

חיפוש GREP באינדיזיין CS3

סקירה כללית

- חיפוש GREP הוא חיפוש באמצעות קודים. באמצעותו אפשר לחפש טקסט בהתניות שונות, להחליף אותו, ו/או להחל עליו סגנונות וכד'.
 - חיפוש GREP רגיש לאותיות גדולות/קטנות.
 - אפשר לשמור שאילתות GREP לשימוש חוזר. לחץ על סמל הדיסקט שבראש התיבה.
- ❖ חשוב מאוד: את הקודים במדריך זה יש להקליד לפי סדרם מימין לשמאל. מילים באנגלית יש להקליד כרגיל.

1. קבוצות תווים

- **()** סוגריים עגולים מגדירים את מה שביניהם כקבוצה (=תת-ביטוי). אפשר להחליף את מיקום הקבוצות בשדה 'שנה ל:'. מספרי הקבוצות ב'שנה ל:': מוגדרים בסימן דולר ומספר. דוגמה: חפש (בוקר) (טוב) החלף \$2\$3\$. **1**. בוקר טוב ישתנה ל-טוב בוקר. החלף \$3\$ ישנה את בוקר טוב ל-טוב.
- **\$** ב'שנה ל:': מסמל את כל הטקסט שב'חפש'.
- **(?:קבוצה)** סימן שאלה ונקודתיים לאחר הסוגר הימני, מגדירים שקבוצה זו לא תחשב כקבוצה בשדה 'שנה ל:'. דוגמה: חפש (ראובן) (שמעון) (לוי) החל סגנון \$2\$1\$ יחיל סגנון רק על ראובן ו-לוי.

2. תווים כלליים

- נקודה מגדירה כל תו שהוא: אות, מספר, סימני פיסוק וכו'. דוגמה: **ים** מחפש **יום**, **יזם**, או **י-ם**. (בשדה 'שנה ל:', נקודה היא תו רגיל).
- כאמור, GREP רגיש לאותיות גדולות/קטנות. בדרך כלל אות קטנה משמשת לחיפוש חיובי, ואות גדולה משמשת לחיפוש שלילי. קודים אלו פועלים רק בשדה 'חפש' ולא בשדה 'שנה ל:'.
 - **w** מחפש כל תו שהוא חלק ממילה. **W** מחפש כל תו שאינו חלק ממילה.
 - **d** מחפש כל תו ספרה. **D** מחפש כל תו שאינו ספרה.
 - **u** מחפש כל אות לועזית רישית (=גדולה). **U** מחפש כל תו שאינו אות רישית.
 - **l** מחפש כל אות לועזית קטנה. **L** מחפש כל תו שאינו אות קטנה.
 - **s** מחפש כל תו לבן (שאינו מופיע בדפוס: רווח, טאב, שבירת שורה, סוף פיסקה וכו'). **S** מחפש כל תו שאינו תו לבן.
- **תווים כלליים נוספים (Posix)**. גם קודים אלו פועלים רק בשדה 'חפש'.
 - **[[:alnum:]]** כל תו אות או מספר.
 - **[[:alpha:]]** כל תו אות.
 - **[[:upper:]]** כל תו אות לועזית רישית.
 - **[[:lower:]]** כל תו אות לועזית קטנה.
 - **[[:digit:]]** כל תו ספרה.
 - **[[:space:]]** כל סוגי הרווחים.
 - **[[:blank:]]** כל תו לבן (=שאינו מופיע בדפוס).
 - **[[:graph:]]** כל תו שמופיע בדפוס, ללא רווחים.
 - **[[:print:]]** כל תו שמופיע בדפוס, כולל רווחים.
 - **[[:control:]]** כל תו בקרה (=סימן פיסקה, שבירת שורה, סוף עמוד וכו').
 - **[[:punct:]]** כל סימן פיסוק.
 - **[[:word:]]** כל תו שהוא חלק ממילה (כולל מספרים).
 - **[[:x=]]** כל תו מאותו סט גליפים. דוגמה: **[[:a=]]** יחפש a, à, á, â, ã, ä, å, Å, A.
 - **[[:unicode:]]** כל תו שקוד יוניקוד שלו גבוה מ-255.
 - **[[:xdigit:]]** כל תו ספרה הקסדצימלית (0-9, A-F).

3. תווים פרטיים

- קודים לתווים וסמנים פרטיים כגון סימן פיסקה; טאב; זכויות יוצרים; מספור אוטומטי; סמן אובייקט מעוגן; רווחים למיניהם; ועוד רבים, אפשר למצוא בתפריט שבכפתור @ ליד שדה החיפוש/ההחלפה.
- תווים שאינם מופיעים בתפריט זה, אפשר להעתיק מהמסמך ולהדביק בשדה החיפוש/ההחלפה.

4. פרמטרים. פרמטרים מציינים את תנאי החיפוש. אפשר לשלב מספר פרמטרים בנוסחה אחת.

- חלק מן התווים הרגילים (כגון נקודה וסוגריים), משמשים ב-GREP כקודים. כדי לחפש אותם עצמם הקלד לוכסן הפוך \ לפני כל תו בנפרד.
 - כדי להגדיר קבוצה של תווים כ'רגילים', הקלד לפנייהם \Q ולאחריהם E.
- | קו מפריד מגדיר 'או'. דוגמה: ראובן|שמעון|לוי מחפש ראובן או שמעון או לוי.
- [] סוגריים מרובעים מגדירים ערכת תווים: כל תו או קוד שנמצא ביניהם יוגדר כ'או'. דוגמה: ע[למד]י יחפש עלי או עמי או עדי. ש[w\s\]ם יחפש כל ש ו-ם שיש ביניהם אות כלשהי או תו לבן כלשהו.
 - מקף בסוגריים מרובעים מגדיר טווח אותיות. דוגמה: מ[ב-ה]ה יחפש מבה, מבה, מדה, ו-מהה.
 - ^ בסוגריים מרובעים מגדיר תו לחיפוש שלילי. דוגמה: מ[^אד]ה יחפש כל תו שבין מ ל-ה, אך לא מאה ו-מדה. גם כאן מקף יגדיר טווח אותיות.
- פרמטרים למיקום הטקסט. יש להקליד לפני הטקסט או לאחוריו, בהתאם להקשר.
 - <\> סוגריים משולשים (עם לוכסן הפוך לפני) מגדירים תחילת מילה וסופה. דוגמה: <סוס לא יחפש הסוס, אך יחפש סוסים. חלל> לא יחפש חללים, אך יחפש החלל.
 - b\ מגדיר גם הוא גבולות מילה. דוגמה: בחלל יחפש חלל, ולא החלל. חלל\ב לא יחפש חללים.
 - B\ מגדיר מילה לחיפוש שלילי. דוגמה: \בחלל לא יחפש חלל בתחילת מילה, אלא רק באמצעה ובסופה. חלל\B יחפש חלל רק בתחילת מילה ובאמצעה.
 - ^ מתנה את החיפוש לתחילת פיסקה. דוגמה: ^שלום יחפש רק שלום שבתחילת פיסקה.
 - \$ מתנה את החיפוש לסוף פיסקה. דוגמה: \$שלום יחפש רק שלום שבסוף פיסקה.
 - z\ מתנה את החיפוש לסוף כתבה (Story). דוגמה: שלוםz\ יחפש רק שלום שבסוף כתבה.
- פרמטרים להתניית הטקסט. פרמטרים אלו פועלים רק בתוך סוגריים עגולים.
 - (?=>טקסט)טקסט. חיפוש מאחור חיובי מחפש את הטקסט שמחוץ לסוגריים, רק כשיש לפניו את הטקסט שבתוכם - אך בוחר רק את הטקסט שמחוץ לסוגריים. דוגמה: (?=>דוד)כהן יחפש את כהן רק כשלפניו יש 'דוד', אך יבחר רק את כהן.
 - (!>טקסט)טקסט. חיפוש מאחור שלילי מחפש את הטקסט שמחוץ לסוגריים, רק כשהטקסט שבתוכם אינו לפניו. דוגמה: (!>דוד)כהן יחפש את כהן רק כשאין לפניו 'דוד'.
 - טקסט(=?טקסט). חיפוש מקדימה חיובי מחפש את הטקסט שמחוץ לסוגריים, רק כשהטקסט שבתוכם לאחוריו - אך בוחר רק את הטקסט שמחוץ לסוגריים. דוגמה: דוד(=?כהן) יחפש את 'דוד' רק כשאחוריו יש כהן, אך יבחר רק את דוד.
 - טקסט(!?טקסט). חיפוש מקדימה שלילי מחפש את הטקסט שמחוץ לסוגריים, רק כשהטקסט שבתוכם אינו לאחוריו. דוגמה: דוד(!?כהן) יחפש את 'דוד' רק כשאין לאחוריו כהן.

● פרמטרים להישנות מספר התוים. פרמטרים אלו אפשר להחל על תו, קוד, או קבוצה. יש להקליד לפני הטקסט.

- ? יחפש אפס פעמים או פעם אחת. דוגמה: א?בשלום יחפש גם בשלום וגם אבשלום. \w?בשלום יחפש גם בשלום וגם אבשלום, בבשלום וכו'. (אב)?שלום יחפש שלום או אבשלום.
- * יחפש אפס פעמים או יותר. דוגמה: א*בשלום יחפש גם בשלום, גם אבשלום, וגם אאבשלום, אאאבשלום וכו'.
- + יחפש פעם אחת או יותר. דוגמה: א+בשלום יחפש גם אבשלום וגם אאבשלום, אאאבשלום וכו' - אך לא בשלום.
- חיפוש רגיל מחפש את התוצאה הארוכה ביותר. הוספת סימן שאלה לאחר סימנים אלו תגרום לחיפוש האפשרות הקצרה ביותר. דוגמה: \(.+)\) מחפש טקסט שבין סוגריים. הוא יבחר בתוצאה הארוכה יותר (שבאותה פיסקה): (מר (אברהם) כהן). או: אין (מצוה) להיות (בשמחה) תמיד. לעומתו \(.+?) (מר (אברהם) כהן). או: אין (מצוה) להיות (בשמחה) תמיד.

- {} בסוגריים מסולסלים אפשר לקבוע את מספר החזרות המדויק של התו/קוד לחיפוש.
- {מספר} יחפש את התו/קוד/קבוצה שלפניו כשהוא חוזר על עצמו כפי המספר בדיוק. דוגמה: אב{3}א יחפש רק אבבבא. א{2}א יחפש רק אבאבא.
- {מספר} יחפש לפחות מספר זה. דוגמה: ז{2}(ער) יחפש זערער, זערערער וכו'.
- {מספר,מספר} יחפש בין מספר זה למספר זה. דוגמה: א{4,2}א יחפש אממא, אמממא, אממממא.
- הוספת סימן שאלה לאחר הסוגריים תגרום לבחירת האפשרות הקצרה ביותר. דוגמה: פ{2,2}? יחפש גם פפפ ו-פפפפ - אך יבחר רק 2 פ בו זמנית (=פפ). א{4,2}? יחפש בין 2 ל-4 א - אך יבחר רק 2 א בו זמנית (=אא).
- \1 יחפש את הטקסט הנמצא כשהוא חוזר על עצמו - להבדיל מסוגריים מסולסלים, שמחפשים גם צורות שונות. דוגמה: (ש[כמ]ן)1 יחפש רק שכשכשן ו-שמןשמן. לעומתו (ש[כמ]ן){2} יחפש גם שכששמן ו-שמןשכשן.

● פרמטרים לאותיות גדולות/קטנות. יש להקליד לפני הטקסט.

- text(!?) מפעיל רגישות לאותיות רישיות. דוגמה: text(!?) לא יחפש Text או TEXT.
- (!-?) מבטל רגישות לאותיות רישיות. דוגמה: text(!-?) יחפש גם Text או TEXT.

● פרמטרים לשינוי ההתייחסות לטקסט. יש להקליד לפני הטקסט.

- (m?) הפעל שורות מרובות מתייחס לכל מעבר שורה ידני (כלומר אנטר/שיפט+אנטר) כאל פיסקה נפרדת. דוגמה: w^(m?) יחפש מילה שנמצאת בתחילת פיסקה, או לאחר מעבר שורה. זהו מצב ברירת המחדל.
- (m-?) בטל שורות מרובות יבטל התייחסות זו. דוגמה: w^(m-?) יחפש מילה שנמצאת בתחילת כתבה (Story) בלבד.
- (s?) הפעל שורה יחידה יתייחס לכל הטקסט כאילו הוא בשורה אחת ללא פיסקאות. דוגמה: [+.](s?) יחפש סוגריים מרובעים שביניהם תו אחד או יותר, גם אם הסוגר הראשון בפיסקה אחת והשני באחרת.
- (s-?) בטל שורה יחידה יבטל התייחסות זו. דוגמה: [+.](s-?) יחפש סוגריים מרובעים שביניהם תו אחד או יותר - רק אם הם באותה פיסקה. זהו מצב ברירת המחדל.
- (x?) הפעל התעלמות מרווח לבן מתעלם מרווחים לבנים שבאמצע הטקסט - גם אם הם מופיעים בנוסחת החיפוש. דוגמה: w\w\ w\w\w(x?) יחפש כל שלושה תוים במילה, למרות שאין ביניהם רווח.
- (x-?) בטל התעלמות מרווח לבן יבטל התייחסות זו. דוגמה: w\w\ w\w\w(x-?) יחפש תו במילה, רווח, ושני תוים במילה. זהו מצב ברירת המחדל.
- הערה: בנוסחת GREP משולבת ממספר פרמטרים, אפשר לבטל ולהפעיל לסירוגין פרמטרים אלו.

את המונחים באנגלית אפשר למצוא במדריכים האנגליים המצורפים למדריך זה.