

בניית אתרים וסמנטיקה (Semantic HTML)

מאת ניר אדר

פתיחה

Semantic HTML פירושו להשתמש ב-HTML כדי להגדיר משמעות (סמנטיקה) לחלקים בטקסט שלנו, בניגוד להגדרה בה משתמשים מתכנתים רבים, המגדירה רק את עיצוב הטקסט שלנו. נציג דוגמה כדי להעביר את הכוונה. נניח שאנחנו רוצים לעשות כותרת לדף הבית שלנו. דרך אחת יכולה להיות:

```
<font size="14px" color="#000"><b>Welcome to the site</b></font>
```

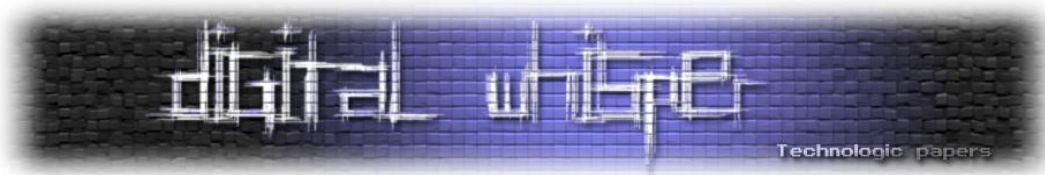
במקרה הזה קשה למי שקורא את הקוד שלנו להבין שמדובר בכותרת. לעומת זאת, אנחנו יכולים לכתוב:

```
<h1>Welcome to the site</h1>
```

במקרה הזה ברור לחלוטין שמדובר בכותרת של האתר. לתג יש משמעות של כותרת. יתכן שמבחינת התצוגה על המסך שני קטעי הקוד יתנו תוצאה זהה, אבל קטע הקוד השני נותן לקורא מידע נוסף על הטקסט שהוא מעצב - הוא אומר לו "זו הכותרת של הדף".

למה חשוב לנו ליצור קוד עם משמעות?

1. מנועי חיפוש יבינו יותר טוב את הדף שלנו ואת התכנים שיש בו.
2. קוד הדף יהיה ברור יותר עבורנו, ועבור מתכנתים נוספים שיצפו בקוד הדף. המשמעות של חלקי הקוד השונים והחלקים השונים בדף תהיה מהירה יותר לקורא.
3. הפרדת העיצוב מהתוכן!
4. קוד נוח לתחזוקה ולשינוי:
 - a. אין טבלאות לצורך עיצוב (כי העיצוב נעשה ב-CSS) ומכאן - אין עומס בדף עצמו - הרבה יותר קל לקרוא את הקוד.
 - b. הקוד נקי יותר וקצר יותר, ומכאן - ברור יותר להבנה.
 - c. הרבה יותר קל לבצע שינויים בעיצוב האתר.



5. אין לכם ברירה - תגים המיועדים לעיצוב בלבד, ללא משמעות, צפויים להעלם בגרסאות הבאות של HTML.

6. סיבות רבות נוספות.

איך נושא זה מתקשר לגליון שלנו? המאמר המצויין של אפיק דיבר על Meta Data. שימוש ב-Semantic HTML הוא בעצם הוספת Meta Data לדפי האינטרנט שאנחנו כותבים. אחרי המאמר שהציג את הסכנות ב-Meta Data, במאמר זה אני רוצה להראות לכם כמה מהיתרונות, ובדרך להציג את הנושא החשוב של Semantic HTML, שלטעמי חייב להיות הרבה יותר מוכר ונפוץ ממה שהוא היום.

המטרה של שפת ה-HTML לא היתה ואינה גם כיום לעצב את הדפים שלנו. המטרה של HTML היתה להוסיף לטקסט מידע שיגדיר את חלוקת הטקסט, ושיתן לו משמעות (סמנטיקה) נוספת. האתרים הראשונים בשפת HTML היו אתרים נכונים מאוד מבחינה סמנטית, אבל מכוערים - מתכנתים כתבו אותם, והשתמשו בתגיות השונות עם עיצוב ברירת המחדל. תגי העיצוב הוכנסו לשפת HTML במהלך שנות התשעים של המאה ה-20. הם גרמו לרווח מצד אחד מבחינת מראה האתרים, אך גרמו מהצד השני לקוד שערבב עיצוב ותוכן, והפך פחות קריא ופחות ברור.

CSS בא כדי לעזור למתכנתים להפריד את העיצוב מהתוכן. עם זאת, באתרים רבים ברשת עד היום הנושא לא מבוצע כראוי.

מעט רקע בנושא HTML

אני רוצה להתחיל את הדיון על סמנטיקה במעט רקע על שפת HTML, ועל המרכיב הבסיסי בה - HTML Element. באופן בסיסי, מסמכי HTML מורכבים מתחילתם ועד סופם מ-"רכיבי HTML". מבנה של רכיב HTML טיפוסי:

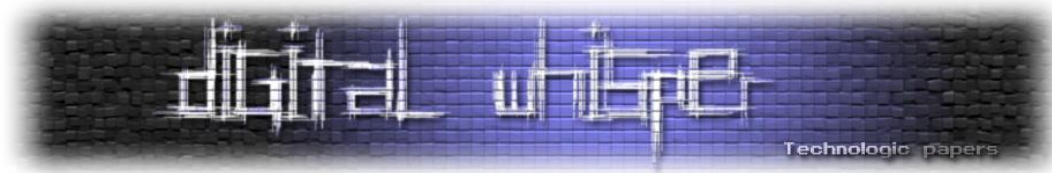
```
<tag>content to render</tag>
```

מבנה כללי יותר:

```
<tag attribute1="value1" attribute2="value2">content to render</tag>
```

אני מניח שכולכם מכירים את השימוש ב-HTML ושאינן צורך להסביר מהם רכיבי HTML מעבר לכך. מה שאולי אינכם יודעים היא העובדה שהתגים עצמם מחולקים ל-3 קבוצות:

1. Block Elements - תופסים את כל הרוחב האפשרי בדף. לפנייהם ואחריהם מתחילה שורה חדשה. (לדוגמא: <div>, , <h1>, <p>)



- 2. Inline Elements - לוקחים בדיוק את המקום שהם צריכים, ולא מתחילים שורה חדשה. (לדוגמא: `<a>`, ``, ``)
- 3. בלתי נראים - לא מופיעים בדף עצמו. (לדוגמא: `<meta>`)

מבחינת משמעות התגים השונים, מוגדר כי:

- 1. Block Elements יכולים להכיל Block Elements אחרים, או Inline Elements.
 - 2. Inline Elements יכולים להכיל בתוכם Inline Elements אחרים וטקסט בלבד.
- תגי ה-Block הם הנותנים את "התמונה הגדולה" של הדף. הם מחלקים את המסמך לחלקים שונים. תגי ה-Inline הם המרכיבים של החלקים השונים.

בתיאוריה זה נשמע ברור והגיוני. איפה הרבה אנשים טועים? יתכן שבעבר כתבתם קוד כזה:

```
<b><h1> Welcome to my site </h1></b>
```

יש כאן כותרת, ואנחנו רוצים גם שהיא תהיה מודגשת. זו טעות! כותרת היא תג Block ואילו `` הוא תג inline. למרות שרוב הדפדפנים היום יציגו את הטקסט בכותרת מודגש, כמו שביקשנו, מבחינת המבנה הנכון של מסמך HTML זו פשוט טעות. הצורה היותר הנכונה היא:

```
<h1><b> Welcome to my site </b></h1>
```

אני אדגיש "היותר נכונה" מכיוון שהשימוש בתג `` בעייתי באופן כללי - זהו תג של עיצוב ולא של משמעות. התג הנכון יותר לשימוש, במידה ואנחנו רוצים להבליט את הכותרת, הוא ``.

התמונה הגדולה

דיברנו על סמנטיקה, הצגנו מעט טעימות מה זה לבנות אתר "בצורה נכונה סמנטית", עכשיו בואו נדבר על התמונה הגדולה. כשאנחנו מדברים על בניה נכונה מבחינה סמנטית, על מה אנחנו מדברים?

מבנה מסמך הגיוני, זרימת מסמך הגיונית:

- מבנה קוד נכון - חלוקה גדולה לתגיות בלוקים. בתוכם חלוקה לבלוקים משניים.
- כחלק ממבנה הקוד - תגיות Hx במבנה היררכי נכון. תגיות הכותרת מסבירות את המבנה הלוגי של העמוד ואפשר להבין את נושאי העמוד מקריאת הכותרות בלבד.
- כשמסתכלים על קוד המסמך, ללא עיצוב - הסדר של הדברים במסמך צריך להיות הגיוני. למשל: ה-footer לא צריך להיות החלק הראשון מבחינת הקוד, המופיע בדף שלנו.

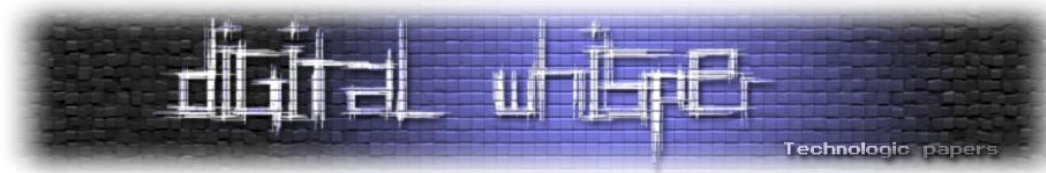
תיג מושכל:

- שימוש בתגיות שיסבירו את התוכן. הדוגמא הכי בסיסית: תגית הכותרת, שהדגמנו בפרק הפתיחה. דוגמא נוספת לתגית בעלת משמעות היא תגית הפסקה <p>. שימוש בתגית <p> במקום ב-
 כדי להפריד בין פסקאות מבהירה לקורא שמדובר בפסקה חדשה בדף שלנו, לעומת
 שאין לו משמעות סמנטית. יש להשתמש בתגיות המסבירות את התוכן בכל מקום שרק ניתן להשתמש בו בהן.
- לא משתמשים בתגיות במשמעות שונה מהמשמעות הסמנטית שלהם! טעות זו חוזרת במקרים רבים כשמתכנתים משתמשים בתגית עם משמעות סמנטית כדי לקבל אפקט עיצובי כלשהו.

הפרדת עיצוב מתוכן:

- במצב הטוב ביותר - אין שום סימן של עיצוב כשמסירים את ה-CSS - הדף נראה כמו דף של שנות ה-90 ללא ה-CSS.

לעתים קיים בלבול בין קוד תקני לקוד נכון סמנטית. שימו לב כי קוד תקני הוא קוד ללא שגיאות syntax, אבל הוא לא מציין בהכרח שמוקמו בו התגיות הסמנטיות הנכונות. אתר המעוצב באמצעות טבלאות יכול להכיל קוד תקני, אבל הוא אינו נכון סמנטית.



מבנה מסמך נכון סמנטית

שלבים ליצירת מבנה דף נכון סמנטית:

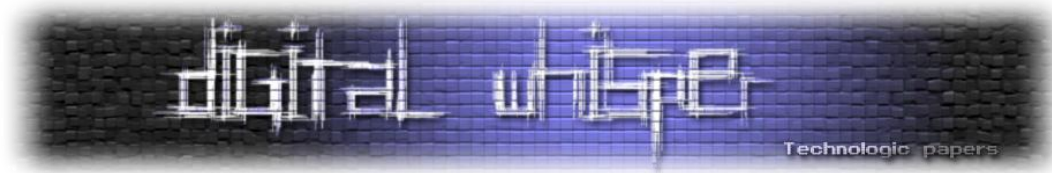
1. המסמך שלנו צריך להתחיל בתגית DOCTYPE המסבירה לדפדפנים איך רצוי להתייחס למסמך שלנו. ביחרו את התג המתאים.
2. מבנה דף ה-html הבסיסי:

```
<html>
<head>
  <title> </title>
</head>
<body>
  Content here
</body>
</html>
```

חשוב לשמור על מבנה זה - כמה שהוא בסיסי, אפילו הוא מפוספס באתרים רבים.

3. ה-body צריך להיות מחולק ל-div שונים, המציינים את חלקיו השונים של הדף שלכם. ה-div השונים צריכים להיות בסדר הגיוני. ה-div בו נמצא ה-footer של הדף, למשל, צריך להופיע בסוף הקוד.
4. מבנה הכותרות וסדר הכותרות צריך להיות הגיוני. למשל, אסור שתהיה קפיצה מ-h1 ל-h4. דוגמה למבנה:

```
h1 -Page heading
h2 -Section heading
h3 -Section subheading level 1
h3 -Section subheading level 1
h2 -Section heading
h3 -Section subheading level 1
h3 -Section subheading level 1
h4 -Section subheading level 2
h5 -Section subheading level 3
h3 -Section subheading level 1
h4 -Subsection subheading level 2
h2 -Section heading
```



משמעות סמנטית של תגיות HTML

תגיות בלתי נראות

<title>

- תגית ה-title לא מופיעה בדף עצמו, אך היא אחת התגיות החשובות ביותר הקיימות.
- תקן W3 מציין לגבי ה-title שעליה לציין את תוכן הדף. בנוסף הוא מציין שיש לתת תיאור משמעותי ולא תיאור קצר מכיוון שהרבה פעמים הגולש מגיע לדף ישירות ולא יודע את ההקשר. הכותרת "Introduction", למשל, היא כותרת גרועה לדף. כותרת מוצלחת יותר תהיה "Introduction to HTML" - כעת ברור מה יש בדף שלנו.
- שימוש ב-title טוב מועיל הרבה כאשר מדובר במנועי חיפוש: רוב מנועי החיפוש ישתמשו ב-title כדי להציג את הדף שלנו בתוצאות החיפוש. בנוסף מנועי החיפוש משתמשים בתג זה רבות כדי להחליט עבור אילו מילים הדף שלנו יופיע גבוה בתוצאות החיפוש.

תגיות Block

<div>

- divider - חלוקת הדף לחלקים. התוכן ב-div מייצג חלק בדיד משאר חלקי הדף. לרוב ישמש כדי לציין חלוקת הדף ל-header, footer, איזור התפריט ואיזור התוכן. נעשה בו גם שימוש לחלק איזור לאיזור משנה - למשל - חלוקת כתבה לאיזור הטקסט המרכזי ולאיזור התגובות שלה.
- אם נכתוב את הדף עם המון div, כאילו עיצבנו את הדף בעזרת טבלאות. חשוב להשתמש בתג זה לפי המשמעות הסמנטיות שלו.

h1-h6

- כל הכותרות באות לייצג חלוקה של התוכן לחלקים שונים (המופרדים על ידי <p> ו-<div>). הכותרות מגדירות למעשה את מבנה הדף.
- h1 היא הכותרת הראשית של הדף. היא מופיעה רק פעם אחת. כותרות אחרות יכולות להופיע מספר פעמים, בהתאם למבנה ההיררכי הנכון שהוצג בפרק הקודם.

<p>

- אלמנט הפסקה הוא האלמנט הבסיסי המחלק טקסט. מבחינה סמנטית, היא מייצגת קטע בדיד של טקסט.
- בניגוד ל- <div> המגדיר חלוקה כללית בדף, <p> משמשת לחלוקת טקסט לפסקאות.

<blockquote>

- תג הבא לציין שקטע הטקסט שבתוכו מצוטט ממקור אחר.
- תג זה תמיד צריך לבוא עם תג cite המציין את המקור ממנו הציטוט מגיע, ולעיתים גם מכיל מאפיין cite המציין URI של הטקסט המצוטט.

תגיית Inline

<a>

- תג זה מייצג קישור למשאב חיצוני, או לחילופין קישור לקטע פנימי בדף הנוכחי.
- חשוב לשים לב שבעזרת קוד סקריפט אפשר להקנות פונקציונליות מגוונת לתג <a>. מבחינה סמנטית חשוב לנו לשמור את המשמעות שלו כקישור.

- תג זה מייצג דגש קל. מטרתו הסמנטית היא להדגיש שהטקסט שהוא מקיף חשוב מעט מהטקסט הסובב אותו.
- הבחירה האם טקסט כלשהו דורש דגש קל היא בידי כותב הטקסט.
- כברירת מחדל, דגש קל מופיע בטקסט כטקסט מוטה, אך בעזרת CSS אפשר לשנות זאת לכל צורה אחרת.

- תג זה מייצג דגש חזק. מטרתו הסמנטית היא להדגיש שהטקסט שהוא מקיף חשוב מהטקסט הסובב אותו. זה בעצם גרסה חזקה יותר של
- הבחירה האם טקסט כלשהו דורש דגש קל היא בידי כותב הטקסט.

- כברירת מחדל, דגש חזק מופיע בטקסט כטקסט מודגש, אך בעזרת CSS אפשר לשנות זאת לכל צורה אחרת.

<address>

- תג המציין קטע בו מופיעים פרטי יצירת הקשר של כותב הדף.
- כברירת מחדל הטקסט בו מופיע מוטה.
- מקובל שתג זה מופיע ב-header או ב-footer של העמוד.
- תג זה מכיל לרוב את הכתובת הפיזית של העסק/כותב הדפים.

<code>

- מציין קטע קוד (של שפת תכנות) המוצג בתוך המסמך.

<samp>

- מציין פלט של תוכנית מחשב.
- נניח התוכנית שלכם מדפיסה "hello world", אז קוד התוכנית יוקף בתג <code> והפלט שלה, "hello world", אם תציגו אותו בדף, יוקף בתג <samp>.

<cite>

- מקור ציטוט - משמש כדי לציין מהו המקור של המידע.

<q>

- טקסט מצוטט קצר. לדוגמא:

```
<p>As <cite>Bill Gates</cite> said, <q>640K should be enough for anybody.</q></p>
```

- הטקסט המצוטט יופיע באופן אוטומטי בין מרכאות עוטפות.
- חשוב לציין ש-Internet Explorer מפגין יישום גרוע של תג זה, ולפיכך אתרים רבים לא משתמשים בו.

- ל-span יש תפקיד סמנטי - מפריד פנימי כללי של בלוק. תג זה הוא מפריד כללי - אין לו משמעות ספציפית (כמו תגים אחרים שראינו בפרק זה).
- נשתמש בו כאשר אין תג ספציפי בעל משמעות מתאימה יותר.

, ,

- תגים אלו מייצגים רשימות של פריטים. מייצג רשימה חסרת סדר ו- רשימה בעלת סדר. הוא פריט ברשימה.
- תפריטים, לדוגמא, הם דוגמא קלאסית של נתונים המתאימים להצגה כרשימה - תפריטים הם הרי רשימה של קישורים. הערות בבלוג מתאימות להיות רשימה בעלת סדר (לרוב הן ממוספרות בסדר לפי זמן שליחתן).

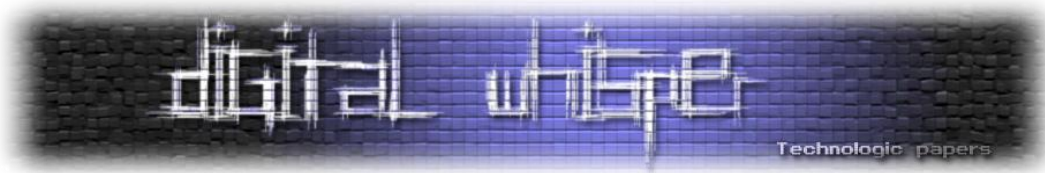
<dl>, <dd>, <dt>

- תגים אלו מייצגים "רשימת הגדרות" - ובאופן פרקטי - זוגות של מונחים (<dd>) ושל ההגדרות שלהם (<dt>).
- תג זה יחסית פחות מוכר מאשר הרשימה הרגילה וחשוב להכירו.
- שימוש בתג זה נכון מבחינה סמנטית בכל פעם שיש לך רשימה המורכבת מזוגות בעלי אופי של שם+ערך. ברוב שאר המקרים, הוא הפתרון הנכון.

תגיות בהן לא משתמשים

התגיות הבאות הוכרזו כמיושנות (Deprecated) ואין להשתמש בהן כלל: dir, font, center, applet, u, strike, s, menu, isindex.

תגיות בהן כמעט אף פעם לא נשתמש: big, small, b, i, tt, pre. תגיות אלו הינן חלק מהשפה, אבל אין להן משמעות סמנטית. לפיכך, נשתמש בהם רק במקרים נדירים בהם אחרי מחשבה נחליט שהשימוש הטוב ביותר הוא בהן. למשל הוא הדגשה, אך ללא מתן חשיבות לטקסט המודגש.



חשוב לדעת שיש עוד תגיות רבות ב-HTML עם משמעות סמנטית. אמליץ לכל הקוראים להעמיק ברשת בנושא זה כדי לקבל בו שליטה מלאה.

הפרדת העיצוב מהתוכן

הכלי המרכזי שיש לנו כדי לבנות מסמך סמנטי ולהפריד את העיצוב מהתוכן הוא CSS Positioning. הוא אינו קשה מאוד ללמידה ומומלץ להבנה לעומק.

העיצוב של האתר יבוצע תמיד על ידי קובץ CSS חיצוני. כאשר המסמך מחולק נכון - לא נדרשים אלמנטי עיצוב בתוך המסמך.

סיכום

המשמעות של התוכן מובנת על ידי הקורא של התוכן. חשוב לשים לב שהקורא הוא לא תמיד אנושי. רובוטים ותוכנות שונות יקראו את התוכן שלכם פעמים רבות.

כתיבה נכונה סמנטית נכונה גם במקומות נוספים, שאינם רק HTML. גם ב-Word, למשל, ניתן לקבוע כי שורה מסויימת היא מודגשת וגדולה, או לחילופין ניתן לסמן אותה ככותרת במסמך. אם תסמנו את השורות ככותרות, Word יוכל, למשל, לייצר לכם תוכן עניינים למסמך באופן אוטומטי. עניין הכתיבה הסמנטית צריך להיות בראש בכל פעם שכותבים מסמך, ובכל פלטפורמה.

ברצוני להמליץ גם על הבלוג המצויין של מר יובל רז, www.yuvalraz.com, חבר ואחד האנשים המבריקים בתחום ה-UX בארץ. במיוחד אפנה אתכם להרצאת וידאו המדברת על הנושאים במסמך זה, ומציגה דגשים שבחלקם שונים מאלו שהצגתי לכם היום. [הרצאת הוידאו במפגש איגוד האינטרנט.](#)

חשוב לי להדגיש כי לא לכל דבר יש ל-HTML תג יעודי, אך עדיין קיימים ב-HTML הרבה תגים ייעודיים בהם אתרים לרוב לא מבצעים שימוש. כשאתם חושבים על קוד - נסו לתת לו כל העת משמעות רבה ככל שניתן. כשאתם כותבים קוד HTML חישבו על מה התוכן ומה המשמעות שלו, ולא על איך אתם רוצים שהוא יעוצב, כאשר אתם בוחרים את התגים.