

## טכנולוגיית המידע – הסביבה העברית עבור מְנַשֵּׁק מערכות הפעלה יבילות

Information technology: Hebrew local for portable operating systems interface (POSIX)

תקן זה הוכן ואושר על ידי הוועדה הטכנית 2109 - עברית במערכות ממוחשבות ואינטרנט, בהרכב זה:

- |   |  |
|---|--|
| איגוד האינטרנט הישראלי                                  | - מתתיהו אלוש (יו"ר), אורן בר, שחר שמש, דורון שקמוני |
| איל"א - איגוד ישראלי לטכנולוגיית המידע                  | - ישראל גידלי  |
| אמצ"י - ארגון הגג למשתתפים ושותפים עסקיים של יבמ בישראל | - מנחם דישון   |
| המועצה הישראלית לצרכנות                                 | - מירי ברור לוי                                      |
| המקור - עמותה ישראלית לקוד מקור פתוח ותוכנה חופשית      | - עמית אהרונוביץ'                                    |
| לשכת מנתחי מערכות המידע בישראל                          | - זיו תובל, דן תל ניר                                |
| מיקרוסופט ישראל   | - שמואל יאיר   |
| רשות ההסתדרות לצרכנות                                   | - עמית שניזק   |

כמו כן תרמו להכנת התקן: שי ברגר, אפרים ברקוביץ', אהרון לנין, תומר מכלין, לינה קימל, יובל רבינוביץ'

עמית אסא ריכזה את עבודת הכנת התקן.

---

הודעה על רויזיה  
תקן ישראלי זה בא במקום  
התקן הישראלי ת"י 1748 מאוגוסט 2007

---

**מילות מפתח:**  
חליפת מידע, תמסורת נתונים, תווים עבריים, עברית, יביל.

**Descriptors:**

information exchange, data transmission, hebraic characters, hebrew, portable.

---

**עדכניות התקן**

התקנים הישראליים עומדים לבדיקה מזמן לזמן, ולפחות אחת לחמש שנים, כדי להתאימם להתפתחות המדע והטכנולוגיה. המשתמשים בתקנים יודאו שבידיהם המהדורה המעודכנת של התקן על גיליונות התיקון שלו. מסמך המתפרסם ברשומות כגיליון תיקון, יכול להיות גיליון תיקון נפרד או תיקון המשולב בתקן.

---

**תוקף התקן**

תקן ישראלי על עדכוני נכנס לתוקף החל ממועד פרסומו ברשומות. יש לבדוק אם המסמך רשמי או אם חלקים ממנו רשמיים. תקן רשמי או גיליון תיקון רשמי (במלואם או בחלקם) נכנסים לתוקף 60 יום מפרסום ההודעה ברשומות, אלא אם בהודעה נקבע מועד מאוחר יותר לכניסה לתוקף.

---

**סימון בתו תקן**

כל המייצר מוצר, המתאים לדרישות התקנים הישראליים החלים עליו, רשאי, לפי היתר ממכון התקנים הישראלי, לסמנו בתו תקן:



---

**זכויות יוצרים**

© אין לצלם, להעתיק או לפרסם, בכל אמצעי שהוא, תקן זה או קטעים ממנו, ללא רשות מראש ובכתב ממכון התקנים הישראלי.

## תוכן העניינים

1	1. תחום התקן.....	1
1	2. אזכורים.....	1
1	3. מונחים והגדרות.....	1
2	4. הסביבה העברית.....	2
2	4.1. מערכת תווים.....	2
3	4.2. סדר המיון.....	3
4	4.3. ערכים כספיים.....	4
4	4.4. מספרים.....	4
4	4.5. זמנים.....	4
4	4.6. הודעות.....	4
4	5. הגדרות הסביבה העברית.....	4

## 1. תחום התקן

תקן זה קובע הגדרות המבוססות על המוסכמות התרבותיות המקומיות עבור מערכת התווים המקובלת בסביבת עבודה עברית, עבור סיווג התווים ומיונם, ועבור הצגת מספרים, ערכים כספיים, תאריך, זמן והודעות מערכת.

תקן זה מתאים לדרישות המקובלות בטכנולוגיית המידע, כנקוב בתקן הבין-לאומי למערכות הפעלה יבילות (POSIX) ISO/IEC 9945-2.

התקן מאפשר פיתוח יישומים במערכות שונות המממשות את מנגנון הסביבה העברית (הגדרה 3.1) באופן אחיד וללא צורך בהתאמה מיוחדת של הפרמטרים העבריים.

התקן מתייחס למידע המאוחסן בסדר לוגי ולשימוש בכיווניות משתמעת (הגדרה 3.3).

התקן מבוסס על ההנחה שהכיווניות הכללית היא מימין-לשמאל.

כיווניות התווים מוגדרת בתקן הישראלי ת"י 1489.

## 2. אזכורים

תקנים ומסמכים המאזכרים בתקן זה (תקנים ומסמכים לא מתוארכים - מהדורתם האחרונה היא הקובעת):

### תקנים ישראליים

ת"י 1489 - טכנולוגיית מידע: ארכיטקטורה לאחסון ולהצגה של טקסט בשפה העברית במערכות תקשורת ומחשבים (מערכות תקשוב)

ת"י 3166 - קודים לייצוג שמות של ארצות

### תקנים בין-לאומיים

ISO 639	-	Code for the representation of names of languages
ISO/IEC 4217	-	Codes for the representation of currencies and funds
ISO/IEC 9945-2	-	Information technology - Portable operating system interface (POSIX): Shell and utilities
ISO/IEC 10646-1	-	Information technology - Universal multiple-octet coded character set (UCS): Architecture and basic multilingual plane

## 3. מונחים והגדרות

מונחים והגדרות אלה כוחם יפה בתקן זה:

### 3.1 סביבה עברית (Hebrew locale) (להלן: locale)

מערכת הגדרות, כנקוב בתקן הבין-לאומי למערכות הפעלה יבילות (POSIX) ISO/IEC 9945-2 - 1993, המבוססת על המוסכמות התרבותיות המקומיות סדר לוגי הסדר שבו הטקסט נקרא או נכתב.

### 3.2 כיווניות משתמעת

כמוגדר בתקן הישראלי ת"י 1489 מיולי 2012:

שיטה לקליטה ולהצגה של טקסט שבה הכיוון נקבע על פי אלגוריתם, המתייחס לסוגי התווים ולמיקומם ביחס לתווים הסמוכים להם וביחס לכיוון הראשי. כיווניות משתמעת כרוכה בדרך כלל באחסון טקסט לפי סדר לוגי.

- התווים מסווגים בהתאם לסוג האלגוריתם המופעל, כמפורט להלן:
- תווים "בעלי כיוון" - תווים שעצם מהותם מגדירה כיווניות, כגון:
    - תווים עבריים - מימין-לשמאל;
    - תווים לטיניים - משמאל-לימין;
    - ספרות - משמאל-לימין.
  - תווים "חסרי כיוון", כגון רווח או מקף, שאין להם כיווניות עצמית, אלא כיווניותם נרכשת מהסביבה (תווים קודמים או תווים עוקבים) לפי אלגוריתם פיענוח הכיווניות.

#### 4. הסביבה העברית

בהתאם למקובל, שם ה-locale יהיה he\_IL, כאשר he מסמן את השפה העברית על פי התקן הבין-לאומי ISO 639 ו-IL מציין את מדינת ישראל על פי התקן הישראלי ת"י 3166. הערה: בעבר, שם ה-locale היה iw\_IL. שם זה אינו מקובל כיום, והוא בגדר מונח שיצא מכלל שימוש (deprecated).

#### 4.1 מערכת תווים

בתקן זה מובאת מערכת התווים המקובלת בסביבת עבודה עברית, על פי התקן הבין-לאומי ISO/IEC 10646-1. לצורך התאמת התקן הבין-לאומי למערכת תווים מסוימת, יש להתייחס אך ורק לתווים הכלולים במערכת. להלן מפורטות קבוצות התווים המוגדרות עבור ה-locale. שמות הקבוצות, כפי שהם מובאים בקובץ המקור להגדרות הסביבה העברית (ראו סעיף 5), מצוינים בסוגריים. מערכת התווים המקובלת לסביבה פתוחה תכלול את התווים האלה: האותיות העבריות והלטיניות, הספרות וסימני הפיסוק.

##### 4.1.1 אותיות (alpha)

קבוצה זו כוללת את האותיות העבריות והלטיניות.

##### 4.1.2 תווי רווח (space)

קבוצה זו כוללת את התווים המשמשים להפרדה בין מילים. בקבוצה זו נכללים גם תווי מרווח (ראו סעיף 4.1.9).

##### 4.1.3 תווי בקרה (ctrl)

קבוצה זו כוללת תווים המשמשים לבקרה, שבדרך כלל אין להם ייצוג גרפי.

##### 4.1.4 סימני פיסוק (punct)

קבוצה זו כוללת סימני פיסוק וסימנים מיוחדים.

##### 4.1.5 תווים גרפיים (graph)

קבוצה זו כוללת את כל התווים שיש להם ייצוג גרפי.

##### 4.1.6 תווים מודפסים (print)

קבוצה זו כוללת את כל התווים המופיעים בהדפסה, לרבות רווחים.

##### 4.1.7 ספרות (digit)

קבוצה זו כוללת את הספרות 0 עד 9.

- 4.1.8 ספרות הקסאדצימליות (xdigit)** . קבוצה זו כוללת את הספרות הלטיניות ההקסאדצימליות.
- 4.1.9 תווי מרווח (blank)** . קבוצה זו כוללת את התווים רווח וטבלר<sup>(1)</sup> (tab).
- 4.1.10 המרה לאותיות רישיות (toupper)** . קבוצה זו מתייחסת לאותיות לטיניות בלבד, וכוללת טבלת המרה של אותיות קטנות לאותיות רישיות.
- 4.1.11 אותיות עבריות (Hebrew)** . קבוצה זו כוללת את האותיות העבריות.
- 4.1.12 סימני ניקוד (point)** . קבוצה זו כוללת את סימני הניקוד העבריים.
- 4.1.13 טעמי המקרא (accent)** . קבוצה זו כוללת את טעמי המקרא.

## 4.2 סדר המיון

- באופן עקרוני, סדר המיון יהיה כמפורט להלן: תווי בקרה מסוימים, סימני פיסוק, ספרות, אותיות עבריות, אותיות לטיניות.
- אין במיון התייחסות לטעמי המקרא ולתווי עיצוב הכיווניות. להגדרה מדויקת של סדר המיון, ראו סעיף 5 (בקובץ המקור להגדרות).
- סדר המיון של טקסט עברי יתבצע כמפורט בשלבים להלן:
1. פירוק צירופי האותיות עבור יידיש (בתחום יוניקוד 05F0 עד 05F2) ועבור עברית (בתחום יוניקוד FB1D עד FB4F).
  2. נרמול: כאשר מופיעה אות בסיס עם תוספות, המיון מבוצע כך שאות הבסיס מופיעה ראשונה, אחריה נקודת שי"ן ימנית או שמאלית, אחריה דגש, ואחריו ניקוד.
  3. מיון האותיות ללא ניקוד (בתחום יוניקוד 05D0-05EA): אותיות מנצפ"ך סופיות (סןףך) מופיעות לפני אותיות מנצפ"כ רגילות (לדוגמה, חך לפני חכ).
  4. מיון השי"נים: ש ללא ניקוד (יוניקוד 05E9), ש ימנית (בצירוף יוניקוד 05C1), ש שמאלית (בצירוף יוניקוד 05C2).
  5. מיון לפי סימני הניקוד (בתחום יוניקוד 05B0 עד 05BC), לרבות דגש.
  6. גרש וגרשיים: גרש וגרשיים משפיעים על המיון (זהו המצב הקיים).
  7. מקף, רווח ומינוס: בשלב המיון המתחשב בסימני הניקוד, מתמיינים כערכים שווים וקטנים מכל האותיות. במיון של סימני הניקוד יש ביניהם הבדל: רווח, מינוס, מקף בסדר עולה.

(1) מוכר גם כליווח אופקי או עזר.

#### 4.3 ערכים כספיים

סמל המטבע הוא ₪ או ILS (על פי התקן הבין-לאומי ISO/IEC 4217).

#### 4.4 מספרים

- הנקודה העשרונית מיוצגת על ידי נקודה.
- במספרים ארוכים, בין כל שלוש ספרות מפריד פסיק.

#### 4.5 זמנים

##### ימי השבוע

השמות המקוצרים של ימי השבוע (5 תווים לכל היותר): יום א, יום ב, יום ג, יום ד, יום ה, יום ו, שבת.

השמות המלאים של ימי השבוע: יום ראשון, יום שני, יום שלישי, יום רביעי, יום חמישי, יום שישי, שבת.

##### חודשים

השמות העבריים המקוצרים לחודשים בלוח השנה הכללי (3 תווים): ינו, פבר, מרס, אפר, מאי, יוני, יול, אוג, ספט, אוק, נוב, דצמ.

השמות העבריים המלאים לחודשים בלוח השנה הכללי: ינואר, פברואר, מרס, אפריל, מאי, יוני, יולי, אוגוסט, ספטמבר, אוקטובר, נובמבר, דצמבר.

##### שעות

ציון השעה הוא לפי מערכת של 24 שעות, ולא לפי ציון am/pm.

#### 4.6 הודעות

הודעת "כן" מוגדרת בצורות אלה: כן, כ, YES, Y, y או yes.

הודעת "לא" מוגדרת בצורות אלה: לא, ל, NO, N, n או no.

#### 5 הגדרות הסביבה העברית

להלן קובץ המקור להגדרות הסביבה העברית:

[https://portal.sii.org.il/UploadedFiles/12\\_2013/he\\_IL.txt](https://portal.sii.org.il/UploadedFiles/12_2013/he_IL.txt)

##### הערה:

בעת פרסום תקן זה הכתובת שלעיל היא הכתובת שבתוקף.