

UTF8-ו Wordpress

מה זה יוניקוד (Unicode) ו-UTF8

יוניקוד (Unicode) הוא תקן בינלאומי לייצוג טקסט במערכות מחשב. התקן מגדיר מערכת תווים המקיפה את כל מערכות הכתב הנמצאות כיום בשימוש פעיל בשפות העולם. לפני שהחל השימוש בתקן זה, התקן המקובל לייצוג טקסט היה ASCII. תקן זה הכיל 256 תווים: 128 תווי ASCII ו-128 תווים אחרים לפי האזור, כך שאותו מספר בעצם היה יכול להיות אותיות שונות. למשל, המספר 244 הכיל גם את האות à בקידוד של מערב אירופה וגם את האות א בקידוד העברי. מי שעייין בטקסט ראה את האות בהתאם לקידוד בו הוא צפה. אם היה בוחר לצפות בקידוד מערב אירופה היה רואה את האות à, אם היה בוחר לצפות בקידוד עברי, היה רואה את אותו טקסט עם האות א. דבר זה גרם לבעיות בהעברה של טקסט ממערכות שונות. בשנת 1990 התכנסו חברות תוכנה ותקשורת עולמיות כדי לפתור את הבעיה וכך החלה התקינה של יוניקוד. כיום, השימוש ביוניקוד מאפשר לייצג כל אות, בכל שפה, במספר ייחודי לה, ולכן ניתן לכתוב באותו הטקסט אותיות שפעם "יישוב" על אותו מספר.

UTF-8 הוא למעשה שם של צורת דחיסה עבור יוניקוד המשתמשת ב-8 תווים: 8 bit **Unicode Transformation Format**. כיום יש גם דחיסה ב-16 תווים וב-32 תווים.

Wordpress יכולה לעבוד בקידוד utf8 לעיתים יש לבצע התאמות בקבצי המערכת ובקבצי התבנית.

שינוי קובץ wp-db.php

ראשית, יש לבצע שינוי באחד מקבצי המערכת, קובץ **wp-db.php** הנמצא בתיקיית **includes**.

יש לפתוח קובץ זה בכל עורך טקסט או עורך HTML ולמצוא את השורה:

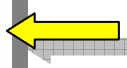
```
$this->select($dbname);
```

ולוסיף מעליה את השורה:

```
$this->query('SET NAMES UTF8');
```

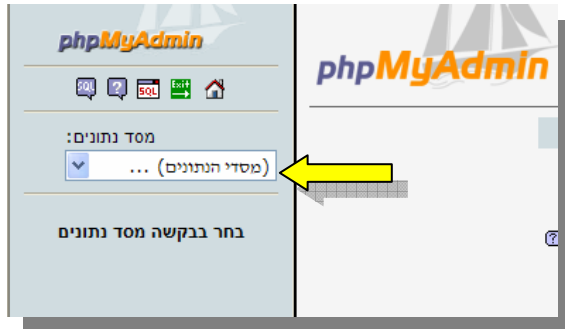
כך:

```
}  
$this->query('SET NAMES UTF8');  
$this->select($dbname);  
}
```

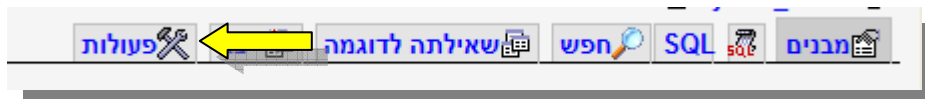


הכנת בסיס הנתונים עבור UTF

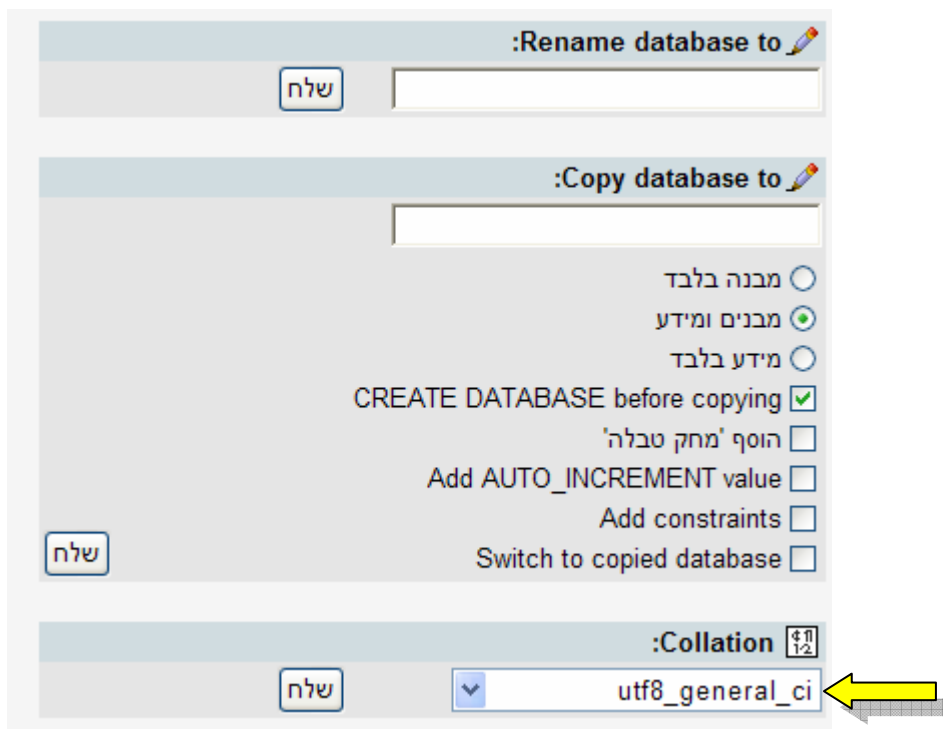
לפני התקנת האתר, בעת יצירת בסיס הנתונים יש לוודא כי הוא בקידוד UTF-8. בפאנל הניהול של חשבון האירוח (למשל בפאנל מסוג direct admin) ניתן ליצור בסיס נתונים בצורה מאד קלה, אולם בשלב היצירה בפאנל זה, לא ניתן לציין את קידוד בסיס הנתונים והוא ייוצר תחת הגדרות ברירת המחדל של השרת. ניתן להיכנס ל-phpMyAdmin ולשנות את קידוד בסיס הנתונים שכבר נוצר. יש לבחור את בסיס הנתונים הרצוי:



לאחר שנבחר בסיס הנתונים, יש להיכנס ל"פעולות" שבסרגל הניווט העליון

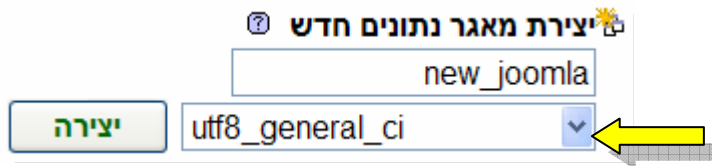


בתחתית החלון שיפתח, תחת Collation יש לבחור את הקידוד: utf8_general_ci



כעת בסיס הנתונים מוכן להתקנה.

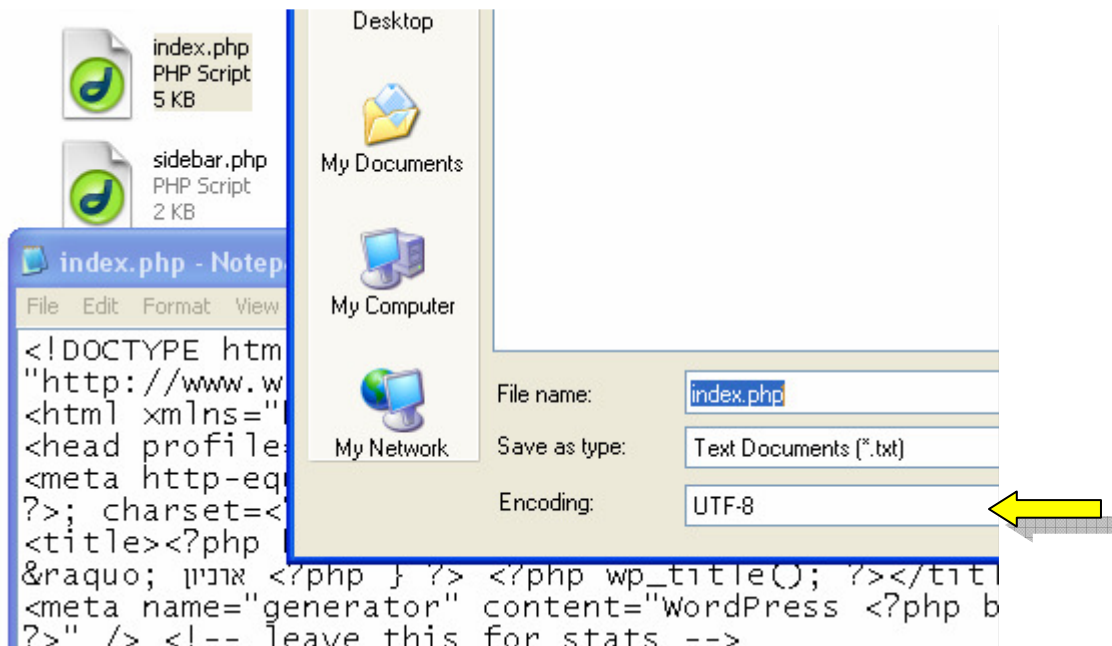
בעת יצירת בסיס נתונים חדש ב- phpMyAdmin ולא דרך ממשק ניהול אחר, ניתן לציין את הקידוד כבר בעת יצירת בסיס הנתונים.



התאמת התבנית לקידוד UTF-8

גם אם הגדרות האתר מורות על שימוש בקידוד utf8 ייתכן כי תבנית האתר נשמרה תחת קידוד אחר. במקרה כזה, יש להתאים את תבנית העיצוב לקידוד utf8, באמצעות שמירת הקבצים מחדש תחת קידוד זה.

תבניות העיצוב שמורות בתיקיית wp-content/themes, כאשר לכל תבנית קיימת תיקייה הנושאת את שמה. יש לפתוח את קבצי ה-PHP שבתבנית ולשמור אותם תחת קידוד utf8. ניתן לעשות זאת בקלות באמצעות Notepad. לאחר שהקובץ נפתח ב-Notepad יש לבחור ב"שמירה בשם" ובחלון השמירה, מתחת לשם הקובץ וסוג הקובץ יש לבחור בקידוד UTF-8.



המרת אתר שהותקן על קידוד ASCII ל-UTF8

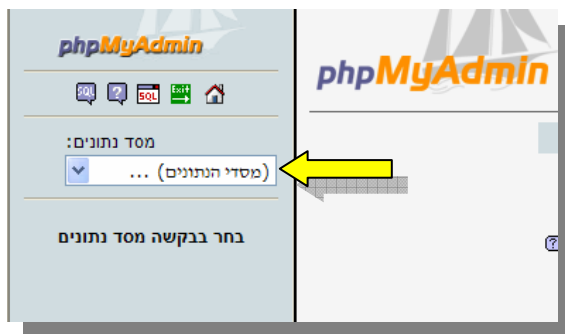
שים לב:

- יש לבצע גיבוי מלא של האתר ושל בסיס הנתונים לפני ביצוע פעולה זו.
- הפעולה היא על אחריות המשתמש.

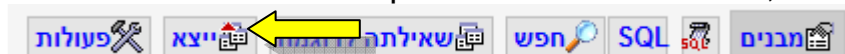
יש לבצע התאמה של קבצי התבנית כפי שהוסבר קודם לכן.

התאמת בסיס הנתונים לאתר קיים

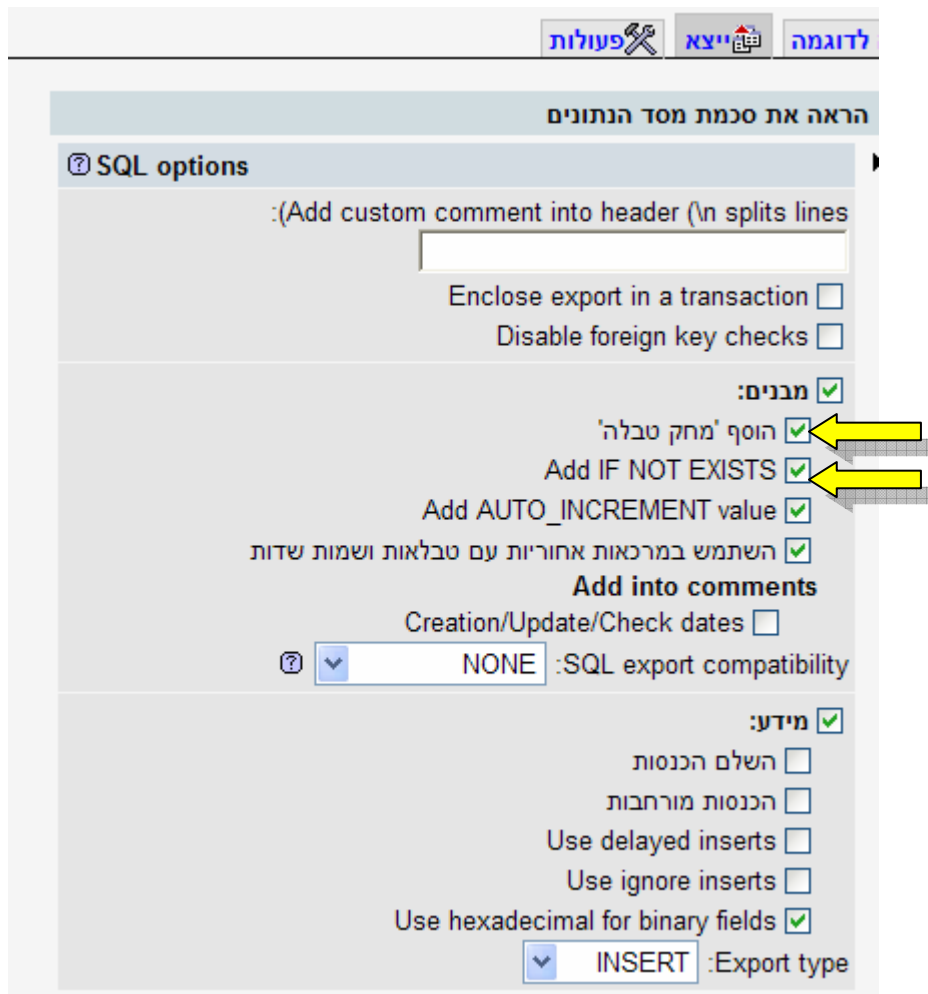
ב-PhpMyAdmin ניתן לייצא את בסיס הנתונים לקובץ SQL. יש לבחור בסיס הנתונים הרצוי



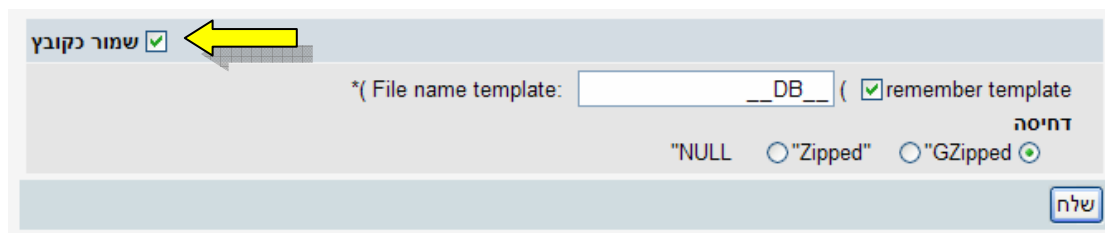
לאחר שנבחר בסיס הנתונים, יש לבחור בסרגל הניווט העליון ב"ייצא".



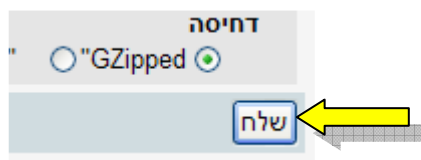
במסך הייצוא יש לסמן את האפשרויות "הוסף 'מחק טבלה'" ואת " Add " AUTO_INCREMENT value".



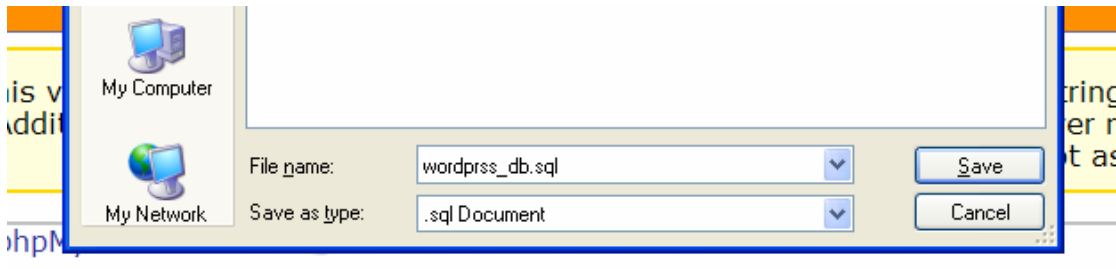
בתחתית מסך הייצוא יש לסמן את "שמור כקובץ"



ולבצע שמירה



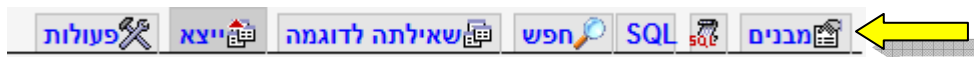
קובץ מסוג SQL ישמר אל המחשב



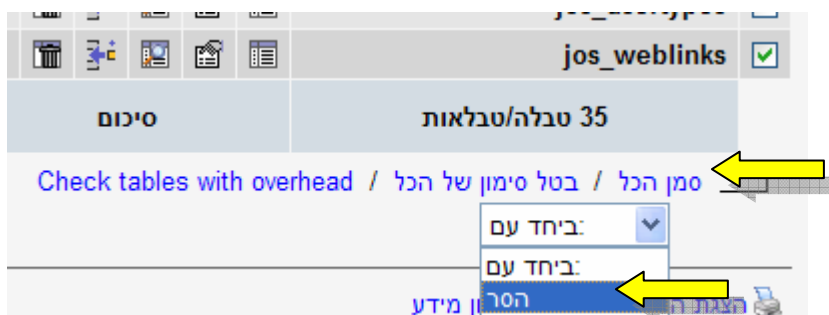
כעת יש לגשת אל קובץ זה, לפתוח אותו באמצעות NotePad ולבצע שמירה חוזרת ולשמור בקידוד Unicode (כפי שהוסבר עבור קבצי התבנית, "שמירה בשם"...)

כעת בסיס הנתונים שבקובץ ה-SQL הוסב ל-UTF8 ויש להתאים את בסיס הנתונים ל-UTF8.

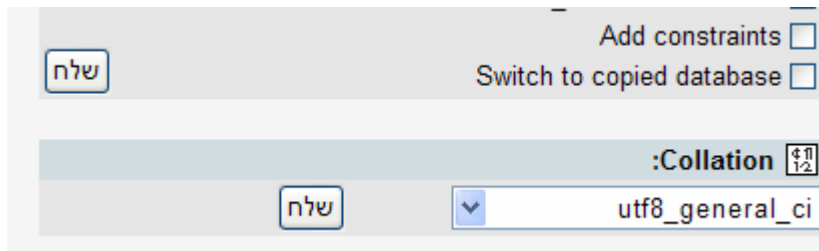
ראשית, ניתן למחוק את כל הטבלאות מבסיס הנתונים. בסרגל הניווט העליון יש לעבור ל"מבנים"



במסך שיופיעו, יוצגו כל הטבלאות של בסיס הנתונים. בתחתית רשימת הטבלאות יופיע הכיתוב "סמן הכל". לחיצה עליו תסמן את כל הטבלאות. כעת, למחיקת הטבלאות, יש לבחור מתוך תיבת הבחירה שמתחת, את האפשרות "הסר". **שים לב: פעולה זו תמחק את כל טבלאות ואת נתוני האתר! וודא שביצעת גיבוי לבסיס הנתונים ובצע פעולה זו על אחריותך בלבד!**



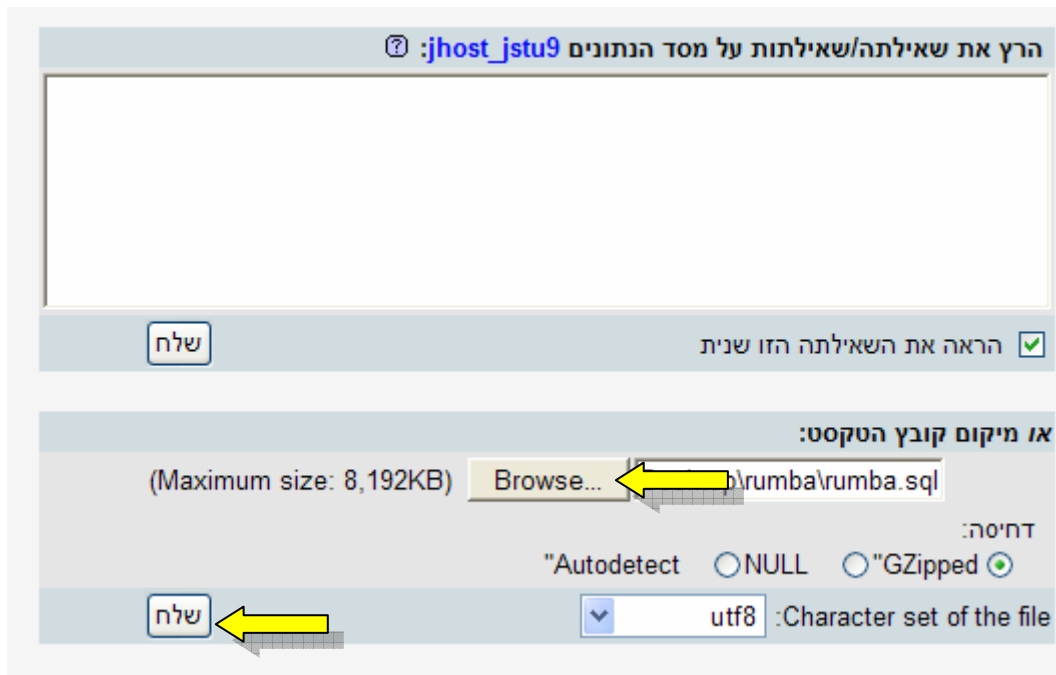
לאחר שנמחקו כל הטבלאות, ניתן להמיר את קידוד בסיס הנתונים ל-UTF8, כפי שהוסבר קודם לכן, עבור בסיס הנתונים חדש. יש לעבור ל"פעולות", ולבחור במסך זה, בקידוד: utf8_general_ci



כעת, בסיס הנתונים הותאם לקידוד הרצוי, וניתן להעלות אליו את קובץ ה-SQL שנשמר על המחשב והותאם לקידוד. יש לעבור בסרגל הניווט העליון למסך "SQL".



במסך זה, יש לבחור את קובץ ה-SQL שנשמר על המחשב ועבר הסבה וללחוץ על "שלח" כדי להעלות אותו אל בסיס הנתונים.



לאחר שהסתיים ייבוא הקובץ, וכל הפעולות הנוספות בוצעו – האתר הותאם ל-UTF8!