

# מערכות אחסון ביג-דאטה - סילבוס

מרצה: דר' מרים אללוף

דוא"ל: [miriamal@ice.ac.il](mailto:miriamal@ice.ac.il)

משרד: מסדרון מרצים קומה 1 חדר 012.

שעת קבלה: על פי תאום מראש באמצעות דוא"ל

## תקציר

תחום האחסון עבר שינוי עצום בעשור האחרון. העולם הדיגיטלי של היום מלווה בכמות הולכת וגדלה של מידע מסוגים שונים הדורשת יכולות אחסון ועיבוד מודרניות המאפשרות שאלות וניתוח יעילים של הנתונים. הקורס עוסק בנושאים קלאסיים ומתקדמים במערכות אחסון סוקר את האתגרים העומדים בפני אלו הבונים וממשים מערכות ענק ובאבני הבניין המאפשרים סילומיות (scalability) וזמינות במערכות אלו.

הקורס יכלול את הנושאים הבאים:

- עקרונות הפעולה של דיסקים מכניים ו- solid-state-disks, RAID ומערכות מרכזיות מסוג SAN המרכזות מספר רב של דיסקים
- מערכות קבצים מקומיות ומבוזרות: Ex2/3 fs, NFS, LFS, Ceph & Rados, GoogleFS, HDFS
- דרכי התמודדות עם כמויות המידע (Big Data) העצומות הדורשות אחסון וארכיטקטורות אחסון בענן (Storage Cloud) כוללים נושאים רחביים כמו:
  - טופולוגיות וניתוב בין שרתים המחזיקים מידע: Distributed Hash Tables (DHT)
  - מבנה וסביבת חישוב המאפשרים טיפול בהרבה מידע: Hadoop stack
  - חישוב מקבילי בסביבה מרובת מחשבים mapReduce
  - Replication & Consistency - שכפול מידע לצורך Availability, Scalability and Reliability
  - מסדי NoSQL והשוואה בין ACID מול BASE
  - CAP Theorem
  - אמינות המידע וטיפול בנפילות שרתים: Failures and Fault Tolerance
  - שיטות לאחסון יעיל: Deduplication ודחיסה.
  - NOSQL databases: Column family, Graph DB and Document DBs

## דרישות הקורס

הקורס הינו במתכונת מעורבת ומורכב מהרצאות המרצה בכתשע מההרצאות הראשונות ומהרצאות הסטודנטים בשליש האחרון של הקורס. הסטודנטים יבחרו מאמר מרשימת מאמרים שתניתן בתחילת הקורס. המאמרים נלקחים מכנסים מוערכים בתחום ה-storage המדברים על מערכות עכשוויות. הסטודנטים ייפגשו עם המרצה במועד שייקבע לפני ההצגה כדי לדון במאמר שלהם ובמצגת בכיתה - על אחריות הסטודנטים לקבוע פגישה לקראת ההרצאה.

הציון מורכב:

- נוכחות בקורס חובה 5% - יש נוכחות חובה בחמשת השיעורים אחרונים (שיעורי הסמינר) ועוד 4 שיעורים לפחות
- מתן הרצאה בנושא מסוים לפי רשימה - 40%
- תרגיל ב SPARK - 10%
- בחינת בית 45% - עבודה זו תינתן בסוף הסמסטר - רק ציון עובר ומעלה בבחינת הבית יקנה ציון מעבר (הציון יהיה ממוצע המטלות)