

POLSKIE ZNAKI

Jedną z ważnych decyzji które należy podjąć rozpoczynając tworzenie nowej strony WWW jest określenie sposobu kodowania polskich znaków. Trzy standardy które zazwyczaj bierze się pod uwagę w takiej sytuacji to ISO-8859-2, Windows-1250 (zwykle odradzany) i UTF-8 (standard bazujący na Unicode, polecany do wielojęzycznych stron WWW). Obecnie najpopularniejszy jest system kodowania UTF-8 i jego będziemy używać.

Wyświetlanie polskich liter na stronie

Standard kodowania ustawiamy w sekcji HEAD:

```
<!DOCTYPE HTML PUBLIC "-//W3//DTD HTML 4.01//EN" http://www.w3.org/TR/html4/strict.dtd">

<html>

<head>

<meta http-equiv="content-type" content="text/html; charset=UTF-8">

</head>
```

Przeglądarki wyświetlają stronę w takim kodowaniu, jakie zostało im przedstawione w nagłówku protokołu. Bardzo często PHP ma wymuszone wysyłanie dokumentów w jednym, określonym systemie kodowania znaków. należy więc samemu nadpisać nagłówek HTTP w skrypcie PHP. Służy do tego polecenie [header](#) :

```
<?header('Content-Type: text/html; charset=utf-8');?>
```

Ustawienie parametru nagłówka HTTP przy pomocy w/w komendy, musi nastąpić zanim wyślemy jakiegokolwiek inne dane do przeglądarki, czyli ostatecznie :

```
<?header('Content-Type: text/html; charset=utf-8');?>
<!DOCTYPE HTML PUBLIC "-//W3//DTD HTML 4.01//EN" http://www.w3.org/TR/html4/strict.dtd">

<html>

<head>

<meta http-equiv="content-type" content="text/html; charset=UTF-8">

</head>
```

Kolejna rzecz, to **edytor**, w którym piszemy stronę. Należy w nim wybrać sposób kodowania (UTF-8).

Niestety wykonanie tych czynności nie zawsze oznacza koniec problemów z polskimi znakami - może zdarzyć się sytuacja że dane odczytywane z bazy danych, pliku XML lub też pobierane z Internetu będą używać innego standardu kodowania znaków. W takiej sytuacji konieczne będzie dokonanie odpowiedniej konwersji.

Zapis polskich liter do bazy

W celu zapisania polskich znaków do bazy, przed ich wysłaniem, należy wysłać zapytanie:

```
mysql_query("SET NAMES 'utf8'");
```

Należy również sprawdzić w bazie danych sposób kodowania i jeżeli konieczne, to również zmienić na UTF-8.

Wyświetlanie polskich liter w PDF-ie

Czcionka wygenerowana do zapisu polskich znaków w PDF-ie korzysta ze standardu ISO-8859-2 (patrz oddzielna instrukcja). Jeżeli chcemy więc odczytać poprawnie polskie znaki z bazy (zapisane w UTF-8) musimy dokonać ich konwersji na ISO-8859-2. Do tego celu przydaje się funkcja **iconv()**.

Nagłówek funkcji `iconv()` w PHP wygląda następująco:

```
string iconv ( string in_charset, string out_charset, string str )
```

Oznacza to że funkcja ta przyjmuje trzy parametry:

1. sposób kodowania znaków w tekście przekazanym jako trzeci parametr;
2. docelowy sposób kodowania znaków;
3. tekst dla którego ma zostać zmienione kodowanie znaków.

Przykładowo konwersję tekstu z UTF-8 na ISO-8859-2 można wykonać następująco:

```
$str = iconv("utf-8", "iso-8859-2", $str);
```