

שם הקורס: Big Data Technologies (2 נק')
מספר הקורס: 236620 (תחת נושאים מתקדמים באלגוריתמים)
סמסטר אביב תשע"ז

ד"ר אשכר הלל, Yahoo! ד"ר רוני למפל, Outbrain	מרצים:
יום ראשון, 9:30-11:30	שעות הרצאה:
<ul style="list-style-type: none"> • 234322 מערכות קבצים <u>א</u> • 236363 מערכות מסדי נתונים <u>א</u> • 236370 תכנות מקבילי ומבוזר 	דרישות קדם:
http://webcourse.cs.technion.ac.il/236620/Spring2016	אתר הקורס מ- 2016:

תאור הקורס

אתגרי המחשוב היחודיים של חברות אינטרנט גדולות בעלות מאות מליוני משתמשים ופטה-בייטים של מידע דחפו לפיתוחו של דור חדש של מערכות לעיבוד עתיר-נתונים. קורס זה מכסה שלושה תחומים של עיבוד עתיר-נתונים: (1) עיבוד של משימות גדולות ובדידות במודל התכנותי של map-reduce ומימושו בקוד פתוח, Hadoop; (2) עיבוד אינקרמנטלי עם מערכות OLTP/NoSQL, כפי שמודגם ע"י המערכת Big Table של גוגל ומקבילתה בקוד הפתוח, HBase; ו- (3) עיבוד שטפי נתונים, כפי שמודגם באמצעות מערכת הקוד הפתוח Storm. כל אחד מהתחומים הנ"ל ילמד באמצעות סקירת צרכים חישוביים "מהחיים" אשר היוו כר להתפתחות המערכות הנלמדות, וכן באמצעות ניתוח היתרונות והחסרונות של אותן מערכות. בנוסף, הקורס יכסה מספר תחומים משיקים.

דרישות הקורס

הקורס יכלול 3 תרגילי בית, מתוכם 2 רטובים. להבנת הקורס תידרש לעיתים קריאה של מאמרים מהתחום. בסיום הקורס יערך מבחן סופי.

רשימת ספרות

מאמרים מכנסים, ותיעוד של מספר מערכות קוד פתוח והשימוש בהן.