת תבנית (Sub Expression) מתוחמת באמצעות סוגריים עגולות.

דוגמא:

```
/a (fgh.) f/
```

תבנית הטקסט (Regular Expression) הזו תואמת כל מחרוזת שמתחילה באות a ולאחריה fgh ולאחריה כל תו ולאחריה האות f.

```
<?php
    $rgx = "/a (fgh.) f/";
    $str = "afgh9f";
    if (preg_match($rgx, $str))
    {
        echo "match";
    }
    else
    {
        echo "no match";
    }
?>
```

הפלט של התכנית יהיה:

```
match
```

ניתן להשתמש בתת תבנית (Sub Expression) בצירוף של quantifier.

דוגמא:

```
/a (fg)?b/
```

תבנית טקסט (Regular Expression) זו תואמת כל מחרוזת תווים שמתחילה ב-a ולאחריה צירוף האותיות fg פעם אחת (או בכלל לא) ולאחריה האות b. במילים פשוטות, תבנית טקסט זו תואמת את כל אחת ממחרוזות התווים הבאות:

ab

afgb