

MindCET Snapshot #2

ביג דטה וחינוך

פברואר 2014



דוח MindCET יוצא לאור 4 פעמים בשנה.
כל גיליון מאיר תחום אחד בשדה המתפתח של
טכנולוגיה בחינוך. הדוח מסמן מגמות מרכזיות
ממגוון נקודות מבט פדגוגית, טכנולוגית, עסקית
ועוד. הדוח שלפניכם עוסק בתחום המשחקים
הדיגיטליים והממשק שלהם לשדה החינוכי.

תוכן עניינים

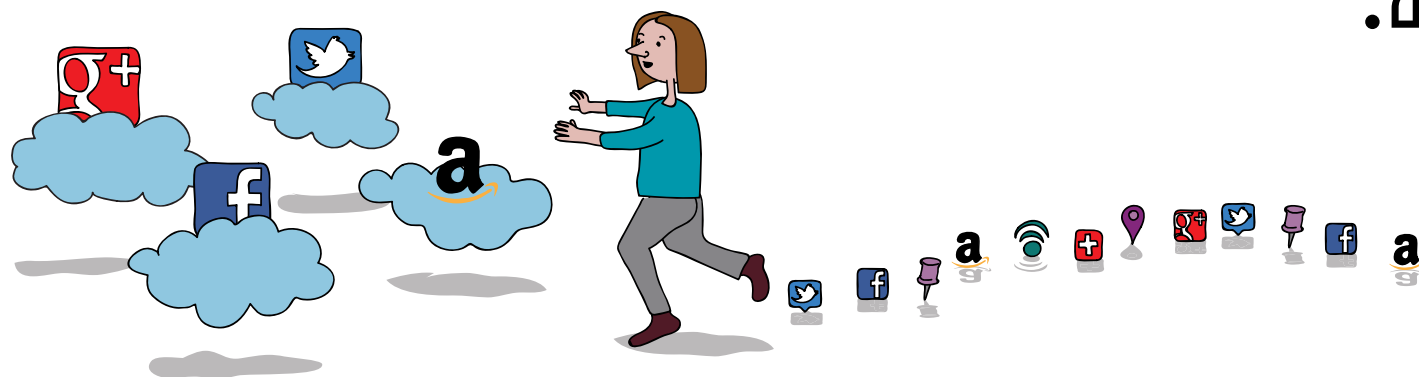
4-23	במבט כולל.....
24-39	חינוך וביג דטה.....
40-58	נתונים חדשים: פרטיות ומודעות לסביבות מקוונות.....
59-68	ביג דטה וחינוך - הזדמנות או איום?.....
69-73	המילון החסר.....
74-79	במבט מסכם.....



במבט כולל:

במבט כולל

חיינו מושפעים יותר ויותר מהחלטות המבוססות על נתונים.



בצורה פסיבית, כמויות עצומות של נתונים, ונתונים אלה הופכים למידע בעל משמעות ויוצרים את הזהויות הדיגיטליות האישיות והמשותפות שלנו. מערכות חכמות אלו לאיסוף ולפיצוח נתונים מיועדות לעזור לנו ללמוד יותר על העולם שאנו חיים בו. זהו הרקע למושג ביג דטה.

מערכות חכמות אלו לאיסוף ולפיצוח נתונים מיועדות לעזור לנו ללמוד

יותר על העולם שאנו חיים בו. זהו הרקע למושג ביג דטה.

בכל פעם שאנו מקלידים, מקליקים, מחליקים על המסך, מפעילים אפליקציה, ובמילים אחרות - משתמשים במכשירים דיגיטליים, אנו משאירים מאחורינו עקבות של נתונים למעקב, לניתוח ואף לשימוש מסחרי. במקום כלשהו נאספים כל הנתונים עלינו לכדי תמונה המאפשרת למכשיר להכיר כל אחד מאתנו באופן "אישי" ומובחן מההמון. שימוש בכלים שמקנים לנו המכשירים הדיגיטליים באמצעות רשתות חברתיות, מנועי חיפוש, אתרי קניות ואפליקציות סלולריות, מעניק לנו תחושת יעילות שאיננו יכולים לחיות בלעדיה. ה-GPS עוזר לנו להגיע הביתה מכל מקום, דף החדשות ב-Facebook בוחר את הפוסטים של החברים הקרובים ביותר, Google הופך את החיפושים שלנו לקלים יותר על ידי סינון תוצאות לא רלוונטיות וכו'. כדי שכל זה יקרה, אנו מזינים באופן קבוע, לרוב

מושג "חמקמק"



השימוש הנרחב במושג "ביג דטה" בכל אמצעי התקשורת ממחיש את השפעתו הבלתי מעורערת על כל תחומי החברה, מעסקים ועד לאקדמיה. עם זאת, עדיין קשה למצוא לו הגדרה ברורה ואחידה. על פי מגזין MIT Technology Review

[אוקטובר 2013], "ביג דטה" מחולל מהפיכה בעסקים של המאה ה-21 אף

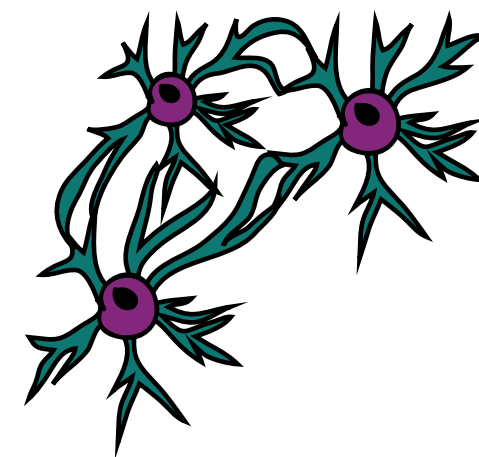
על פי שאף אחד אינו יודע בבירור את משמעותו. מעבר לכך, החשיבות של נפח הנתונים הולכת ופוחתת: "ביג דטה הוא אחד המונחים הגרועים ביותר שהומצאו בתעשייה. לא רק שהוא אינו מתאר כראוי את תפקידם ההולך וגדל של נתונים בחיינו, אלא הוא גם יוצר אובססיה למשתנה הלא נכון: נפח הנתונים".¹ טים סמית אומר שבחמש השנים האחרונות המידע הדיגיטלי, הנמצא כל הזמן במגמת צמיחה, מציב בפנינו אתגר - כיצד ליצור אמצעים חדשים לאחסון, לחיבור ולניתוח נתונים. סמית מזכיר לנו את מאיץ החלקיקים במעבדת המחקר CERN, שבה פיתח המדען הבריטי טים ברנרס-לי את ה-World Wide Web, ואת השפעתו החשובה על התפתחות "ביג דטה".

וורד ובייקר מאוניברסיטת קורנל פרסמו לאחרונה מסמך המנסה לנסח הגדרה ברורה, המבוססת על סקר שבדק כיצד חברות מובילות תופסות את המושג "ביג דטה". החוקרים הגיעו למסקנה כי "ביג דטה הוא מונח המתאר אחסון וניתוח של נתונים מקבצים גדולים ומורכבים תוך שימוש בטכניקות שונות, לרבות, אולם ללא הגבלה ל: NoSQL, MapReduce ולמידת מכונה (machine learning)".²

כדי לעזור לנו לצייר תמונה ברורה יותר, בדקנו מה
אומרים כמה מהאנשים הבולטים בתחום

"ביג דטה - עורק החיים של מערכת העצבים העולמית החדשה שלנו
- הוא המקור לטיפול באתגרים העולמיים הגדולים של ימינו. ביג דטה
מנצל את כוחם ההולך וגדל של המחשבים המרושתים כדי לבחון את
פרטי הפרטים של התפקוד של החברה"

(אלכס "סנדי" פנטלנד, פרופסור למדיה, אמנות ומדעים במכון הטכנולוגי של מסצ'וסטס (MIT).³)



"משחר הציוויליזציה ועד 2003 ייצר המין האנושי 5 הקסה-בייט
של נתונים. כיום אנו מייצרים 5 הקסה-בייט כל יומיים... והקצב רק
הולך וגדל"

(יו"ר Google, אריק שמידט).⁴

"ביג דטה מתייחס ליכולתנו לאסוף ולנתח את הכמויות העצומות של
הנתונים שאנו מייצרים כיום בעולם. הנתונים המתקבלים מניתוח
הביג דטה הם שמשנים את עולמנו, לא הביג דטה לכשעצמו

[ברנרד מאר].⁵

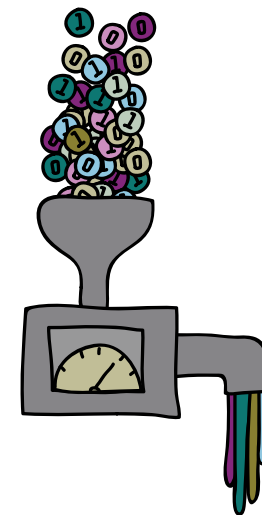


התפתחות הביג דטה היא כמו "לראות כיצד כדור הארץ מפתח
מערכת עצבים"

(מנכ"לית Yahoo, מריסה מאייר).⁶

"ביג דטה" מאפשר לנו למצוא מחט בערימת שחת"

(מייקל אייזנברג, קרן הון סיכון Aleph).



"אין אלגוריתם אחד שיכול לחזות את כל הנתונים ואין סדרת משתנים שיכולה לאפשר לנו לחזות הכול. זוהי עבודה מורכבת, בלאגן שלא לימדו אותי להתמודד אתו בשיעורי סטטיסטיקה"

(שרי ווסלי, EDUCASE, 2013).



ביג דטה "צפוי להפוך לאחד ממקורות הכוח הגדולים ביותר של

המאה ה-21"

(BBC Horizon 2013).

"עולם העסקים החדש בעידן הביג דטה דורש חשיבה שונה בתכלית, מבנים ותהליכים ארגוניים חדשים ומיומנויות חדשות של מנהיגות, כדי לפרש ולחבר נתונים בדרכים יצירתיות ומשמעותיות יותר."

(סר טרי ליהי, מנכ"ל Tesco לשעבר, 2013).



"ביג דטה הפך למשאב טבע חדש, משאב טבע מדהים"

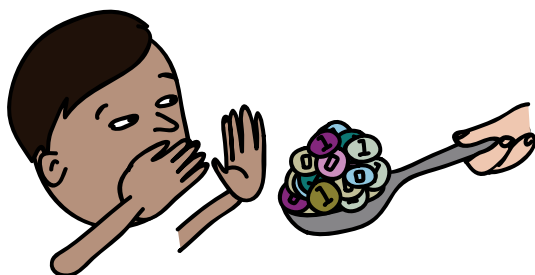
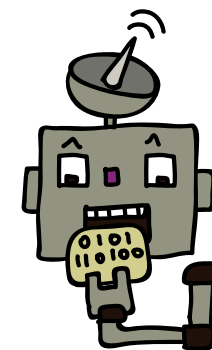
(ג'ים ספורד, מנהל IBM Global University Projects).



"ביג דטה מאפשר לגזור ערך מתהליך קבלת החלטות עסקיות, המבוסס על מסד נתונים יחסי מסורתי, בתוספת משאבים חדשים של נתונים שאינם מובנים."

(Oracle)

אנליסט Gartner, דאג לייני, הציע בשנת 2001 הגדרה הכוללת שלושה היבטים: נפח, מהירות ומגוון (The Three V's: Volume, Velocity and the Variety). הגדרה זו של ביג דטה צברה פופולריות, ונוסף לה היבט רביעי, מהימנות (Veracity), המאפשר לכלול שאלות של אמינות ואי-ודאות.



"ביג דטה הוא מונח שמשמשים בו יותר ויותר כדי לתאר תהליך של יישום כוחות מחשוב רציניים -המילה האחרונה בתחום למידת מכונה ובינה מלאכותית - על מערכות מסיביות, ולרוב מורכבות ביותר, של מידע"¹¹(Microsoft)

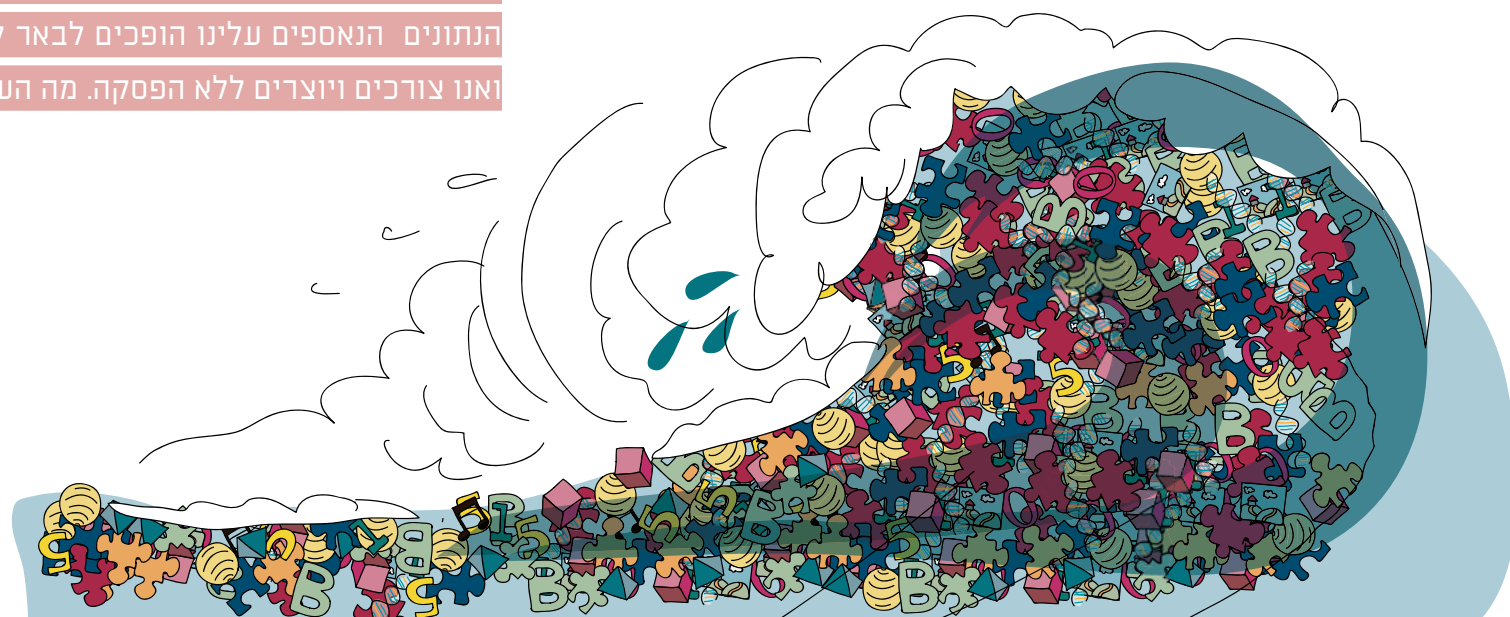
המכון האמריקני הלאומי לתקנים ולטכנולוגיה (NIST) טוען, כי ביג דטה הם נתונים ש"חורגים מעבר לקיבולת או ליכולת של שיטות ושל מערכות קיימות או קונבנציונליות". במילים אחרות, המושג "ביג" הוא יחסי לסטנדרט המחשוב הנוכחי.¹²

צונמי של נתונים



חידושי לשון נולדים מדי יום כדי לנסות להגדיר חוויות חדשות המבוססות על נתונים. דוגמה טובה הוא המונח "dataself",¹³ מונח שמנסה להגדיר את חוסר היכולת שלנו, בעידן הנוכחי, להפריד בין רגשות סובייקטיביים ובין חוויות אישיות "מונעות נתונים" - (כלומר, מבוססות על נתונים). גם המושג "webdata" נועד לבטא את הגידול חסר הרסן בכמות הנתונים הנאספים מהאינטראקציות הדיגיטליות היומיומיות שלנו. יש כיום גדגיטים המתעדים כמה חשמל צורך כל מכשיר חשמלי בביתנו, גנום צרכני המאפשר להתאים טיפול רפואי בהתאמה אישית, וצמיד Nike Fuel Band העוקב אחר נתונים אישיים שלנו בזמן שאנו עושים פעילות גופנית. "גדגיטים כאלו "רעבים" לנתונים, והם מנצלים את היכולת שלהם לחבר אנשים לנתונים שהם משתמשים בהם ולגרום להם לראות כיצד הם יכולים לשנות את חייהם.¹⁴ "הנתונים הדיגיטליים שלנו והשיקוף שלהם חזרה אלינו "הפכו לממד חדש שהופך את החוויות שלנו לאמיתיות".¹⁵

אנו נמצאים תחת צונמי של נתונים, לדברי כריס דה לה טורה: "היום אנו חיים ואוהבים בתוך המכשירים שלנו - בכל דקה משעות היום אנו משתמשים במחשבים ניידים, בטלפונים חכמים, בטבלטים ובקרוב ב Google Glass. הנתונים הנאספים עלינו הופכים לבאר ללא תחתית שממנה שואבים רעיונות, ואנו צורכים ויוצרים ללא הפסקה. מה העולם חושב? החליקו, הקליקו ותגלו."¹⁶





משוב בזמן אמת על נתונים

יתרה מזאת, באמצעות ניתוח הנתונים אפשר לפתח דרכי פעולה יעילות יותר לשיפור השקיפות וחובת הדיווח של המוסדות (McKinsey Report 2013, Open data).

בפני האזרחים נפתחות אפשרויות גישה למידע רב יותר מאי פעם. ההשפעה על המערכות הציבוריות, כגון שירותי הבריאות, משמעותית ביותר. היא מאפשרת לכל אדם להגיע לנתונים הרפואיים שלו, מידע אישי ומותאם לו אישית, וכן למידע הכולל גם נתונים על טיפול מונע, מידע על מגיפות, מגמות בריאות ואפשרויות שונות של טיפולים ונתונים על אנשי מקצוע מתאימים. באמצעות הנתונים האלה כל אחד יכול לשלוט על הטיפול הרפואי הניתן לו.

כיום אפשר להשוות בזמן אמת בין נתונים לבין מה שמתרחש במציאות. הנחות ותחזיות מאומתות ומתוקנות עוד במהלך איסוף המידע. **משוב מידי כזה מחולל מהפכה ביכולתנו להבין ולפעול על פי אירועים אמיתיים.** "המספרים והניתוח שלהם מאפשרים לנו לעשות קישורים מבוססי נתונים, אשר אנשים המשתמשים רק בתחושות הבטן שלהם, מנוסים ככל שיהיו, לעולם לא יוכלו לנחש", אומר פרופסור וויליאם דלונה, מומחה לאפקטיביות של מערכות מידע. "זה כוחו של המידע, הוא אינו משקר".¹⁷

בניתוח ביג דטה אפשר לעבור על כמויות עצומות של נתונים ומידע, שנאספים בזמן אמת ובאופן שוטף, למצוא דפוסים סמויים ולגלות אי התאמות, ואלה עשויים להוביל לידע חדש ולהצביע על הזדמנויות לשירותים ולמוצרים חדשים.

מאליטה של מעטים לכל משתמש

פרופסור מאייר-שונברגר מאוניברסיטת אוקספורד אמר לאחרונה כי יופיים של נתונים טמון בכך ש"ערכם אינו מתמצה לאחר שהשתמשו בהם. יש לשתף נתונים, ולא לשמור אותם, וכל גוף המנסה לעשות זאת קרוב באופן מסוכן להתנהגות רודנית" (דצמבר 2013, Online Educa Berlin).

למידה מנתונים, אשר בעבר הייתה תחום בלעדי למומחים, מוצעת כיום לכל דורש בסביבות שונות, ומהווה זרז לשינויים בהתנהגות של מערכות. "המעקב האישי עושה לשירותי הבריאות מה שהמחשב האישי עשה למחשב: הוא שחרר את הלמידה מהנתונים מהיותה נחלתה של אליטה של מעטים והפך אותה לכלי המיועד להמונים" (מפתח אפליקציות, רגייב מהטה¹⁸. עצם העובדה שהמשתמשים יכולים לגשת למידע, המאומת על ידי נתונים אמיתיים, מעניק שקיפות ומאפשר קבלת החלטות מושכלת יותר. Google analytics, לוחות מחוונים של משחקים דיגיטליים, סטטיסטיקות של Facebook או Twitter - אלו הם רק כמה דוגמאות כיצד אפשר להשתמש בביג דטה על מנת לספק מידע משמעותי בזמן אמת למשתמשים כלשהם.

על פי דו"ח של חברת מקינזי משנת 2013, מגמה חדשה תגביר עוד יותר את נגישות הציבור לנתונים: נתונים פתוחים ונזילים. נתונים פתוחים (open data) משמעותם "שחרור מידע על ידי ממשלות ומוסדות פרטיים והנגשה לציבור של נתונים פרטיים (לא מזהים), כדי לאפשר תובנות בענפים שונים".¹⁹ נתונים נזילים (liquid data) משמעותם הפיכת נתונים פתוחים לנתונים הזמינים לכול בפורמט שניתן לשיתוף. יזמים מנצלים הזדמנות זו ויוצרים ערך מתוך אותם נתונים נזילים.



קישוריות כבסיס לחדשנות

אפשרויות פעולה משותפת

פרופסור אלכס פנטלנד מ-MIT מדגיש את היבט הקישוריות. לדעתו הקישוריות הופכת רשתות, שהיו בעבר בשליטה ריכוזית, לבלתי מתאימות לאתגרים של ימינו. "במקום להתמקד רק בגישה ובמערכות הפצה, אנו זקוקים למערכות דינמיות, מרושתות, בעלות ויסות עצמי ועמידות, המתחשבות בתלות החברתית-כלכלית ההדדית המורכבת של העולם המקושר של ימינו".²⁰ לטענתו, שטף המידע יכול לגרום לנו למצוא דפוסים חדשים, שהינם מרכיב בסיסי לחדשנות. ריצ'רד מרציאנו, ממעבדת החדשנות הדיגיטלית באוניברסיטת צפון קרוליינה בצ'אפל היל, מדבר על הזדמנויות חדשות לשיתוף פעולה: "זה לא רק הבלגן של כל הנתונים האלו, אלא הרעיון שביג דטה יכול ליצור שיתופי פעולה נרחבים, המעוררים שאלת מפתח: כיצד אנשים יכולים להסתדר ולהעלות לדיון סוגיות שונות? שיתופי פעולה בקנה מידה נרחב מובילים גם לרעיונות גדולים יותר. על כן צריך לחשוב כיצד נוכל להתוות כיווני מחקר ולפתח גישות חדשניות הנהנות ממגוון זה?"²¹

ישנם כמה פרויקטים המנסים להשתמש בכוחו של הביג דטה לטובת הכלל, כגון: העלאת המודעות לבעיות עולמיות שונות, הקצאת מקום להשפעה על מערכות ציבוריות כמו בריאות או חינוך, או מתן גישה לפיתוח כלים שימושיים לחברה (OpenStreetmap בהאיטי או Code for America בארצות הברית).²² יוזמות כגון Data without Borders (DataKind.org) תורמות להשפעת הביג דטה ולהפיכתו לכלי לטובת הכלל, לשימושם של אנשים ולמען אנשים. בפרויקט זה עובדים יחד מדעני נתונים וארגונים חברתיים המאמינים כי שיפור הגישה וההבנה של נתונים בתחום החברתי יכולים לתרום לקבלת החלטות טובה יותר ולהשפעה חברתית רבה יותר. "בנתונים טמון פוטנציאל להפוך קשרים סמויים לגלויים, לשמש שפה משותפת בין אנשים אשר אחרת אולי לא היו מדברים זה עם זה, להוות קרש קפיצה לשיתוף פעולה, להציע מדדים לקבלת החלטות, ולהפוך רעיונות, שלכאורה נראים לא קשורים זה לזה, לתובנות רבות עוצמה שיש בהן פוטנציאל לפתור את הבעיות המורכבות והקשות ביותר הנצבות לפנינו".

השפעה על המדע

איסוף נתונים, לפני הנחת המחקר?

המתמטיקאי סימון דדאו מצביע על השפעה משמעותית נוספת על המחקר. "כעת יש לנו נתונים רבים חדשים (שלוקטו) ממערכות ביולוגיות וממערכות חברתיות אנושיות, ואלה נאספים לפני שיש לנו אפילו הנחת מחקר. הנתונים נמצאים שם בצורה מרובת ממדים וחסרת סדר, והם מחכים שיחקרו אותם. אולם כיצד אפשר לדעת אילו שאלות לשאול כאשר השיטה המדעית נהפכה על פיה?"²⁵

האפשרויות החדשות שמביא הביג דטה מאיצות את המדע ומחוללות מהפך בהבנתנו את מורכבותם האינסופית של נושאים, החל מננו-ביולוגיה וכלה בגילוי יקומים חדשים. "מערכות מקרו, או מערכות אקולוגיות גדולות (big ecology)", כפי שמכנה אותן דיוויד שימל,²³ מדען מחשב בכיר בנאס"א, "הופכות אפשריות אך ורק כאשר יש גישה לנתונים בקנה מידה רחב מספיק. קבוצות נתונים גדולות ועשירות מאפשרות למדענים לשלב את המורכבות ואת השונות שבעולם האמיתי בדגמים המתארים תופעות בקנה מידה רחב. CERN, מעבדת הפיזיקה הגרעינית השווייצרית, משתמשת בכוחות מחשוב של אלפי מחשבים, הפזורים ב-150 מרכזי נתונים ברחבי העולם, כדי לנתח נתונים שיחשפו את סודות היקום. מדעני המחשב הוד ליפסון ומייקל שמידט מאוניברסיטת קורנל פיתחו תוכנת בינה מלאכותית לרובוטיקה, תוך שימוש בנתונים רבים ומורכבים, אשר בני אדם לא יכולים להשתלט עליהם."²⁴



תזכיר לי שוב מה אנחנו
רוצים להוכיח?

שינויים מבניים

ההתפתחויות הטכנולוגיות האחרונות, במיוחד שרתי ענן והתפתחות רשתות נתונים, מרחיבות את הפוטנציאל הגלום בביג דטה. הפיזיקאי הארוי ניומן מצביע על שינויים מבניים משמעותיים שדורש הביג דטה. אם המגמות הקיימות יימשכו, צורכי המחשוב של ניתוח הנתונים יטילו עומס משמעותי על מארג המחשוב. "ביג דטה דורש חשיבה על מערכת מסוג שונה. הארכיטקטורה המועדפת כבר אינה בנויה מיחידת עיבוד מרכזית (CPU) בודדת ומתוספת של זיכרון גישה אקראית (RAM) וכונן קשיח לאחסון ארוך טווח. גם מחשבי-על גדולים ומרוכזים מקבילים, ששלטו בשנות השמונים והתשעים, מפנים את מקומם למרכזי נתונים מבוזרים ולמחשוב ענן, המרושתים לעתים קרובות בתוך ארגונים רבים ועל פני מרחקים גאוגרפיים עצומים".²⁶

מחשוב ענן מספק לכלל הציבור גישה קלה, הניתנת להרחבה, למשאבי מחשוב. Amazon Web Services הוא כיום ספק הענן הציבורי הגדול ביותר.²⁷

לדברי שונברגר וקוקייה (Shonberger and Cukier),²⁸ שלושה חידושים טכנולוגיים גורמים לכך שביג דטה משנה את האופן שבו אנו חיים, עובדים וחושבים: דטפיקציה (datafication) מוגברת של דברים, קיבולת אחסון זיכרון מוגדלת וכוח עיבוד מוגבר.



ויזואליזציה: הפיכת אוריינות ביג דטה לאוניברסלית

העידן הדיגיטלי הביא עמו את ההכרה כי ויזואליזציה צריכה להיות תכונה מרכזית של העברת מידע בזמינות אוניברסלית ושיתוף רעיונות עם אחרים. במקביל להתפתחות ניתוח הנתונים חלה התפתחות יצירתית וחכמה של אוריינות דיגיטלית ויזואלית, וזו מציעה לכל המעוניין מידע שנחשב עד כה קשה להבנה ובלעדי למומחים. בוועידת TED בשנת 2006 הדהים האנס רוזלינג²⁹ את הקהל כאשר הפיח חיים בסטטיסטיקה. הוא רצה להעביר לסטודנטים שלו מסר חשוב בנושא בריאות וכלכלה בעולם המתפתח, ועל כן פיתח תוכנה שבה בועות נעות ועקומות משיטות הציגו נתונים מורכבים בצורה ברורה ואינטואיטיבית. בשנת 2011, בוועידת TED אחרת, השאיר דב רוי את הקהל פעור פה, כאשר הציג תוצאות מחקר על רכישת שפה, והדגים בצורה ויזואלית ניתוח של נתונים מורכבים, שנאספו מ- 90,000 שעות צילום בכמה מצלמות שונות שהסריטו תנועות של ילד ומשפחתו.³⁰

ספקנות

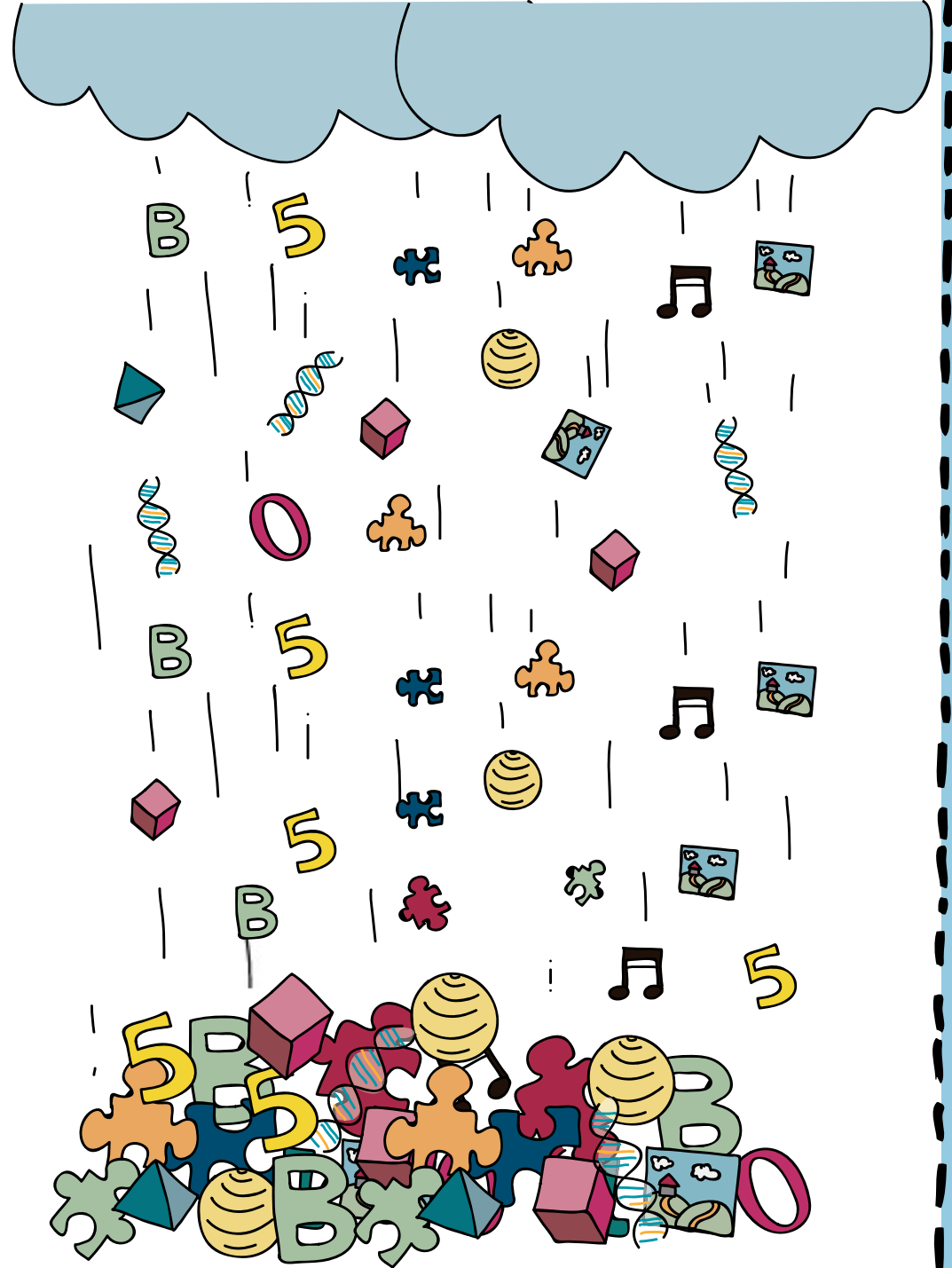
ספקנות היא חלק חשוב מהרעש התקשורתי סביב ביג דטה. מומחים רבים מסכימים כי דרך ארוכה לפנינו עד אשר נוכל בקלות להפוך את הנתונים המורכבים הנאספים כיום למידע בעל משמעות. "הביג דטה של היום רועש, לא מובנה, דינמי ולא סטטי. הוא עשוי להיות גם פגום" (אלסנדרו וספיניאני).³¹ "חלק גדול משיגעון הנתונים המתחולל לאחרונה, החל במדעי החיים והפיזיקה וכלה בתוכן שמייצרים המשתמשים ונצבר על ידי Google, Facebook ו-Twitter, הוא למעשה שטף לא מובנה של ערב רב דיגיטלי, הדורש מסדי נתונים חדשים וגמישים, כוח מחשוב מסיבי ואלגוריתמים מתוחכמים כדי לחלץ מהם פיסות משמעות" (מאט למאי, Bitly).³² "ביג דטה אינו מעשה קסמים. אין משמעות לכמות שיש לך, אם אינך יודע כיצד למצוא את ההיגיון הטמון בהם." היזם סטיבן רוזנבאום (Magnify.net) מתריע, כי אנו זקוקים ל"גיבורי על" במהירות הבזק כדי להכניס סדר והיגיון בגל הגוואה של נתונים ומידע, ויש לזכור כי העובדה שאנו מייצרים נתונים טובים אינה בהכרח אומרת שאנו מבינים אותם טוב יותר. "בעוד מכשירים נאבקים להבחין בין דואר זבל לבין הודעות מחברים, בין מידע קריטי לבין שטויות, בין אות לרעש, כמות הנתונים המגיעה אלינו מדהימה יותר ויותר."³³

פתרונות מוצעים למציאת היגיון

במורכבות הביג דטה

כיום נבדקים כלים מבטיחים וחדשניים, שמטרתם להתמודד עם העולם החדש של הנתונים. המתמטיקאי רונלד קויפמן³⁴ אומר כי מה שדרוש הוא מקבילה למהפכה של ניוטון: "אין זה מספיק רק לאסוף ולאחסן כמויות מסיביות של נתונים. יש לאצור אותם בצורה חכמה, וזה דורש מסגרת עולמית". קויפמן מאמין כי המתמטיקה המודרנית - בעיקר הגאומטריה - יכולה לעזור לזהות את האתגרים העולמיים המשתמעים מכך.³⁵ המתמטיקאי אלסנדרו וספיאני משתמש בכל דבר, החל מניתוח רשת (יצירת רשתות של קשרים בין אנשים, חפצים ומסמכים שיעזרו לגלות את המבנה הטמון בנתונים) ועד ללמידת מכונה ולסטטיסטיקה הוותיקה: "בסופו של דבר, מדע הנתונים הוא יותר מסך כל חלקיו המתודולוגיים", והדבר נכון גם לגבי הכלים האנליטיים שלו: **"כאשר משלבים**

דברים רבים יוצרים משהו גדול יותר, חדש ושונה." האדרי ניומן צופה שהביג דטה יישען על חישובים אוטומטיים באמצעות "צבאות" מתואמים היטב של סוכנים חכמים, העוקבים אחר תנועת נתונים מנקודה אחת ברשת לאחרת. כל נקודה יכולה לתעד רק את מה שקורה באופן מקומי, אך היא תשתף את המידע בצורה שתשפוך אור על מצבה העולמי של הרשת. "אלפי סוכנים ברמות שונות מתואמים כדי לעזור לבני האדם להבין מה קורה במערכת מורכבת ומבוזרת מאוד". קנה המידה יהיה גדול אף יותר בעתיד, כאשר יהיו מיליארדי סוכנים חכמים כאלה שייצרו ישות עולמית מבוזרת חכמה ועצומה. "היכולת ליצור דברים מעין אלה ולגרום להם לעבוד לטובת האחר תקטין את מורכבות הבעיות התפעוליות האלו. כאשר מתעוררת בעיה מסובכת במערכת כזו, אף קבוצה של בני אדם לא יכולה באמת להבין הכול ולגשת לכל המידע".³⁶ מבין הפרויקטים העוסקים כיום בביג דטה, אחד הפרויקטים המשמעותיים הוא Apache Hadoop, פרויקט קוד פתוח המאפשר עיבוד מבוזר של קבוצות נתונים



אנשי מקצוע מעטים וחסרי הכשרה

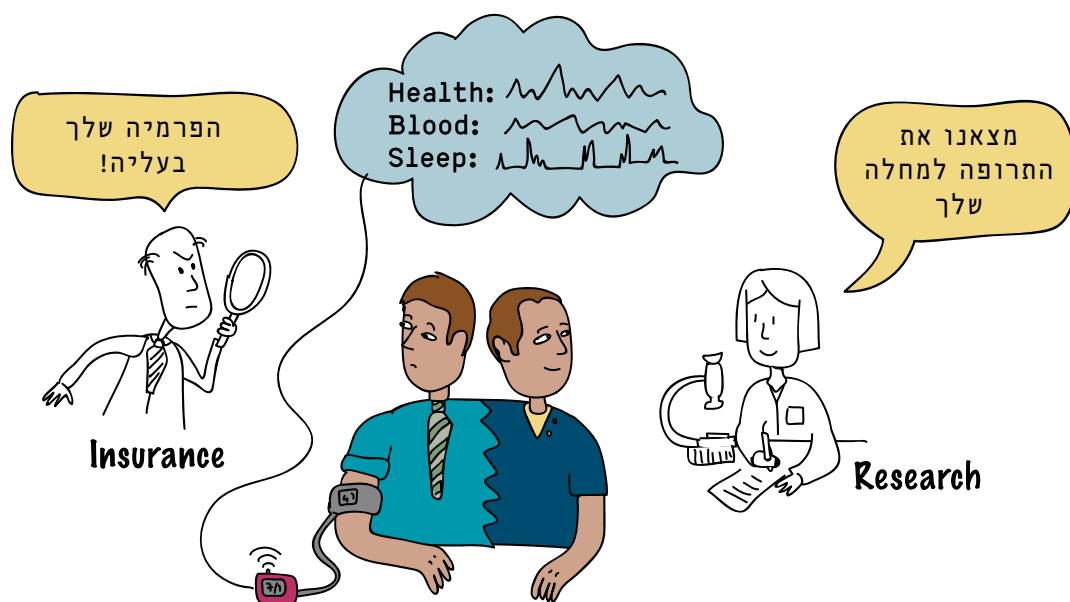
מספיקה

היעדר ידע ואנשי מקצוע שיעסקו במערכות נתונים מורכבות ומגוונות כל כך ויפרשו אותן מעלה סימני שאלה. פאנל של מומחים להשכלה גבוהה (2013) (Campus Computing Project) הביע חששות מציפיות גבוהות ומהשקעות גדולות באיסוף, ניהול וניתוח נתונים ומהצורך להשאיר סטודנטים בתחום ולספק להם הנחיה יעילה יותר. "ביג דטה עשוי לחולל מהפך, אולם הציפייה שמהפך זה יהיה מיידית אינה הוגנת".³⁷ "הבעיה הגדולה ביותר היא שכאשר אנשים במערכת ההשכלה הגבוהה שומעים כיום את המושג ביג דטה, הם חושבים שהם יכולים להדביק את הפער ולהיות כמו כל אחד אחר בשוק בתוך חודש." (פיל אייס, סגן נשיא מחקר ופיתוח ב-American Public University System.³⁸ כמו בכל תחום המנסה ליהנות מיתרונותיהן של טכנולוגיות חדשות, אנשים מתעלמים לעתים קרובות מתהליך היישום וההתמקצעות. ארגונים רבים, אשר יישמו בינה מלאכותית ויוזמות אנליטיות, תוכנה או כוח אדם הולמים ללא הכנה ראויה, לא השיגו את התוצאות הצפויות. חברות יכולות להיסחף עם האפשרויות הטמונות בכלים אלו, מבלי להצליח לפתח את האסטרטגיות המתאימות ליישום הנכון ביותר האפשרי.³⁹ כך קרה בעשור האחרון, כאשר הותקנו מחשבים במערכות החינוך ללא הכנת תשתית מתאימה של פס רחב או הכשרה מקצועית הולמת, והתוצאה הייתה בזבוז משאבים ותוצאות מאכזבות.⁴⁰

חששות אלה מהדהדים במערכת החינוך וההשכלה הגבוהה. בהשכלה הגבוהה ניכרת עלייה ברורה בהשקעות בתשתיות הנוגעות לביג דטה. אוניברסיטת רוצ'סטר השקיעה יותר מ-100 מיליון דולר, אוניברסיטת אינדיאנה השקיעה יותר מ-30 מיליון דולר (7.5 מיליון דולר הושקעו במחשב-על, המיועד לעיבוד כמויות גדולות של נתונים ומכונה Big Red II). קרן גורדון ובטי מור וקרן סלואן התחייבו להעניק 37.8 מיליון דולר לאוניברסיטת קליפורניה בברקלי, אוניברסיטת וושינגטון ואוניברסיטת ניו-יורק לשיתוף פעולה בנושא.

גדולות במקבצי שרתים. Hadoop מאפשר ניתוח נפחים חסרי תקדים של נתונים לא מובנים, שכיום מופקים כווידאו, אודיו, הודעות מדיה חברתית, תמונות וכדומה, וכל זאת במהירות עצומה. "בעולם המחובר של היום, שבו נוצרים יותר ויותר נתונים מדי יום, יתרונותיו פורצי הדרך של Hadoop משמעם שעסקים וארגונים יכולים עתה למצוא ערך בנתונים אשר עד לא מזמן נחשבו לחסרי שימוש."

פרטיות ומודעות



“ביג דטה הפך אמצעי למיקוד כל העולם בבת אחת: הוא מאפשר ליצור מתאמים בין מקבצי ענק של נתונים, הגילויים על סמך נתונים אלה צפויים להיות בעלי ערך למדענים, לאנשי חינוך, לאנשי רפואה וגם לסוכנות לביטחון לאומי, NSA, (המודיעין האמריקאי) ול- GCHQ (המודיעין הבריטי).” (פרופסור ויקטור מאייר-שונברגר, Oxford Internet Institute).

היתרונות של השימוש בקוד פתוח והשימוש החסכוני והיעיל בו, צריכים לעמוד מול חששות בנושא פרטיות. יש לעודד מודעות לנושא השימוש במידע אישי, ובייחוד בשימוש לרעה במידע זה. נורות אדומות בנושא נגישות לנתונים וערכה הממשי של הפרטיות פרצו לתודעה בעקבות פרסום ויקיליקס ומסמכי אדוארד סנאודן, שחשפו את פעילויות הצבא, הדיפלומטיה וגופי המודיעין של ממשלות שונות. Online Educa Berlin 2013 הקדיש דיון פאנל לנושא: “קץ הסודיות ומשמעותו”. ד”ר הרולד אלסטון קבע, כי לאור הנגישות ההולכת וגוברת לנתונים, הסודיות הכוללת הופכת לבלתי מעשית בעידן המודרני, ובעתיד “סודיות ואבטחה יהיו בלתי אפשריות ללא הסכמה”.⁴¹

הבנת אופן השימוש בביג דטה והשפעתו על חיי היומיום:

1. מנוע החיפוש Google הוא הדוגמה הטובה ביותר לביג דטה. Google מציע לכל אדם, שיש לו נגישות למכשיר דיגיטלי, למצוא מידע מכל סוג, ובכך הוא שינה לחלוטין את הרגלי איסוף הנתונים של האנושות. על פי המגזין פורבס, Google מחזיקה ב"מסד הנתונים הגדול ביותר על פני כדור הארץ".⁴²
2. גדג'טים אישיים, כגון שעונים וצמידים חכמים, אוספים נתונים ומידעים אותנו על תפקוד הגוף שלנו באמצעות ניתוח נתונים קולקטיביים. לדוגמה, צמיד Up של חברת Jawbone; הצמיד אוסף נתונים על צריכת הקלוריות שלנו, על רמות הפעילות ועל דפוסי השינה שלנו. החברה מנתחת נפח עצום של נתונים שנאספו במשך 60 שנה על דפוסי שינה, ומפיקה מידע שוטף למשתמשים הפרטיים.
3. רוב ענפי הספורט המקצועי אימצו כיום ניתוחי ביג דטה: IBM SlamTracker המיועד לטורנירי טניס; ניתוח וידאו העוקב אחר ביצועיו של כל שחקן במשחק כדורגל או בייסבול; טכנולוגיית חיישנים בציד ספורט, כגון כדורי סל, או מועדוני גולף (המספקים משוב באמצעות הסמרטפון ושרתי ענן); ושימוש בטכנולוגיה חכמה כדי לעקוב אחר תזונה ושינה של אתלטים או שימוש בשיחות במדיה חברתית כדי לנטר את מצבם הרגשי.
4. רוב אתרי ההיכרויות באינטרנט מיישמים כלי ביג דטה ואלגוריתמים כדי למצוא את השידוכים המתאימים ביותר.



5. Google Ngram Viewer מאפשר לנו להבין מגמות תרבותיות באמצעות חיפוש מילים מסוימות. הכלי מבוסס על מקבצי נתונים עצומים שנבנו על סמך מיליוני ספרים ש-Google העבירה לפורמט דיגיטלי במשך השנים.
6. ניתוח ביג דטה מאפשר ניטור וחיזוי של התפתחויות והתפרצויות של מחלות ומגפות. לדוגמה, אפשר לאתר התפרצות של מחלת השפעת על ידי שילוב נתונים מרישומים רפואיים עם ניתוח של נתונים ממדיה חברתית (מתוכן שאנשים מקלידים, לדוגמה: "מרגיש נורא היום - שוכב במיטה עם הצטננות").
7. Reg4ALL הוא יוזמה המקדמת פעולה אזרחית לרווחת הכלל. זוהי פלטפורמה המאפשרת לאנשים ולקהילות ליצור מסד נתונים פתוח לבריאות הציבור. האנשים מוזמנים לתרום מרצונם תוצאות של בדיקות רפואיות בנושאים מסוימים למסדי נתונים פתוחים, ואלה מאפשרים לאנשי רפואה להבין טוב יותר את הנושא ובסופו של דבר להציע טיפולים יעילים.⁴³
8. אופטימיזציה של זרימת התחבורה מבוססת על מידע בזמן אמת, על מדיה חברתית ועל נתוני מזג האוויר. ישנם פרויקטים ניסיוניים המשתמשים בניתוח של ביג דטה לפיתוח ערים חכמות, שבהן תשתית התחבורה והשירותים הציבוריים משולבים יחד: אוטובוס יכול לחכות לרכבת מאחרת, ורמזורים חוזים נפחי תנועה ופועלים לצמצום פקקי תנועה. דוגמה לכך היא יישום Apache Hadoop של אינטל, שנועד לנסות לעזור לערים בסין שבהן התחבורה עמוסה במיוחד.⁴⁴
9. ממשלת ארצות הברית משקיעה סכומי עתק לשיפור האבטחה באמצעות אכיפת החוק. לדוגמה, היא מאפשרת ל-NSA להשתמש בניתוח ביג דטה כדי לסכל התארגנויות טרור או כדי לאתר ולמנוע מתקפות סייבר.
10. נתוני מדיה חברתית, יומני אירועים, ניתוח טקסט ונתוני חיישנים משמשים לקבלת תמונה על משתמשים ולהבנת התנהגויותיהם והעדפותיהם ליצירת מודלים לחיזוי. חנויות ורשתות שיווק יכולות כיום לכוון את הפרסום בצורה

מדויקת מאוד לקהל היעד. באחד המקרים המפורסמים בעניין זה אב כעס על חנות Target, מפני שבתו קיבלה פרסומות בנושא היריון. לאחר מכן הוא גילה, כי Target "ידעה" שהבת בהיריון עוד לפניו, על פי שינוי בהרגלי קניית מוצרי הקוסמטיקה שלה.⁴⁵

11. אופטימיזציה של תהליכים עסקיים אפשרית על בסיס תחזיות הנוצרות על ידי נתונים, כגון נתוני מדיה חברתית, מגמות חיפוש באינטרנט ותחזית מזג האוויר. תחזיות כאלה עוזרות, לדוגמה, לקמעונאים להתאים את המלאי שלהם. דוגמה נוספת היא חיישני מיקום גאוגרפי וזיהוי תדרי רדיו - אפשר להשתמש בהם כדי לעקוב אחר סחורות או רכבי משלוחים ולשפר את מסלול הנסיעה שלהם על ידי שילוב נתוני תנועה בזמן אמת.

12. פרופסור סבסטיאן תרון (אוניברסיטת סטנפורד) ופיטר נורוויג (מדען נתונים, Google) עומדים בראש פרויקט לבניית מכונית ללא נהג, המבוססת על אלגוריתמים של בינה מלאכותית ועל כל הנתונים הנאספים מתייעוד וממידות של כלי רכב שנעשים על ידי Street View של Google.⁴⁶



מדוע ביג דטה כל כך גדול?

הגודל של הביג דטה נמדד בגודל השפעתו על החברה: האפשרויות הגדולות והולכות שהוא מעניק לנו, כמשתמשים דיגיטליים, תוך חיזוק ערך המידע בחיי היומיום שלנו. נתונים הם מקור כוח בעולם המודרני, ומצרך יקר ערך למי שיכול לנתחם. מעבר לכך, הביג דטה נותן הזדמנות לכולם לגשת למידע שבעבר נחשב מחוץ להישג ידו של הציבור והיה נחלתו של מיעוט מיוחס. כיום כולנו משתמשים בנתונים דיגיטליים ואף תורמים נתונים משל עצמנו. אולם ביג דטה יצר גם ציפיות וחששות, כפי שקרה ברוב המהפכות שהתחוללו במאה שלנו, ויש להתמודד עמן ומהר. הנתונים מאפשרים לנו לדעת יותר, אבל לא לדעת הכול! ערכם האמתי של הנתונים נעוץ במה שאנו עושים איתם. גל הנתונים כבר שוטף אותנו, ועלינו להיות חכמים דיינו כדי להפוך את האתגרים להזדמנויות, בעיקר בתחומים התורמים לציבור ולרווחת הכלל, כמו חינוך!

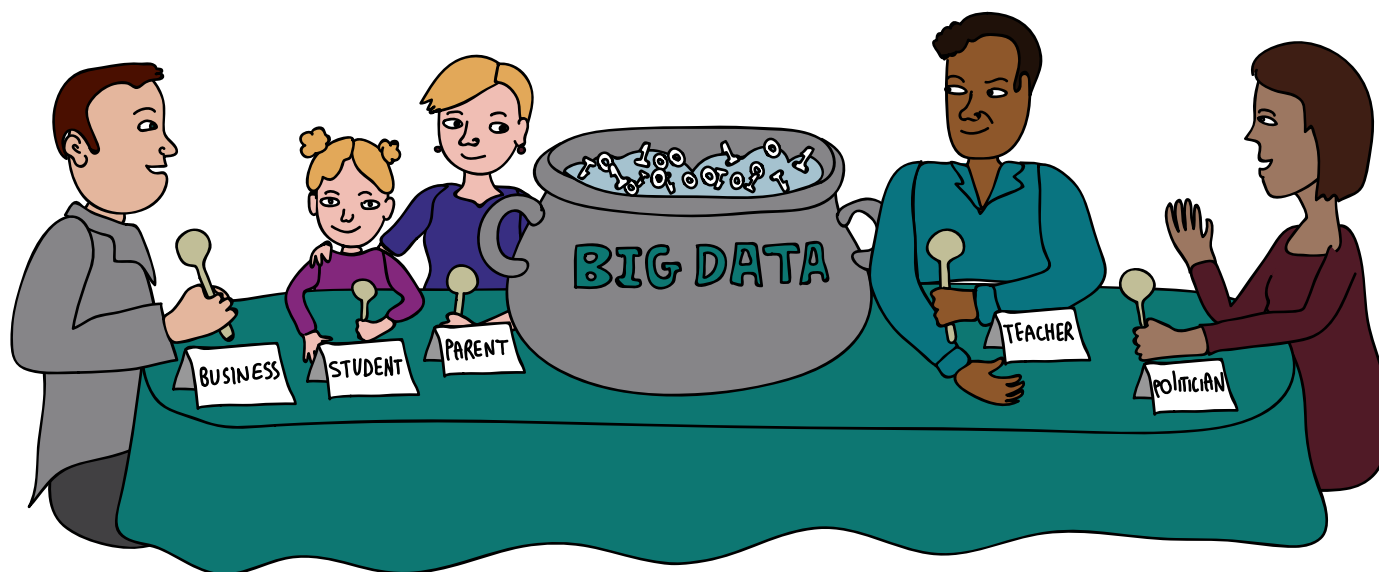
1. <http://www.wired.com/insights/2013/08/why-big-is-blinding-us-to-the-real-value-of-big-data/>
2. [Undefined By Data: A Survey of Big Data Definitions. Jonathan Ward, Adam Baker, Sept 2013.](#)
3. <http://openthoughtsmarter.blogs.uoc.edu/rethinking-the-approach/>
4. <http://smartdatacollective.com/bernardmarr/141351/what-really-big-data-and-why-it-will-change-world#!>
5. <http://smartdatacollective.com/bernardmarr/141351/what-really-big-data-and-why-it-will-change-world#!>
6. <http://www.wired.com/wiredscience/2012/10/big-data-is-transforming-healthcare/>
7. <http://www.youtube.com/watch?v=CO2mGny6fFs>
8. [The Intersection of Big Data and Leadership: Lessons from Sir Terry Leahy, Stern Speakers, 10 Dec 2013.](#)
9. <http://www.ecampusnews.com/featured/featured-on-ecampus-news/big-data-bang-344/2/>
10. [MetaGroup, 3D data management: Controlling data volume, variety and velocity. 2001](#)
11. <http://www.technologyreview.com/view/519851/the-big-data-conundrum-how-to-define-it/>
12. <http://www.technologyreview.com/view/519851/the-big-data-conundrum-how-to-define-it/>
13. <http://thenewinquiry.com/blogs/marginal-utility/dumb-bullshit/>
14. <http://www.wired.com/wiredscience/2012/10/big-data-is-transforming-healthcare/>
15. <http://dmlcentral.net/blog/lyndsay-grant/understanding-education-through-big-data>
16. [Delatorre, Christopher. "Chasing Innovation—On data, disciplines, and ditching the rules." Urbanmolecule., 20 Jul. 2013.](#)
17. <http://kogodnow.com/2013/03/big-data-ignites-revolution-in-decision-making/>
18. <http://www.wired.com/wiredscience/2012/10/big-data-is-transforming-healthcare/>
19. [McKinsey Report 2013: Open Data: Unlocking Innovation and Performance with Liquid Information](#)
20. <http://openthoughtsmarter.blogs.uoc.edu/rethinking-the-approach/>
21. <http://www.hastac.org/blogs/slgrant/2013/01/15/socializing-big-data-collaborative-opportunities-computer-science-social-sc>
22. [McKinsey Report 2013: Open Data: Unlocking Innovation and Performance with Liquid Information](#)
23. <http://www.wired.com/wiredscience/2013/10/big-data-science/all/>
24. <http://www.forbes.com/pictures/lmm45emkh/7-hod-lipson-and-michael-schmidt-computer-scientists-cornell-university/>
25. <http://www.wired.com/wiredscience/2013/10/topology-data-sets/all/>
26. <https://www.simonsfoundation.org/quanta/20131009-the-future-fabric-of-data-analysis/>
27. [EdTech Powered by Big Data. Report by Astra. 2013](#)
28. [BIG DATA: A REVOLUTION THAT WILL TRANSFORM THE WAY WE LIVE, WORK AND THINK, book by Shonberger and Cukier, 2013](#)
29. http://www.ted.com/speakers/hans_rosling.html
30. http://www.ted.com/talks/deb_roy_the_birth_of_a_word.html
31. <http://www.wired.com/wiredscience/2013/10/topology-data-sets/all/>
32. <http://www.wired.com/wiredscience/2013/10/big-data-science/all/>
33. <http://www.fastcompany.com/1834177/content-curators-are-new-superheros-web>
34. <http://www.wired.com/wiredscience/2013/10/topology-data-sets/all/>
35. <http://www.wired.com/wiredscience/2013/10/computers-big-data/>
36. <https://www.simonsfoundation.org/quanta/20131009-the-future-fabric-of-data-analysis/>
37. <http://www.campuscomputing.net/item/2013-campus-computing-survey-0>

חינוך וביג דטה

חינוך וביג דטה



אנו מזמינים אתכם לדמיין לרגע מערכת חינוך שבה כל
השחקנים - תלמידים, מורים, הורים, פוליטיקאים, מוציאים
לאור, מפתחים, חוקרים - כולם משתתפים פעילים, כלומר
כולם לא רק משתמשים. האם תהיה זו מערכת יעילה ורלוונטית
יותר? האם ביג דטה מאפשר מערכת כזו?



חברתיים. גופים אלה ביחד יכולים להביא את החדשנות ואת המשאבים כדי

ליצור מציאות של למידה לכולם הנמשכת כל החיים" (אנדראס שליישר, יועץ

מיוחד לענייני מדיניות חינוך של הארגון לשיתוף פעולה ולפיתוח כלכלי).² גל הביג דטה מגיע לאטו למערכת החינוך, במיוחד דרך מיזמים המציעים מגוון של פתרונות למידה ופתרונות מערכתיים רבים ושונים. התפתחותן של מערכות למידה מותאמות אישית היא הנושא החם כיום. אנשי חינוך טוענים כי מערכות אלו עשויות להיות המפתח להשגת מעורבות של התלמיד בלמידה, ויותר מכך - ליצירת מערכת היכולה לענות על צורכי הלמידה האמיתיים של כל תלמיד. יתרה מזאת, אפשר לבנות מאגרי נתונים של ביג דטה שיספקו מידע לכל קהילת החינוך, ומהלך כזה עשוי להוביל לשיתוף פעולה יעיל בין החברים בקהילה. קובעי מדיניות מאמינים כי יוכלו בחלק מהמקרים לבסס את החלטותיהם בנושא חינוך על נתונים הנאספים ממקומות רלוונטיים רבים, במקביל ובזמן אמת. שחקנים גדולים וותיקים בתחום החינוך משקיעים סכומי עתק בביג דטה. הוצאות

מייסד מעבדת המדיה של MIT, ניקולס נגרופונטה, ראה במחשב אמצעי להעצמת התקשורת בין אנשים למכונות - "המפתח הוא להיות מחובר". ג'וי איטו, מנהל מעבדת המדיה הנוכחי, ממשיך חזון זה באמצעות מחקר והרחבת האפשרויות הגלומות בביג דטה לעולם החינוך. איטו מחפש "עולם שבו 7 מיליארד מורים, שבו המוני משתמשים חכמים מאמצים גישה גמישה ורוח מרדנית ופותרים כמה מבעיותיו המרכזיות של העולם; עולם של רשתות ומערכות גומלין, שבו יצירתיות פורצת גבולות יכולה להתמודד עם כל דבר, החל מתמותת תינוקות וכלה בשינויי האקלים. אנו רוצים לקחת את ה-DNA (של המעבדה), את המרכיב הסודי, ולהטמיעו בקהילות, בחברות ובמשלוח. זה הייעוד שלי, הייעוד שלנו - להפיץ את ה-DNA הזה. אי אפשר לדרוש מאנשים לחשוב בעצמם או להיות יצירתיים. עלינו לעבוד עמם ולגרום להם ללמוד זאת בעצמם."³

"ביג דטה הוא היסוד שעליו יכול עולם החינוך להמציא מחדש את המודל העסקי שלו - ברור? ולבנות קואליציה של ממשלות, של ארגונים ושל יזמים

לאור בולטות בתחום החינוך כגון McGraw-Hill ו-Pearson מפנות את מאמציהן לפלטפורמות מקוונות דינמיות, היכולות לאסוף נתונים מתלמידים המשתמשים בהן ולספק תגובות מותאמות אישית. לאחרונה פועלות ההוצאות לאור האלה בשיתוף פעולה עם חברות צעירות אחרות בתחום הלמידה האדפטיבית (Adaptive learning) כגון Knewton³ ו-Aleks⁴. ספקי תוכנת תשתית, כגון Blackboard⁵, המיועדת למגוון גורמים במערכת החינוך, ו-Ellucian⁶, הפועלת בתחום ההשכלה הגבוהה, מבססים את מערכותיהם על כלים לניתוח נתונים. כולם מנסים לחזות הצלחה בקרב תלמידים על בסיס הנתונים הנרשמים/הנאספים על ידי מערכות התוכנה של לקוחותיהם. קרנות כמו קרן ביל ומלינדה גייטס⁷ מקדמות שימוש בביג דטה למדידה ולשיפור תוצאות הלמידה של תלמידים. הם השקיעו 100 מיליון דולר בחברה ללא כוונת רווח בשם inBloom, שמטרתה להשתמש בניתוח הנתונים על מנת לקדם למידה פרסונלית ואדפטיבית⁸ (לאחרונה החברה נסגרה עקב ביקורת על שמירת פרטיותם של התלמידים).

אנשי חינוך רבים טוענים כי יש ליישם במערכת החינוך יתרונות של תחומים אחרים, שבאים לידי ביטוי בשטח. "החנות הקמעונאית הממוצעת יודעת יותר על קופסת דגני הבוקר שעל המדפים שלה מאשר אנו יודעים על התלמידים שלנו. תכנון מחקר הבודק מה גורם לאדם לקנות מוצר כלשהו, באיזו מידה אתה רוצה שהוא יקנה אותו, ואילו מערכי פרסום תציג לו, אינו שונה באופן מהותי מתכנון תוצאות או הישגי למידה."⁹ (פיל איים, American Public University System).

ניתוח נתוני למידה וכריית נתונים

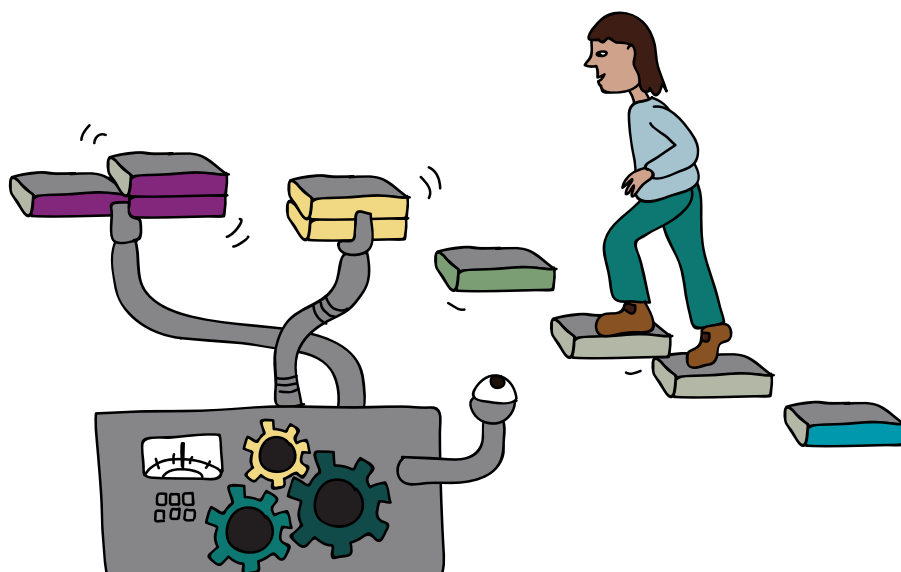
חינוך

והקשר ממלאים גם הם תפקיד חשוב בחקר נתוני חינוך. "חברות סטארט אפ, כגון inBloom ו-Knewton, מציעות שירותים לקיבוץ נתונים קיימים ממגוון מקורות, וכתוצר נלווה - נתונים המופקים מהשימוש של התלמידים בטכנולוגיה. נתונים יחידניים מנותחים יחד עם נתונים ממאות אלפי תלמידים ויוצרים פרופילי למידה, המאבחנים את החוזקות ואת החולשות, משקפים אתגרים וכן מציעים מסלולי למידה מותאמים אישית.¹²

החינוך מנסה להשלים את הפער בתחום מדעי הנתונים שכבר קיים בתחומים אחרים, באמצעות ניתוח נתוני למידה (Learning Analytics) וכריית (mining) של נתונים: תהליכים של יצירת ידע יישומי מכמויות עצומות של נתונים. דו"ח הורייזון לשנת 2013¹³ מתאר ניתוח של נתוני למידה כ"פענוח מגמות ודפוסים הקשורים לביג דטה בנושא חינוך ולמידה, או מקבצים עצומים של נתונים הנוגעים לתלמידים, כדי לקדם מערכת חינוך מותאמת אישית ותומכת". הרעיון הבסיסי העומד ביסוד ניתוח נתוני למידה הוא שימוש בניתוחי נתונים כדי להתאים את ההוראה לצרכיו האישיים של הלומד בזמן אמת, כשם ש-Amazon, Netflix ו-Google משתמשות בשיטות מדידה כדי לשלוח לצרכנים המלצות ופרסומות בהתאמה אישית. ניתוח של נתוני הלמידה מאפשר לחזות את ביצועיו העתידיים של התלמיד (על בסיס דפוסים קודמים של הלמידה שלו במערכות למידה שונות), להמליץ המלצות ולתת משוב המותאם לתשובות התלמיד, להתאים לו את אפשרויות הלמידה ואת סגנונות ההוראה והלמידה. כריית נתונים של תלמידים ודרכי הלמידה שלהם, המאפשרים ניתוח של נתוני הלמידה, מיועדת לפיתוח דרכים לגילוי דפוסים הקיימים בנתונים, בניסיון לזהות תופעות חינוכיות מעניינות. חוקרים בתחום זה מחפשים יישומים המסייעים ללומדים, מוסיפים ידע ומעשירים את מדעי הלמידה.¹⁴ ההגדרה על פי "האגודה הבינלאומית לכריית נתוני חינוך" היא "דיסציפלינה חדשה הנוגעת לפיתוח שיטות לחקר סוגים ייחודיים של נתונים, המגיעים מסביבות חינוכיות, ושימוש בשיטות אלה כדי להבין טוב יותר את התלמידים ואת סביבות הלמידה שלהם. בין אם נתוני החינוך נלקחים מהשימוש של התלמידים בסביבות למידה אינטראקטיביות, בלמידה שיתופית בסיוע מחשב ובין אם אלה נתונים מנהלתיים מבתי ספר ואוניברסיטאות - יש להם לעתים קרובות כמה רמות היררכיה בעלות משמעות, אשר יש להגדירן על פי תכונות הנתונים עצמם. סוגיות של זמן, רצף

הלמידה שלהם. חידוש מלהיב זה יכול להיות "גורם משווה" המעניק הזדמנויות רבות יותר לתלמידים בחברה שלנו, בלי קשר לרקע או למקום שממנו באו.¹³

[טום ואנדר ארק, מנכ"ל Getting Smart].¹⁴ Dreambox Learning מציעה סביבת למידה אדפטיבית ללימוד מתמטיקה בבתי הספר היסודיים. התלמידים משקפים את תהליך העבודה והחשיבה שלהם בתוך סביבת הרפתקאות הדומה למשחק, ותוך כדי המשחק, המערכת מעודדת אותם להסביר, לדון ולהגן על החשיבה המתמטית שלהם. Dreambox טוענת, כי פיתחה מערכת למידה אדפטיבית חכמה, דור חדש של טכנולוגיה חינוכית, המאפשרת חוויות למידה חדשות, מתאימה את המסלול והקצב ללמידה



למידה מותאמת אישית ואדפטיבית

התפיסה של למידה מותאמת אישית חדרה לשוק ה-EdTech (Education Technology - טכנולוגיה בחינוך) בעוצמה רבה, והיא מאלצת את ספקי ההוראה והתוכן המסורתיים לחפש דרכים להתמקד בחוויית התלמיד. חברות סטארט אפ רבות מפתחות אמצעים לשימוש בניתוח נתונים כדי לאפשר מערכת חינוך מותאמת אישית לכל תלמיד.

כיום, המערכות העובדות על פיתוח למידה מותאמת אישית מתבססות על כמויות עצומות של נתונים הנאספים כתוצאה מפעילות תלמידים בזמן שהם פועלים בסביבות למידה מקוונות. מערכות אלו מאתרות מה הם יודעים, וכיצד הם לומדים בצורה הטובה ביותר. המערכות יכולות לנתח את הנתונים הללו ולהמליץ מיד מה צריך התלמיד ללמוד הלאה ובאיזה אופן.

המונח למידה אדפטיבית (מסתגלת) נפוץ כאשר מדברים על תכניות המציעות תוכן שונה לכל לומד, על בסיס הערכת הידע שלו (Edsurge). למידה אדפטיבית, בשונה מלמידה מותאמת אישית, דורשת תהליך מתמשך של למידה ואספקה קבועה של נתונים חדשים כדי שהמערכת תתאים את עצמה לצרכיו וליכולותיו של התלמיד. פלטפורמות של למידה אדפטיבית אוספות ללא הרף נתונים מהתלמיד, וכך הן יכולות ללמוד ולהתאים את עצמן למסלול הלמידה שלו, המשתנה ומשתפר במשך הזמן. לעומת זאת למידה מותאמת אישית יכולה לכלול גם מערכות הנגזרות משיטה המבוססת על כללים של עצי החלטות, ושיטה זו מובילה למסלולים קבועים מראש.

למידה אדפטיבית מספקת לתלמיד סביבות למידה מודולריות, ומשמעות הדבר היא שתוכנית הלימודים הופכת להיות יחידנית. כל תלמיד רואה תוכנית לימודים שונה, המותאמת ליכולות הלמידה והקצב שלו, והיא מציעה את השלב

הבא בלמידה של התלמיד על פי מה שלמדה עליו מהנתונים שאספה. "על ידי

כיוול מחדש של כל אינטראקציה במטרה לשמור על אתגרים הולמים, הלומדים

נשארים באזור הלמידה האופטימלי שלהם ויכולים להשיג את מלוא פוטנציאל

נתוני חינוך

כיום, כאשר הלמידה הופכת ליותר ויותר מקוונת, אנו יכולים לקבל ממקורות שונים תיעוד מפורט ומדויק של פעילות הלמידה של הלומד. מקורות אלו כוללים רישומי תלמידים, מערכות ניהול למידה (LMS), לומדות שונות המפורסמות ומשותפות באינטרנט, ועולם שלם של נתונים בעלי תוכן חינוכי הזמינים ברשת. הנתונים מנוהלים בכל הרמות (אישית, בית ספרית, מחוזית, ארצית), במערכות שונות ורבות, ובכל הצורות (מובנות ולא מובנות באמצעות טקסטים, תמונות, סרטוני וידאו, פעולות אינטראקטיביות וכדומה). **האתגר העיקרי הוא לשלב יחד את החלקים השונים של הנתונים כדי ליצור תמונה ברורה.** "כיום, סוגים רבים של נתונים על התלמיד נשמרים במקומות מגוונים, החל מתוכנות הוראה שונות, וכלה בגיליונות ציונים ובמערכות ניהול למידה. אולם, הנתונים לכודים לעתים קרובות בתוך התוכנה או המערכת ואי אפשר לחלצם או לשלבם בקלות עם נתונים אחרים על אותו התלמיד. כדברי השיר Hotel California: הנתונים יכולים להיכנס בכל עת שירצו, אולם הם אינם יכולים לעזוב לעולם, והמורים אינם יכולים להשתמש בהם ביעילות (פרנק קטלנ, אנליסט Edtech).¹⁶ רוב מערכות התוכנה הוותיקות בתחום החינוך נבנו תוך התחשבות מועטה בניידות נתונים. **כיום, מיזמים רבים מתמקדים ביצירת שפה משותפת או אוצר מילים ומבנה משותף כדי לאפשר שיתוף נתונים ממערכות ומיישומים שונים.** חברות רבות מתחילות לפעול בתחום פיתוח של מערכות המאפשרות איסוף וצבירה של נתוני תלמידים לכלים אנליטיים, ואת הכלים האלה יהיה אפשר למכור חזרה לבית הספר.

האופטימלי של הילד, ועוזרת להאיץ את ההבנה ואת החשיבה הביקורתית. המערכת מספקת גם נתונים מסכמים למורה, כדי לאפשר חוויה מותאמת אישית גם בכיתה.

"כאשר אנו מדברים על למידה אדפטיבית, אנו מתכוונים למערכת המשתנה כל הזמן, כזו המגיבה בזמן אמת לביצועיו ולפעולותיו של כל אדם במערכת. מערכת כזו מגדילה למקסימום את הסיכוי שהתלמיד יבין רעיון מסוים, כיוון שהיא ממליצה על צורת ההוראה המתאימה לאותו תלמיד, לנושא הנלמד בזמן

המתאים" (דוויד ליו, סמנכ"ל התפעול של Knewton). חברת Knewton, אחת החברות הבולטות בתחום, משתמשת בביג דטה כדי לפתח מערכות של למידה אדפטיבית וניתוח נתונים לתלמידים, למורים, למחוזות ולמוציאים לאור. ניתוח הנתונים נועד למפות את החוזקות ואת החולשות של התלמיד לאורך זמן, כדי לאפשר למורים להתאים לו אישית את ההוראה ואת התוכן. "Knewton מתאימה אישית קורסים דיגיטליים, כך שכל תלמיד יהיה מעורב, ואף אחד לא ייפול בין הכיסאות"¹⁵. Knewton מאפשרת להתאים אישית את התכנים ולשלב כלים לניהול הכיתה, התרעות על התערבויות נדרשות והמלצות כיצד ליצור צוותי עבודה הומוגניים. בכך היא מספקת למורה ראייה כוללת על הכיתה.

חברה חשובה נוספת היא inBloom. היא מספקת למדינות ולמחוזות ניתוח נתונים של תלמידים ומורים, כדי לאפשר למורים להתאים אישית את אופן ההוראה שלהם, וכדי לסייע למחוזות או למדינות לאתר חולשות ובעיות במערכת החינוך.

מה צריך להיות ההיקף של נתוני

החינוך?

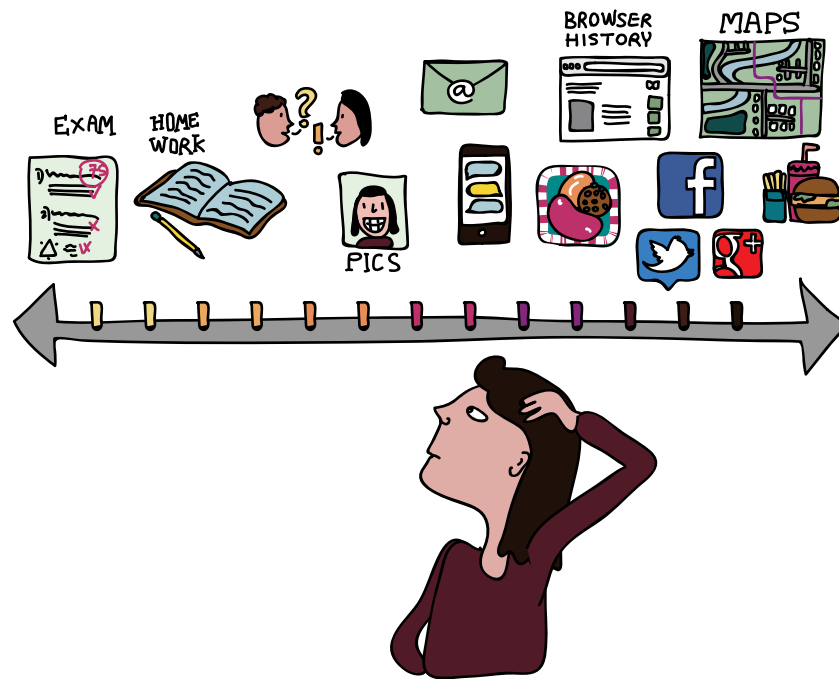
כמה מידע דרוש למערכת כדי להבין את ביצועיו האקדמיים של תלמיד? המערכות מתחילות לכלול את פעילויות היומיום של התלמידים כחלק מהנתונים הנאספים על הלמידה שלהם: הוצאת ספרים מספריות, השתתפות בספורט באולמות סגורים, רכישות בחנויות ספרים, פרוטוקולים משיבות מועצת התלמידים, פעילות במערכות ניהול למידה, היסטוריה של בלוגים ותגובות בקבוצות דיון, שימוש באינטרנט בזמן השהות בבית הספר, דפים שתלמידים קוראים בספרי לימוד דיגיטליים, הקטעים שהם מדגישים, פרופילי המדיה החברתית שלהם, סרטוני הווידאו שבהם הם צופים בקורסים מקוונים, כניסות

ל- Wikipedia וכדומה.¹⁷

"האם זוהי האחריות שלנו לנטר את אתרי המדיה החברתית כדי לעזור להגן על התלמידים מפני סכנות של בריונות, שימוש בסמים, אלימות והתאבדות?" שואל מנהל חטיבת ביניים, בדיון על חדירת המדיה החברתית לבית הספר ולסביבות

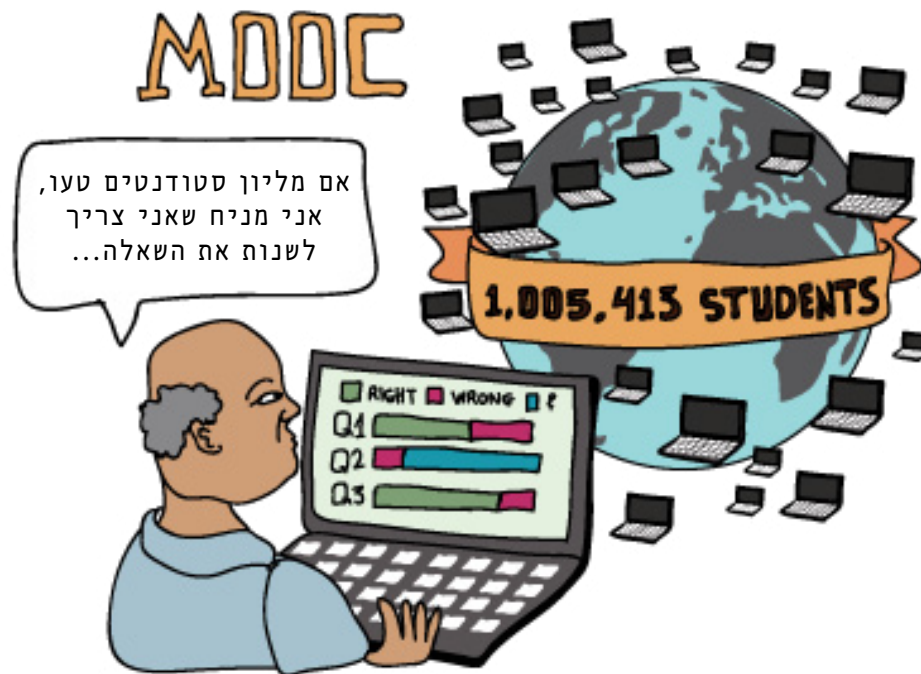
הלמידה. השימוש שעושות פלטפורמות חינוכיות ברשתות חברתיות מעורר

דילמה על גבולות האיסוף והשימוש בנתונים על ידי מערכת החינוך.¹⁸



איסוף מסיבי של נחונים

את מעורבות המשתתפים. מתודולוגיה זו של למידה מבוססת נתונים אפשרית רק כאשר יש מספר גדול של משתמשים - מאות או אלפי משתמשים. "נתונים ממשתמשים רבים יאפשרו למורים ולמנחים לקבל משוב שניתן ליישום, שמציג ומדגים להם מה עובד או לא עובד" (סלמן קאהן, מייסד Khan Academy).²⁰



שוק ה- MOOC (קורסים מקוונים הפתוחים לקהל הרחב) רשם עלייה עצומה בשנתיים האחרונות, וספקי הקורסים המקוונים העיקריים הם Coursera, Khan Academy ו-EdX, Udacity. מטרתם של הקורסים היא הנגשה של קורסים והשתתפות בלתי מוגבלת בהם באמצעות האינטרנט. הם מציגים הרצאות וידאו קצרות המשולבות עם כלי הערכה המותאמים לכמות עצומה של משתתפים (יותר מ-100 אלף סטודנטים בקורסים הפופולריים), כגון: משוב אוטומטי באמצעות הערכות מקוונות (לדוגמה, בחנים ובחינות) והערכות עמיתים ופעילויות קבוצתיות. ב- Coursera, לדוגמה, אוספים נתונים מכל פעולה (או היעדר פעולה) של הסטודנט - כאשר הוא עוצר את הסרטון, מאיץ את מהירות ההשמעה, עונה על שאלות הבוחן, חוזר שוב על מטלה, מגיב בקבוצת דיון וכדומה. רמה מיקרוסקופית כזו של נתונים, הנאספים בתפוצה שבה פועלים קורסי MOOC, מקלים על זיהוי פגמים במערכת. כפי שמציינת דפני קולר,

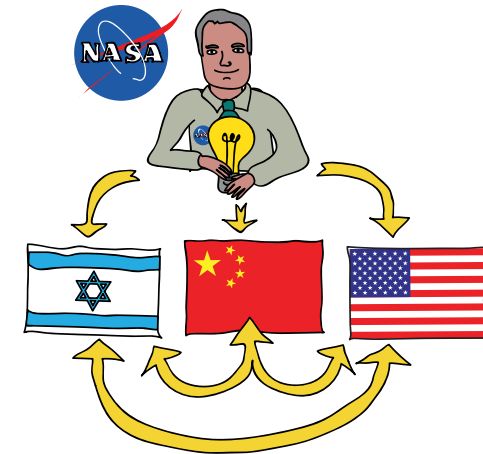
מייסדת שותפה ב- Coursera: "אם שני סטודנטים בקורס באוניברסיטה, המונה מאה איש, כותבים תשובה לא נכונה, לא היינו שם לב. אבל כאשר אלפיים איש

כותבים אותה תשובה שגויה, קשה לפספס את זה."¹⁹

כמויות הנתונים המסיביות, הנוצרות על ידי 10,000 עד 100,000 סטודנטים שנרשמו לקורס, מאפשרות לחברות הבונות את הקורסים לשפרם. אם הסטודנטים עונים תשובה שגויה לשאלה במבחן, או אם הם מאבדים ריכוז בנקודה מסוימת בקורס, נתונים אלה יכולים לכוון את יוצרי הקורס לחזור לחומר הלימוד ולהוסיף בו או לשנות אותו. ספקי MOOC מנצלים את יתרון הגודל כדי לערוך ניסויים בחומרים הנלמדים בקורס, בשיטות הצגת החומר ובדרכי התקשורת עם הסטודנטים. לדוגמה, סבסטיאן תורן, מייסד Udacity, ערך השוואה בין גרסת שיעור צבעונית לבין גרסה של אותו שיעור בשחור לבן, ומצא כי "הגרסה בשחור לבן הייתה הרבה יותר טובה...זה הפתיע אותי". אנדרו נג ערך בדיקות השוואתיות ב- Coursera כדי לבחון אם תזכורות הנשלחות בדוא"ל מגבירות

חילופי נתונים בין מחנכים

הקשר שלנו עם עולם העסקים התפתח ממתן ארון באנשים שיספקו לנו מידע, דרך מסירה מרצון של נתוני כרטיסי אשראי, ועד להתקשרות עם זרים שאנו מאמינים שאפשר לבטוח בהם בתחומי מסחר שונים. בתחום החינוך לעומת זאת, לא נראית מגמה דומה. פרויקט "MyPISA" מנסה לשנות זאת באמצעות גיבוש צוות מחנכים אשר מחליפים ומשתפים מידע באופן פעיל, או בניסוחם: "ביג דטה בונה ארון ובגדול".²¹ מנהלי בתי ספר ומורים מתחילים לראות את עצמם כחברי צוות - ולא רק צופים - במגרש המשחקים העולמי של עידן המידע.



שיתוף נתוני מחקר עם מחנכים

ותלמידים

ה-NASA ו-AWS (Amazon Web Services) מעמידים לרשות משתמשי מחקר וחינוך אוסף גדול של נתוני לוויין של נאס"א בנושא מדעי האקלים וכדור הארץ באמצעות ענן ה-AWS. המערכת מעודדת מחקר והזדמנויות לימודיות על ידי קידום מחקר מבוסס קהילה, חדשנות ושיתוף פעולה. "נאס"א ממשיכה לתמוך ולספק גישה לציבור הרחב לחקר נתונים, ושיתוף הפעולה עם ארון עולה בקנה אחד עם מטרה זו, אומר המדען הראשי של נאס"א, אלן סטופן.²² השימוש בענן מאפשר לחוקרים ולמפתחים מכל העולם לגשת למערכת ניהול נתונים וחישוב, המשולבת בתחום מדעי כדור הארץ, ולהשתמש בנתונים. "אנו שמחים להצמיח מערכת גומלין של חוקרים ושל מפתחים שיכולים לעזור לנו לפתור בעיות סביבתיות חשובות", אומר ראמה נמאני, מדען ראשי בפרויקט NEX. "מטרתנו היא שאנשים יקבלו גישה וישתמשו במגוון שירותי ניתוח הנתונים דרך AWS כדי להוסיף ידע וכלי קוד פתוח למען הזולת."

אחסון נתונים



ארגונים שונים מנסים כיום לפתח מיזמים ושיתופי פעולה ליצירת מאגרי נתונים משותפים וסטנדרטיזציה של מערכות האוספות, מנהלות ומשלבות נתונים מעולם החינוך והלמידה. חברת ²³Schoolzila מציעה שירותי אחסון לשמירת כל הנתונים של בית ספר או מחוז שיש לו מערכות קוד רלוונטיות (מערכות מידע על תלמידים, מערכות ניהול משאבי אנוש, סקרים, מערכות הערכה וכדומה), וכן כלי דיווח ומחקר.

במארס 2013, הפכה inBloom לחברה האחראית לאחסון נתונים המכילים קבצים של מיליוני תלמידים במערכת החינוך הציבורי בארצות הברית. זה היה שיתוף פעולה בין קרן ביל ומלינדה גייטס, Carnegie Corporation of New York וגורמים רשמיים מבתי ספר במדינות שונות. inBloom מפתחת פורטלים המאפשרים ניטור ושימוש בנתונים אלו למגוון מטרות (NHM : K-12 edition) (Horizon Report 2013).

ויזואליזציה: התמקדות בחוויית



המשתמש



חלק מהשפעת השימוש ההולך וגובר בביג דטה על המשתמש היא הביטוי הוויזואלי שלו. חברות המספקות מוצרים ושירותים בתחום החינוך מודעות יותר ויותר לחשיבות ההתמקדות בחוויית המשתמש (user experience-UX), והן מפתחות כלים ויזואליים ידידותיים וקלים להבנה. ניתוח נתונים מעניק תובנות לכל מערכת החינוך, מתלמידים ועד לקובעי המדיניות.

מחשבה ומשאבים רבים הושקעו בפיתוח ויזואליזציה אופטימלית כתכונה מרכזית של מוצרים בתחום החינוך. יכולתו של מורה לראות בצורה ברורה על מסך אחד, בזמן אמת, מה קורה עם כל ילד בכיתה, עשויה לשנות משמעותית את ביצועי המורה. מפות הלמידה של Knewton, לדוגמה, מדגימות את השימוש הייחודי של התלמיד בתכנים שונים כדי להשיג יעד למידה. Blackboard מציעה לוחות מחוונים אינטראקטיביים לניטור ולניתוח של פעילויות באוניברסיטאות או במכללות.

אמצעים למניעה ולחיזוי

עלולים ליצור "עתיד דיסטופי"²⁶, בדומה לסרט מדע בדיוני שבו אדם נידון על פשעים שעדיין לא ביצע. האמונה בנתונים אמפיריים כאמת מוחלטת עשויה לגרום למוסדות להתעלם מהיבטים אישיים ואיכותיים של תלמידים, היבטים שיש להם ערך רב, מדיד פחות.

היכולת לחזות את תוצאות ביצועי התלמידים יכולה להיות משאב יקר ערך, מאחר שהיא מאפשרת למערכת להקצות משאבים כדי למקסם את יכולות התלמידים, למנוע כישלונות מערכתיים, ולספק מידע על האפשרויות האקדמיות העתידיות של התלמידים. מנכ"ל Blackboard Analytics Services, ג'ים הרמנו, מדגים כיצד גישה של מחנכים לניתוח נכון של נתונים יכולה לסייע למנוע נשירה של תלמידים.

"איסוף מידע על תלמידים לפני שהם עושים בגרויות, ושימוש במידע זה כדי לתכנן את הצלחתם הכללית לאחר מכן, הוכחו כטקטיקה

יעילה". בשנת 2013 חברת Blackboard²⁴ הכניסה למערכת ניהול הלמידה שלה כלי המכונה "Retention Center", המיועד להקל על מחנכים לזהות במהירות תלמידים מתקשים. המערכת מתבססת על תוצאות מחקר אשר זיהה ארבעה אינדיקטורים חשובים: היסטוריית הכניסה (login) של התלמיד למערכת, הפעילות שלו בקורס, איחורים במועדי ההגשה וירידה בציונים.

בשל הנגישות לנתונים והאפשרויות לנתחם, אנו מתחילים לראות גל של מחקרים ומוצרים המנסים לחזות את הביצועים ואת ההתנהגות של התלמידים. יש מוסדות חינוך שכבר משתמשים בשיטות אלה כדי לזהות תלמידים הנמצאים בסיכון מוקדם ככל האפשר. EdWeek.org פרסם לאחרונה מקרה של מחנכים ממרילנד שמצאו כי אפשר לזהות סימנים מוקדמים המתריעים על תלמיד הנמצא בסיכון לנשירה, כבר בשלב מוקדם של לימודיו בבית הספר. מחוז מונטגומרי, מחוז אמיד ועתיר טכנולוגיה, המונה 149,000 תלמידים ושוכן בפרבר של וושינגטון, בונה את אחת ממערכות ההתרעה המוקדמת הראשונות בארצות הברית. מערכת זו יכולה לאתר ולהתריע על 75% ממקרי הנשירה העתידיים כבר במחצית השנייה

של כיתה א'.²⁵

עם זאת, יש להביא בחשבון גם את החסרונות של דרך פעולה כזו. פרופסור ויקטור מאייר-שונברגר מאוניברסיטת אוקספורד מזהיר, כי מוסדות חינוך המייחסים משקל יתר לביג דטה כדי לנבא ביצועים אקדמיים של תלמידים

מגינים על התלמידים או משחקים

תפקיד של אלוהים?

את לא צריכה להגיע יותר
לבית הספר. מהנתונים
שלך, ניתן לראות שאין לך
עתיד אקדמי.



המידע הרב והנתונים האישיים הרבים, שנעשים נגישים לכל דורש בשנים האחרונות, מחדדים את הסוגיה כי יש לשקול בכובד ראש את תרומת הביג דטה להבנה טובה יותר של הישגי התלמידים וקשייהם לעומת הסקת מסקנות נמהרות. דילמה זאת נובעת בין היתר מן ההבנה כי אפשר להגיע לתוצאות ולקבל החלטות על עתיד התלמידים על סמך הבנה מוטעית של הנתונים. יש

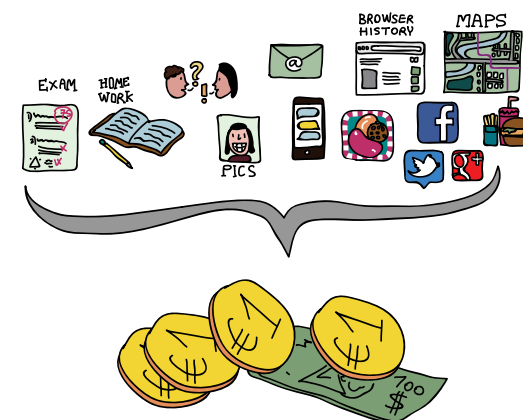
להבין כי תחזיות הן רק ציפיות, גם אם הן מושתתות על נתונים, וזה חשוב במיוחד כאשר אנו מדברים על התנהגות אנושית.

"כאשר זהות הלומד היא משהו שתלמידים מגדירים ביחסיהם עם מורים ועמיתים, יש להם אפשרות בחירה בקביעת איזה סוג של לומד הם, ואיזה סוג לומד הם רוצים להיות. הם יכולים לספק את ההקשר שמכניס היגיון לנתונים שלהם. הם יכולים לקרוא תיגר או להתנגד לפרשנויות של אחרים לפעולות ולמניעים שלהם. כלומר, יש להם שליטה מסוימת ואמירה בנוגע למה שהם רוצים להיות בתור לומדים... עלינו לשקול את ההשלכות ואת תוצאות השימוש בניתוח ביג דטה כדרך העיקרית להשגת מידע על תהליכי למידה. ביג דטה נוטה לפשט שאלות חברתיות ופוליטיות גדולות הנוגעות לסוג הלומדים שאנו רוצים להיות ולתגובת החינוך לאתגרים חברתיים וכלכליים מרכזיים, והוא הופך אותם לתהליך פשוט המכתיב את החלק הבא של תוכנת הלימוד שיש להוריד".²⁷

מבקרי השימוש בביג דטה במערכת החינוך חוששים כי למידה מושתתת נתונים תהפוך את בתי הספר ל"מפעלים", בין היתר בשל ריבוי ההסכמים עם חברות מסחריות. חברות אלה משווקות מוצרים המתבססים על ניתוח ואיסוף נתונים לבתי הספר.

יכולתה של הטכנולוגיה לזהות חוזקות וחולשות של תלמידים ולהתאים את החומרים ואת אופן הלימוד לכל לומד, מוטלת בספק. הביקורת מדברת על שימת משקל יתר לנתונים במקום השקעה רבה יותר בגורם האנושי.²⁸

"ייתכן שגם תפיסתם העצמית של ילדים כלומדים נשלטת על ידי הוויזואליזציה של נתוני הלמידה שלהם (באמצעות אפליקציות, פרופיל אינטרנטי ומידע המוצג בצורה גרפית), יותר מאשר על ידי תהליכי חשיבה ודו-שיח. הפתגם העתיק "דע את עצמך" הופך ל"מדוד את עצמך". אם אמינות הידע שלנו נשענת על ביג דטה, פרופילי הלמידה שלנו יכולים להיתפס - על ידי אחרים ועל ידינו - כתיאורים מוצקים ואובייקטיביים יותר של מי אנחנו "באמת", ואגב כך תקטן חשיבותו של הידע העצמי, הסובייקטיבי וחסר הארגון שלנו."²⁹



חברות אינטרנט בולטות וספקי שירותי תקשורת דנים בפתיחות כבר היום על "מסחר" בנתוני פרופיל המשתמשים שבבעלותם ובניהולם. ב' שיר מאתר האינטרנט Innovation Insights מזהיר מתוצרי הלוואי המסחריים של תוכנות חינוכיות. " מתוך כלל הכנסותיה של Google לשנת 2011, 37.9 מיליארד דולר, 96% נבעו מפרסומות. האם Google מספקת לבתי הספר גישה חופשית ל - Google Apps for Education software (אפליקציות Google לתוכנות חינוכיות חנימיות), בתקווה לרווחים מכריית נתונים מתוך המטלות הדיגיטליות בבתי הספר של ילדינו והשימוש לצורכי פרסומות שהיא יכולה לעשות בנתונים אלו?"³⁰ בתחילת 2013 הפכה מסציוסטס למדינה הראשונה בארצות הברית שאסרה על חברות המספקות שירותי מחשוב ענן לעבד נתונים של תלמידים למטרות מסחריות.

פלטפורמות במימון חברות ללא כוונת רווח, כגון inBloom, צפויות להתחיל לחייב מחוזות על שימוש בתשתית שלהן משנת 2015 (2-5 דולר לתלמיד לשנה). בנוסף, ספקי אפליקציות אשר יישענו על התשתית ועל ענן הנתונים שלהם יחפשו את הדרך להרוויח גם הם. (חברת inbloom נסגרה לאחרונה עקב ביקורת על שמירת פרטיותם של תלמידים).

מי הבעלים של כל הנתונים הללו? חברות? בתי ספר? מורים? תלמידים? האם התלמידים יודעים אילו נתונים נאספים עליהם? כיצד נוכל לוודא שניתוח נתוני למידה וכריית הנתונים מיועדים להוסיף ערך לימודי ולא מסחרי? כיצד נוכל לוודא שבמרוץ להשגת תובנות מכל נתוני החינוך שאנו אוספים ושומרים עתה, התלמיד אינו רק אובייקט הניתוח? כיצד נוכל לוודא שלתלמידים יש סמכות ושליטה על הנתונים ועל הלמידה שלהם?³¹

שלהם. בכוונת הפרויקט ליצור מערכות המשלבות נתונים הנוגעים לאינטראקציה פיזית של התלמידים עם מחשב ונתונים נוספים ממערכת הלימוד, כדי ללמוד יחידות אקדמיות ספציפיות. על ידי שימוש בניתוח נתוני למידה, המערכת תספק מידע שיכול לעזור להתוות את הדרך שבה שיעור ממוחשב יגיב לכל תלמיד. ההתאמה האישית מבוססת על איסוף נתונים בזמן אמת ועל תגובות התלמיד כאשר הוא נמצא באינטראקציה עם השיעור.

מיזמים שונים מדרבנים את עולם החינוך להיות חדשני ולהשתמש בנתונים כדי לחולל שינויים פדגוגיים משמעותיים. קבוצות מחקר רבות בישראל מנסות להשפיע על מערכת החינוך על ידי חקר ההתפתחויות האחרונות בתחום

הניתוח והטכנולוגיה. לדוגמה: האפשרויות הגלומות בחיישנים העוקבים אחר תנועותיהם של התלמידים כאשר הם מקיימים אינטראקציה עם המכשיר; מודלים של בינה מלאכותית מתקדמים לפיתוח מערכות למידה חדשות הרלוונטיות לדרכים המקובלות היום לקיום אינטראקציה, תקשורת ושיתוף מידע. בניסיון לקחת צעד נוסף קדימה את הלמידה האדפטיבית ולבסס מערכת על הידע של הלומד, ולא רק על רמות הישגיו, ש' הרשקוביץ מ"המרכז לטכנולוגיה חינוכית" (מטח) וארנסט ליוביץ' חוקרים כיום מודלים של למידת מכונה, מתוך כוונה לפתח קורס מתמטיקה אינטראקטיבי שבו יהיה אפשר לשנות את ההוראה בהתאם לתשובותיו של התלמיד.

פרופסור קובי גל וצוותו מאוניברסיטת בן גוריון החלו לאחרונה לחקור שיטות ומודלים, הלקוחים מבינה מלאכותית ומדעי הלמידה גם יחד, על מנת להבין טוב יותר כיצד תהליך הלמידה מושפע משיתופיות ומדיה חברתית. פרויקט רב-תחומי זה מפתח טכנולוגיות לניתוח ולתמיכה בלמידה בשיתוף פעולה בסביבות עתירות טכנולוגיה, בתוך הכיתה ומחוצה לה, בדגש על מדיה חברתית הנמצאת בכל מקום (לדוגמה, אתרים חברתיים כגון Facebook ו-Wikipedia). הוא גם מעריך את יתרונות הלמידה הקבוצתית מקבוצות קטנות ועד קבוצות גדולות ואינטראקציות ארוכות טווח.

פרויקט מחקר של אינטל, בהובלת שחר שפיגלמן, מנסה לשלב את הטכניקות החדשות ביותר של מחשוב תפיסתי (ראייה - מעקב אחר תנועות עיניים או זיהוי מחוות תלת ממדי, דיבור, זיהוי רגשות וכדומה) עם למידת מערכות אדפטיביות. הפרויקט חוקר את השפעותיהן של טכנולוגיות חדשות על ניתוח ביג דטה, כדי לנסות ולהבין טוב יותר את ביצועי הלמידה של תלמידים ואת דפוסי הלמידה

1. <http://www.wired.co.uk/magazine/archive/2012/11/features/open-university?page=all>
2. http://www.huffingtonpost.com/andreas-schleicher/big-data-and-pisa_b_3633558.html
3. <http://www.knewton.com/>
4. <http://www.aleks.com/>
5. <http://uki.blackboard.com/sites/international/globalmaster/Platforms/>
6. <http://www.ellucian.com/>
7. <http://dmlcentral.net/blog/lyndsay-grant/understanding-education-through-big-data>
8. <https://www.inbloom.org/>
9. <http://1776dc.com/2013/12/13/how-big-data-is-changing-the-educational-frontier/>
10. New Media Consortium Horizon Report 2013.
11. <http://www.columbia.edu/~rsb2162/BakerSiemensHandbook2013.pdf>
12. <http://dmlcentral.net/blog/lyndsay-grant/understanding-education-through-big-data>
13. <http://www.dreambox.com/white-papers/the-future-of-learning>
14. <http://www.dreambox.com>
15. http://www.eltjam.com/big-data-and-adaptive-learning-in-elt-knewton-interview-part-1/?utm_source=linkedin&utm_medium=social&utm_content=3190976#%21
16. Frank Catalano, How Will Student Data Be Used? GeekWire, July 3, 2012.
17. <http://hackeducation.com/2013/10/17/student-data-is-the-new-oil/>
18. <http://www.eschoolnews.com/2013/12/23/schools-monitor-media-400/2/>
19. http://www.ted.com/talks/daphne_koller_what_we_re_learning_from_online_education.html
20. <http://www.skilledup.com/blog/mooc-data/>
21. http://www.huffingtonpost.com/andreas-schleicher/big-data-and-pisa_b_3633558.html
22. <http://www.nasa.gov/press/2013/november/nasa-brings-earth-science-big-data-to-the-cloud-with-amazon-web-services/#.UoNKIfIT5ZA>
23. <https://schoolzilla.org/>
24. Visualization: The simple way to simplify Big Data. Chris Taylor. Wired. 8.26.13
25. <http://1776dc.com/2013/12/13/how-big-data-is-changing-the-educational-frontier/>
26. Dropout Indicators Found for 1st Graders, By Sarah D. Sparks, edweek.org, 07/29/2013
27. <http://www.timeshighereducation.co.uk/news/big-data-could-create-dystopian-future-for-students/2010061.article>
28. <http://dmlcentral.net/blog/lyndsay-grant/understanding-education-through-big-data>
29. Scientific American August 2013
30. <http://dmlcentral.net/blog/lyndsay-grant/understanding-education-through-big-data>
31. <http://insights.wired.com/profiles/blogs/bill-to-ban-data-mining-of-student-email#axzz2oelHtmKm>
32. <http://hackeducation.com/2012/12/09/top-ed-tech-trends-of-2012-education-data-and-learning-analytics/>

נתונים חדשים: פרטיות ומודעות לסביבות מקוונות

נתונים חדשים: פרטיות ומודעות לסביבות מקוונות



בעולם החינוך. מאידך, היתרונות הרבים הגלומים במערכות למידה המבוססות על ביג דטה (כפי שנסקרו בפרקים הקודמים), יוצרים מצב מורכב ולא ברור לגבי השימוש בנתונים אלה.

בתחילת 2013 התחוללה סערה ציבורית נגד inBloom אשר הפכה לתביעה משפטית¹⁰ והביאה לפרישתן של כמה מדינות בארצות הברית מהפרויקט. פרויקט inBloom נועד ליצירת מאגר ארצי של נתוני חינוך. המאגר מבוסס על ענן שמפעילה Amazon.com ועל מערכת הפעלה שפותחה על ידי News Corporation - Amplify. חברת inBloom הכריזה כי בכוונתה לשתף

Google מגיעה לבית המשפט בגין סריקה של הודעות של Gmail (The Telegraph, ספטמבר 2013).¹ Facebook נתבעת בגין סריקת הודעות 'פרטיות' למטרות רווח (Wired, ינואר 2014).² תביעה - LinkedIn פורצת לדוא"ל של משתמשים, מפיצה דואר זבל לאנשי קשר (GigaOm, ספטמבר 2013).³ "יש חיישנים שעוקבים אחרינו לכל מקום שאנו הולכים אליו. חשבו מהן ההשלכות שלהם על פרטיותו של האדם הממוצע" (אדוארד סנואודן, טלוויזיה, 25 בדצמבר 2013).⁴ הידעתם שהלייקים שאתם עושים ב-Facebook יכולים לחשוף פרטים אינטימיים עליכם ועל תכונות אופי שיתכן שלא תרצו לשתף בהן אחרים? (How Big Data Analytics reveal your most intimate secrets).⁵

שנת 2013 הייתה מלאה כותרות מסעירות על פרטיות מקוונת ועל שימוש וסחר במידע אישי ללא הסכמה. לא פחות דרמטיות היו הכותרות הנוגעות

לתחום החינוך, שהתפרסמו בעקבות מתקפת הסייבר המסיבית בקליפורניה שפגעה גם באוניברסיטאות⁶, או הודאתה של Google לא מזמן, כי היא אכן כורה נתונים מדוא"ל של תלמידים למטרת מיקוד פרסומות ב-Google Apps

for Education⁷. בה בעת, ישנה עלייה משמעותית בהשקעות ובמוצרים (שלל אפליקציות חינוכיות הופכות למקור למידה מרכזי⁸) המבוססים על מערכות ביג דטה בסביבות חינוכיות. הגילויים האחרונים הדליקו נורות אדומות בכל קהילת החינוך. מורים והורים מודאגים לגבי השימוש בנתוני התלמידים בכלל והשימוש לרעה בפרט. "נתוני תלמידים הם הנפט החדש"⁹ היא אמירה הצוברת פופולריות

את הנתונים עם מלכ"רים או עם חברות מסחריות בהסכמת המדינה והמחוז. "הורים, מורים, קבוצות עניין ומומחים לפרטיות בכל רחבי המדינה מחו נגד התוכנית חסרת התקדים לשתף מידע רגיש של ילדים עם חברות פרטיות ועם ספקים מסחריים. הארגונים הניו-יורקים שהתנגדו לכריית נתונים זו היו Class Size Matters, Learning Disability Association of New York, Alliance for Quality Education, New York State Allies for Educational Justice, ו- Coalition for Educational Justice. קבוצות אלה טענו, כי חשיפה של מידע רגיש, או שימוש לא הולם בו, עשויים לסכן את ביטחונם של הילדים."¹¹

מחקר שנערך על ידי Common Sense Media, ארגון המדרג מוצרי טכנולוגיה חינוכית על פי יעילותם והלימותם, הראה כי רוב האפליקציות הסלולריות המיועדות לילדים אוספות נתונים אישיים ומשתפות אותם עם ארגונים מסחריים ללא ידיעתם של ההורים. חשיפה זו הולידה התחייבות של חברות בולטות המציעות טכנולוגיה חינוכית, כגון Google, Pearson, Scholastic, ו-Samsung, לדאוג שהשימוש בנתוני התלמידים יהיה למטרות חינוכיות בלבד ולא לשיווק.¹²

"פיתוח דרכים למעקב אחר נתונים פיזיים ורגשיים, והשימוש בהם על ידי מערכת החינוך, מעלה שאלות נוספות בנושא חדירה לפרטיות התלמידים." השיח על פרטיות יצטרך להשתנות באופן דרמטי בעתיד הקרוב. לא רחוק היום שבו תוכלו לצלם אדם במצלמת הטלפון שלכם, ותוכנה תוכל לזהות גנים מסוימים ב-DNA שלו...."¹³

האם החששות של קהילת החינוך ימנעו פיתוחים מבוססי נתונים בסביבות

חינוכיות? קיים חשש רב בנוגע ליישום תוכניות הדורשות שימוש בנתוני תלמידים, וחשש זה נובע מ"בלבול" או מ"היעדר ידע מסוים" על כל החידושים הטכנולוגיים בביג דטה. על פי דיוויד רובין, עו"ד החבר בפורום האמריקאי Council of School Attorneys, אחד הקשיים הגדולים ביותר של ההגנה על פרטיות הנתונים במחשוב ענן, הוא חוסר הבנה של מועצות בתי הספר והמפקחים המחוזיים. "מתחילים לדבר אתם על פרטיות נתונים ואחסון ענן ורואים אותם עם מבט מזוגג בעיניים. קל מאוד להשתמש במינוח מקצועי ולומר "זו בעיה של IT", אבל כולם צריכים להיות בקיאים ולהבין לעומק את נושא פרטיות הנתונים."¹⁴

השימוש הגובר בביג דטה בסביבות חינוכיות, ובמקביל היעדר המודעות לנושא הפרטיות, מעוררים חשש רב אף יותר בסביבות מקוונות פולשניות, כגון רשתות

חברתיות ואפליקציות סלולריות. "Google ו- Facebook, Verizon, AT&T מוכרות את נתוני השימוש של לקוחותיהם (מיקום, היסטוריית דפדוף ברשת וכדומה). הן מספקות למשתמש גם דרכים "לבחור להימנע" משיתוף הנתונים הללו, אם הלקוח מודע לכך ויודע כיצד לעשות זאת."¹⁵

משתמשים דיגיטליים מספקים נתונים אישיים כאשר הם גולשים באינטרנט, לעתים ביודעין, לעתים מבלי לדעת שהם מספקים מידע לגופים נוספים. היעדר אמון והבנה בקרב משתמשים באשר ליעד הנתונים שלהם יכול לעכב את קצב ההתפתחות של מיזמים חדשניים. הדבר נכון במיוחד בסביבות חינוכיות המיועדות לקטינים. מחקר על פרטיות ברשת ו"מודעות מקוונת" בבריטניה מצביע, כי ההחלטות שצרכנים מקבלים בנוגע לנתונים מקוונים מושפעות מהאופן שהם תופסים את הסיכון ואת התועלת שבחשיפת פרטיותם.¹⁶

ישנם מיתוסים רבים אודות שימוש לרעה בנתונים,¹⁷ וחשוב שחברי קהילת החינוך ידעו שהם יכולים לבחור מתי והיכן יעשה שימוש במידע על התלמידים. על השחקנים הרלוונטיים (מורים, מנהלים, קובעי מדיניות) לבחור בקפידה את סוג הנתונים הנמצאים בשימוש המערכת, כדי להגן על פרטיותו של כל תלמיד.

סקר בנושא מודעות על אופן

שמירת הפרטיות של נתונים מקוונים



המודעות בקרב קהילת החינוך לשימוש ולשימוש לרעה בנתונים מקוונים והחשיבות שהם מייחסים לפרטיות ברשת משפיעים מאוד על ההחלטות שהם מקבלים בנושא יישום מערכות המבוססות על נתוני תלמידים. כדי לשפוך מעט אור על הנושא, נערך סקר שהקיף מדגם מכלל האוכלוסייה, וכן קבוצה נפרדת של מורים ותלמידים. הנבדקים נשאלו על הידע ועל החששות שלהם בנוגע לפרטיות המידע בסביבות מקוונות פופולריות, כגון רשתות חברתיות, אפליקציות סלולריות, מנועי חיפוש ותחומים ספציפיים כגון מערכות הבריאות והחינוך.

1,877 ישראלים השתתפו בסקר, והתוצאות הראו כי רובם מודאגים מנושא

הפרטיות - ברשת, כאשר מקור הדאגה הגדול ביותר הינו השימוש המסחרי במידע פרטי שעושות החברות הסלולריות.

מתודולוגיה

הסקר נערך בשני שלבים: השלב הראשון סקר את כלל האוכלוסייה, והשלב השני התמקד בקהילת החינוך (תלמידים ומורים). הנשאלים ענו על סדרת שאלות בנושא מודעות לפרטיות ושימוש בנתונים (משפטי נכון או לא נכון), וסדרת שאלות נוספת בנושא דאגה לפרטיות (סולם מ-1 עד 5 מ"מסכים מאוד" ועד "לא מסכים כלל"). פרטים נוספים על המתודולוגיה אפשר לבקש מ-MindCET@cet.ac.il או ceciliaw@cet.ac.il.

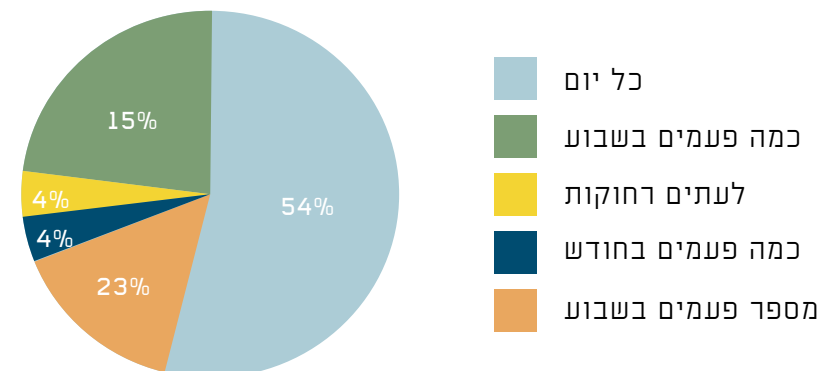
כלל האוכלוסייה

בין 10 בדצמבר 2013 ועד 1 בינואר 2014, נערכו ראיונות טלפוניים עם 1000 ישראלים בגילאי 60 - 18. כדי לקבל מדגם מייצג של האוכלוסייה הישראלית, הנתונים שוקללו לפי מין וגיל בהתאם ליחס האמתי שלהם באוכלוסייה בישראל. המדגם כלל 21% לא יהודים והשאר יהודים, והוא התפלג באופן יחסי בכל אזורי הארץ.

54% מהמשתתפים במדגם השתמשו ברשתות חברתיות מדי יום, 23% מעולם לא השתמשו בהן, וכצפוי - ככל שהגיל היה צעיר יותר, כך תדירות השימוש הייתה גבוהה יותר. גברים ונשים דיווחו על תדירות שימוש דומה ברשתות חברתיות.

התוצאות מצביעות, כי מין וגיל הם משתנים עצמאיים המשפיעים באופן מובהק על רבים מהמשתנים שנסקרו.

באיזו תדירות אתה מתחבר לרשתות חברתיות?



תדירות השימוש ברשתות חברתיות						כל יום או כמה פעמים בשבוע
מין		גיל				
זכר	נקבה	18-29	30-39	40-49	50-60	
69%	70%	85%	65%	62%	55%	

גיל

גיל השפיע באופן מובהק על 7 מתוך 10 ההיגדים בנושא מודעות. עם זאת, נמצאו מגמות שונות לגבי שאלות שונות. נבדקים צעירים הפגינו פחות מודעות לשימוש מסחרי בנתונים על ידי חברות סלולריות, אולם הם הראו מודעות גבוהה יותר למדיניותה של Google בנוגע לתמונות ולמידע אישי, לחשיפה למידע ברשתות חברתיות, כגון Facebook, ולמידע משותף בשירותי הודעות, כגון WhatsApp. צעירים (מתחת לגיל 29) היו פחות מודאגים מסוגיות של פרטיות ברשת. הגיל השפיע באופן מובהק על התפיסה כי אדם יכול להיות אנונימי באינטרנט - ככל שאתה מבוגר יותר כך אתה בוטח פחות באפשרות להיות אנונימי.

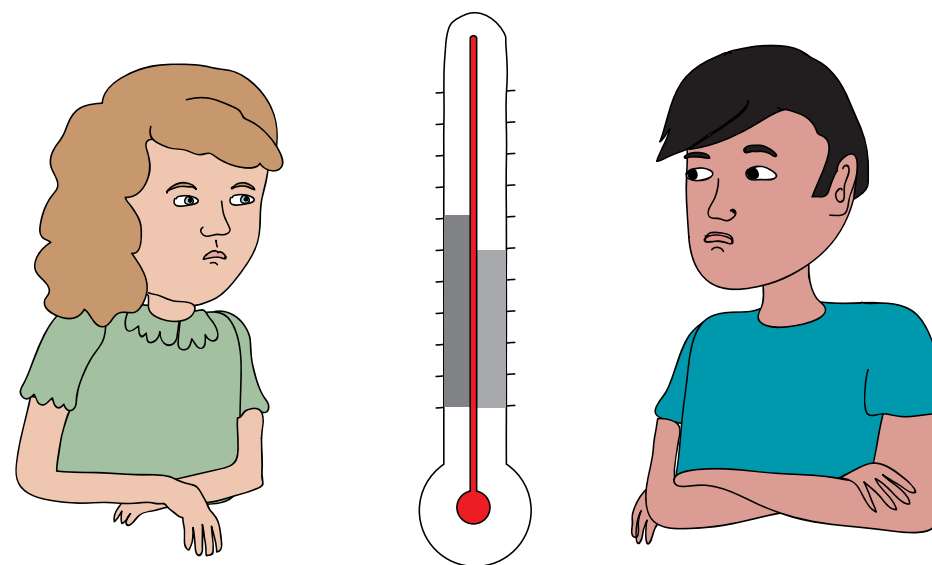


מין

ב- 5 מתוך 10 השאלות בנושא השימוש בנתונים מקוונים גברים היו מודעים יותר מנשים לשימוש המסחרי הנעשה בנתונים, במיוחד לשימוש הנעשה על ידי חברות סלולריות ואפליקציות. המין השפיע באופן מובהק ב-8 מתוך 10 ההיגדים הנוגעים לדאגה לפרטיות - נשים תמיד הביעו דאגה רבה יותר לפרטיותן בתחומים השונים שנסקרו.

60% מהנשים שנבדקו ענו כי הן דואגות לפרטיותן ברשת לעומת 50% מהנבדקים הגברים. לגברים פחות אכפת מאשר לנשים אם למישהו יש גישה לתוכן שלהם ברשתות חברתיות. הם סבורים - יותר מנשים - כי ההזדמנות לחלוק סביבת רשת עם אנשים אחרים מרחיבה את אופקיהם, והם פחות מודאגים לגבי מידע שהם משתפים ב-Facebook, WhatsApp ו-Twitter. באופן מובהק, יותר נשים אמרו שהיו מגיבות באינטרנט לו יכלו לעשות זאת בצורה אנונימית. כמו כן, הן משתמשות יותר, במידה מובהקת, בהגדרות הפרטיות שלהן ב-Facebook כדי להגביל את הגישה לפוסטים שלהן. הן הרבה יותר מודאגות בנוגע לקשרים ברשת מאשר גברים, והן מודאגות יותר לגבי העלאת תמונות שלהן. בקרב קבוצת הנשאלים המשחקים משחקים באינטרנט, 65% מהנשים אינן משחקות עם שחקנים שאינן מכירות בהשוואה ל-44% אחוזים מהגברים.

נשים מודאגות יותר (60%) מפרטיותן
ברשת מאשר גברים (50%)



מודעות



הנבדקים הפגינו מודעות רבה יותר לשימוש מסחרי במידע על ידי רשתות חברתיות (77% מהנבדקים מודעים לפונקציות הגדרות פרטיות ברשתות חברתיות, ו-60% לפרסומות מותאמות אישית על סמך נתונים אישיים) ועל ידי חברות סלולריות/אפליקציות (55% ו-60% מהנשאלים בהתאמה ידעו שנתונים שלהם מועברים לצד שלישי). בכל התחומים האחרים, חלק גדול מהנשאלים לא הראו מודעות לנושאים שהועלו, וסברו שהיגד לא נכון הינו נכון, או שענו "אינני יודע".

בכל הנוגע לסביבות סגורות ברשת (שהמשתמשים נכנסים אליהן באמצעות סיסמה ושם משתמש) הנבדקים הפגינו מודעות נמוכה יותר ליעד המידע המגיע: רק 25% היו מודעים לכך שמידע הנשלח דרך חברות של הודעות מקוונות (כגון WhatsApp) או דוא"ל אינו בלעדי ליעד שאליו הוא אמור להגיע. רק 36% היו מודעים לקיומו של מידע מותאם אישית באתרי אינטרנט, כגון Google (תמונות, נתונים אישיים וכדומה).

מאידך, 22% בלבד מהנבדקים סבורים כי אפשר להיות אנונימי לחלוטין באינטרנט.

פרטיות



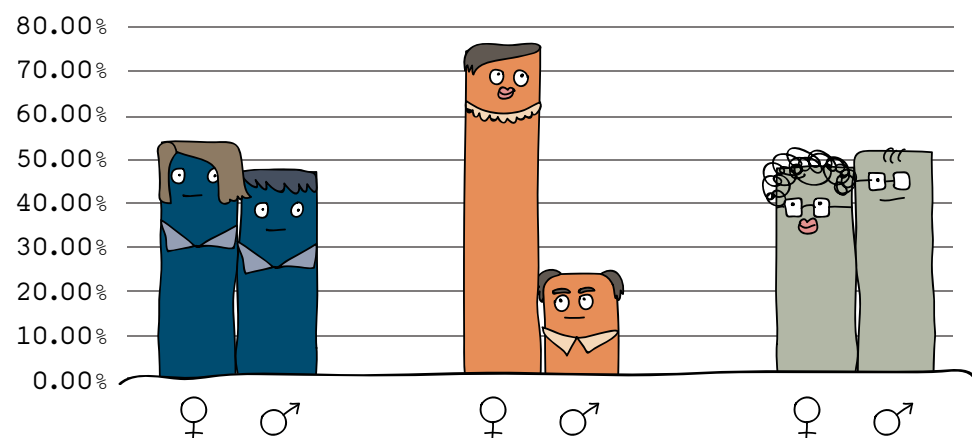
55% מהנבדקים הביעו דאגה רבה מפרטיות ברשת, ו-24% אמרו שאינם מודאגים. בשאלות הנוגעות למערכת החינוך, 34% מסכימים ו-48% אינם מסכימים כי מערכת החינוך צריכה להשתמש או לקבל גישה חופשית למידע אישי של התלמידים או לרשתות החברתיות של התלמידים. הרוב (61%) אינם מסכימים כלל, כי יש לאפשר למערכת הבריאות להשתמש בנתונים האישיים שלהם, גם אם לשם מחקר לקידום הבריאות. 72% אינם מסכימים לשימוש של חברות בנתונים שבטלפון הסלולרי שלהם, גם אם לשם הצעת מבצעים טובים. בשאלות הנוגעות לשיתוף מידע אישי, 68% אמרו שהם מגבילים את הגישה לתמונותיהם המועלות לאינטרנט, 71% אמרו שהם מגבילים את הגישה לפוסטים שלהם ב-Facebook, 50% היו מודאגים לגבי שיתוף מידע ברשתות חברתיות. 36% בלבד אמרו שהם מגיבים באינטרנט (בלוגים, וידאו וכדומה), רק אם התגובה אנונימית.

קהילת החינוך

האוכלוסייה, דוא"ל (מורים) או הקלקה מרצון על פנייה שפורסמה באתר של מוציא לאור בתחום החינוך (תלמידים). התפלגויות הגיל (14 - הצעיר ביותר ו- 63 - המבוגר ביותר) מופיעות בגרפים. ההבדלים בהתפלגות הגיל והמין מוצדקים על ידי הדגימה.

כיצד ענו מורים ותלמידים בהשוואה לכלל האוכלוסייה בכל הנוגע למודעות לשימוש בנתונים ולפרטיות? כדי לענות על שאלה זו, הנתונים נאספו מתלמידים וממורים בנפרד, והתוצאות הושוּו לנתוני כלל האוכלוסייה, וכך נוצר מדגם חדש של 1,887 שאלונים, המחולקים לתת-קבוצות: תלמידים (N=156), מורים (N=721) וכלל האוכלוסייה (N=918¹⁸). הנתונים נאספו באמצעות ראיונות טלפוניים (כלל

מגדר



תלמידים

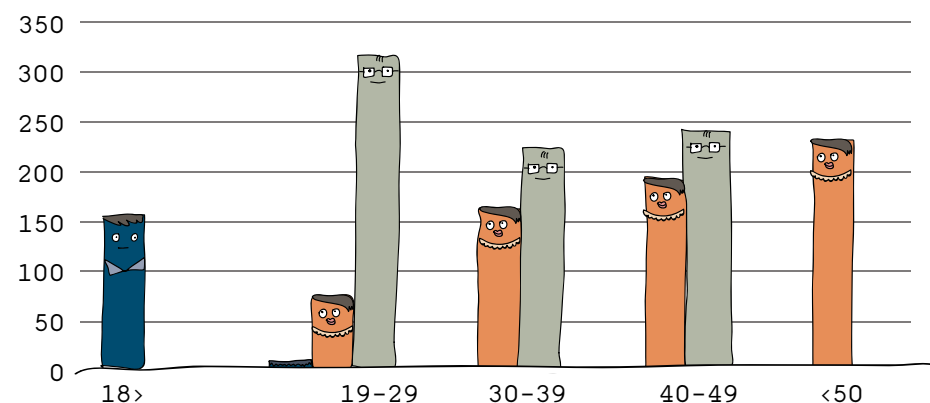


מורים



האוכלוסיה הכללית

גיל



תלמידים



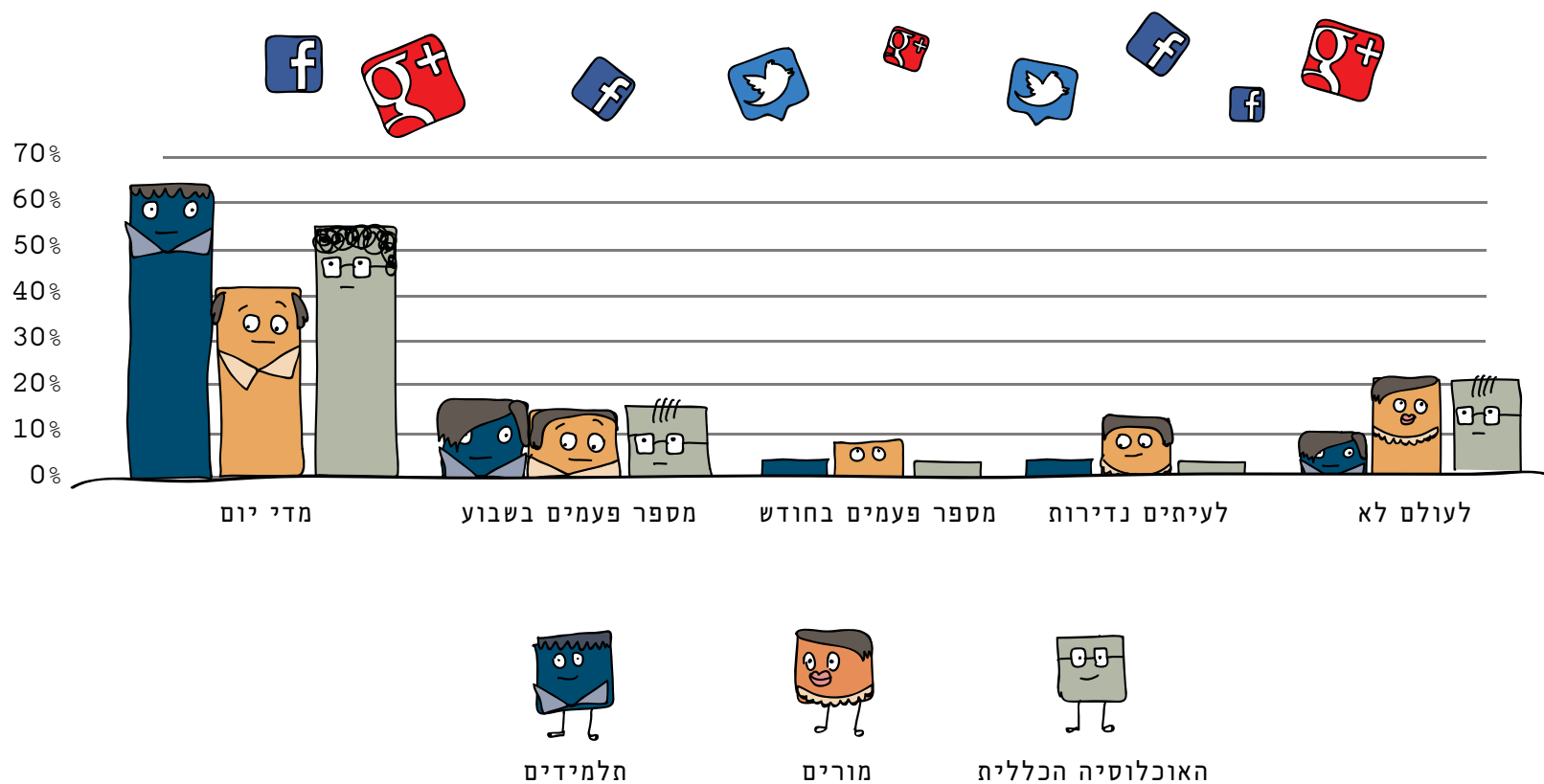
מורים



האוכלוסיה הכללית

מבין המשתמשים ברשתות החברתיות בתדירות הגבוהה ביותר המורים היוו את האחוז הנמוך ביותר (22% מעולם לא השתמשו בהן), והתלמידים - את התדירות הגבוהה ביותר (82% כל יום או כל שבוע). כל אחת משלוש הקבוצות הביעה דאגה בנוגע לרוב התחומים שנסקרו, את נתנה עדיפויות שונות לכל תחום. לשם השוואה, תלמידים היו הרבה פחות מודאגים מפרטיות מקוונת, למעט שימוש בנתונים הנוגעים לבית הספר. כפי שנראה בגרף בנושא פרטיות מקוונת באופן כללי, 50% מהתלמידים אינם מודאגים כלל או מודאגים במידה מסוימת, ואילו 63%

שימוש ברשתות חברתיות

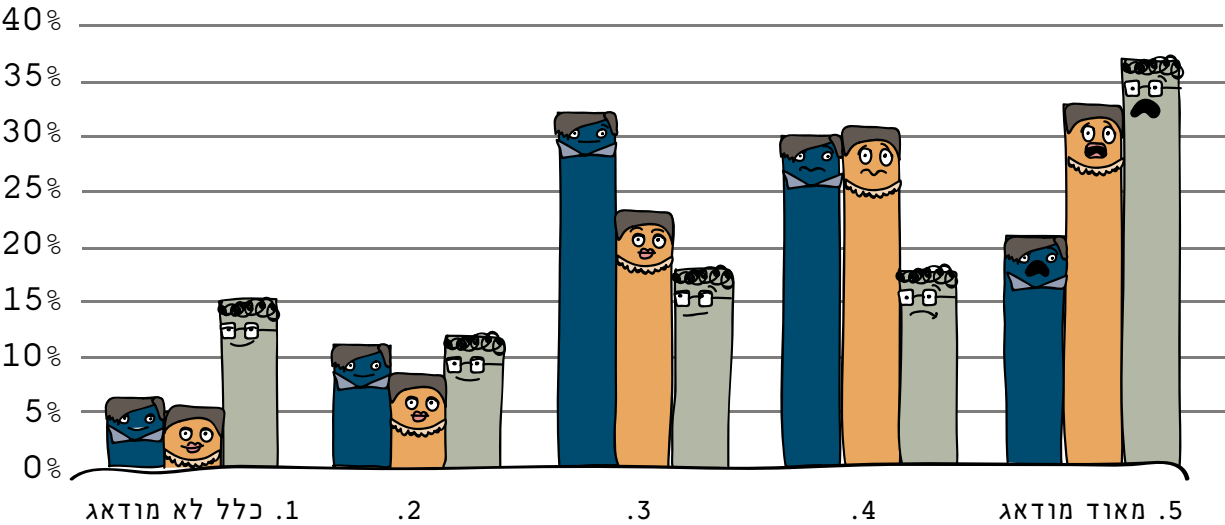


מהמורים ו-60% מכלל האוכלוסייה מודאגים מאוד.

תלמידים ומורים ענו באופן שונה מכלל האוכלוסייה על שאלות הנוגעות לחינוך. רוב התלמידים אינם מסכימים לקביעה כי בתי הספר צריכים להשתמש במידע הלקוח מתוך הרשתות החברתיות שלהם, גם אם המטרה היא לשפר את הלמידה שלהם. מורים ותלמידים כאחד נטו להתנגד לאפשרות להתיר למערכת החינוך

גישה חופשית למידע הפרטי של התלמידים, וזאת לעומת הדעה הרווחת בכלל הציבור, אשר פחות מוטרד ורואה בעיקר את היתרונות בשימוש בנתוני תלמידים במערכת החינוך. מעניין לציין את כמות התשובות "אינני יודע" שהתקבלו, במיוחד מתלמידים ומורים שהינם חלק אינטגרלי ממערכת החינוך.

עד כמה הנך מודאג מפרטיותך ברשת?



תלמידים



מורים

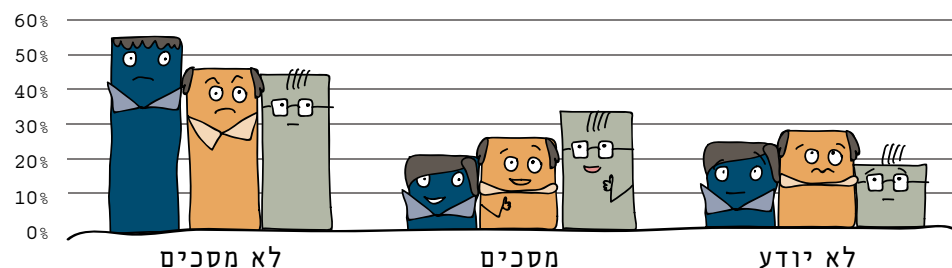


האוכלוסייה הכללית

יחס התלמידים לפרטיותם ברשת אינו חד משמעי. הם הביעו חוסר דאגה בנוגע למשחק עם שחקנים שאינם מכירים או בנוגע למתן תגובות אישיות באינטרנט, אולם רובם כן היו מודאגים בנוגע לשיתוף מידע ברשתות חברתיות. 63% מהתלמידים הביעו דאגה בנוגע להעלאת תמונותיהם, ורובם (69%) משתמשים בהגדרות הפרטיות של Facebook כדי להגביל את המידע המיועד לשיתוף.

אף קבוצה לא הסכימה לתת גישה חופשית למערכת שירותי הבריאות כדי להשתמש במידע אישי, גם אם הוא מיועד למטרות מחקר. מספר גבוה באופן מובהק של מורים היו מודאגים יותר מהשימוש בנתונים על ידי מערכת הבריאות (70%) מאשר על ידי מערכת החינוך (47%). התלמידים, לעומת זאת, היו עקביים והפגינו גישה דומה לשתי המערכות. 50% אינם מסכימים לשימוש חופשי בנתונייהם האישיים על ידי מערכת הבריאות, ו- 55% - על ידי מערכת החינוך.

בתי הספר צריכים להשתמש במידע שתלמידים מפרסמים ברשתות חברתיות, בשביל לשפר את הלמידה שלהם



תלמידים

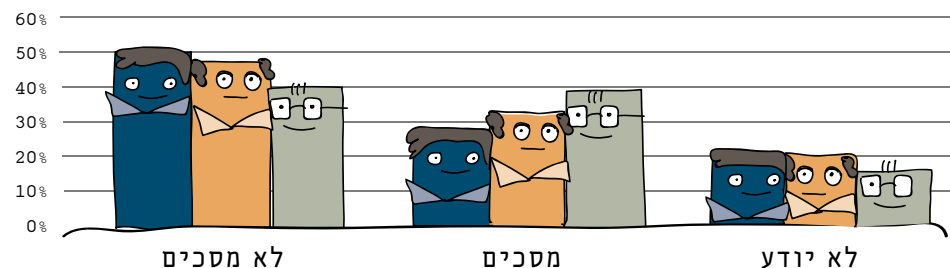


מורים



האוכלוסיה הכללית

אין לאפשר למערכת החינוך גישה חופשית למידע האישי של תלמידים הקיים בבית הספר



תלמידים



מורים



האוכלוסיה הכללית

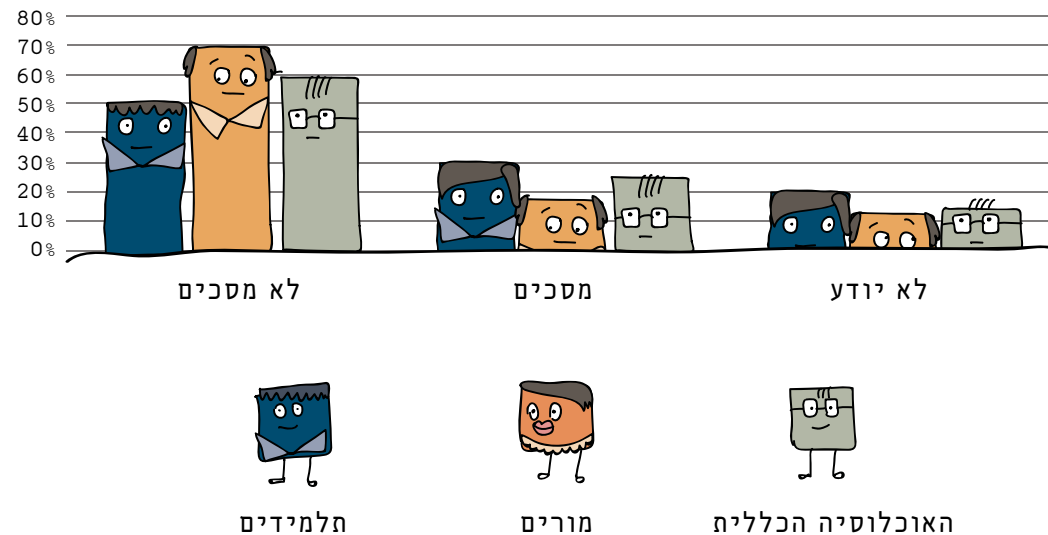
דאגתם העיקרית של המורים (78%) היא השימוש של חברות הסלולר במידע האישי שלהם למטרות מסחריות, דאגה שלה שותפים רוב הנשאלים משתי הקבוצות האחרות (67% מהתלמידים, ו-71% מכלל האוכלוסייה). באופן מפתיע, התוצאות הראו שאין מגמה ספציפית בנוגע לאנונימיות כאשר מדובר במתן תגובות ברשת. שיעור דומה מאוד של אנשים הביעו אכפתיות וחוסר אכפתיות.

מורים הראו מודעות גבוהה יותר לשימוש בנתונים באינטרנט ב-5 מתוך 10 השאלות בהשוואה לשתי הקבוצות האחרות. בחמש השאלות האחרות, התלמידים הפגינו את שיעור המודעות הגבוה ביותר בהשוואה לשתי הקבוצות האחרות. עם זאת, יש להביא בחשבון כי בכל שלוש הקבוצות אחוז נמוך של

אנשים הפגינו מודעות בכלל. מורים היו מודעים יותר באופן מובהק משתי הקבוצות האחרות לשימוש מסחרי בנתונים בטלפונים סלולריים, לדוגמה, 37% מהמורים, 95% מכלל האוכלוסייה ו-94% מהתלמידים היו מודעים לכך שחברות הסלולר יכולות להשתמש בנתונים אישיים לשימושן או להעבירם לצד שלישי.

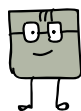
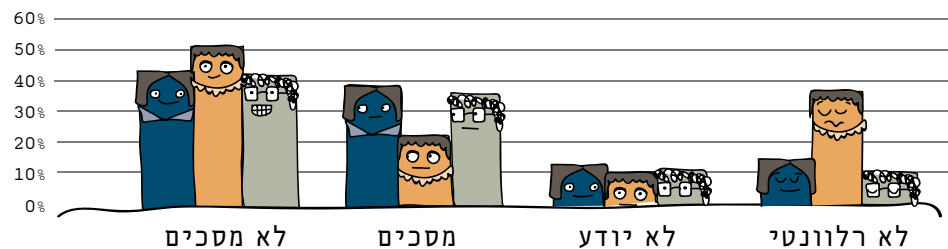
מורים ותלמידים הפגינו מודעות גבוהה בהשוואה לכלל האוכלוסייה לחשיפה ברשתות חברתיות ולשימוש במידע. תלמידים הפגינו מודעות מעט גבוהה יותר מהמורים לגילוי מידע המועבר דרך הודעות מקוונות ודוא"ל.

חשוב לאפשר לקופת החולים שלי לעשות שימוש חופשי בנתוני האישיים, על מנת לחקור נושאים הקשורים לבריאות

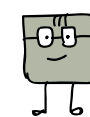
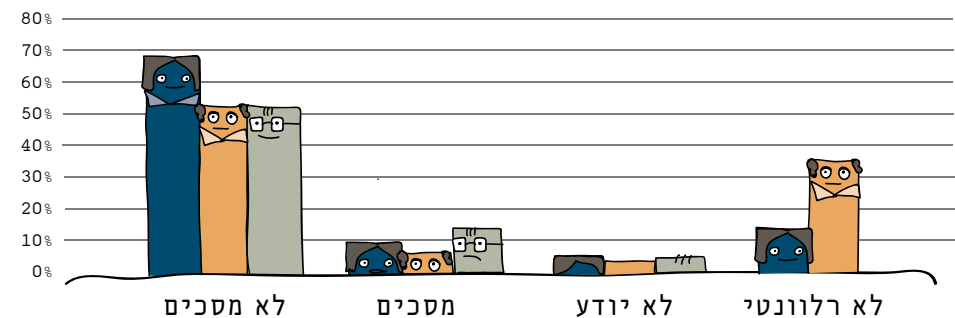


מספר מובהק של נבדקים ענו "אינני יודע" בשאלה הנוגעת למדיניות Google, המאפשרת לחברות אחרות גישה לתמונות או למידע אישי אחר, **והמורים היו מודעים יותר באופן מובהק לעובדה כי המידע המסופק לאתרי אינטרנט הדורשים כניסה מוגנת בסיסמה אינו יוצא מן הכלל.** כל הקבוצות הפגינו מודעות גבוהה לשימוש במידע על ידי רשתות חברתיות (Facebook) משתמשת במידע כדי להתאים אישית פרסומות, והשימוש בהגדרות פרטיות).

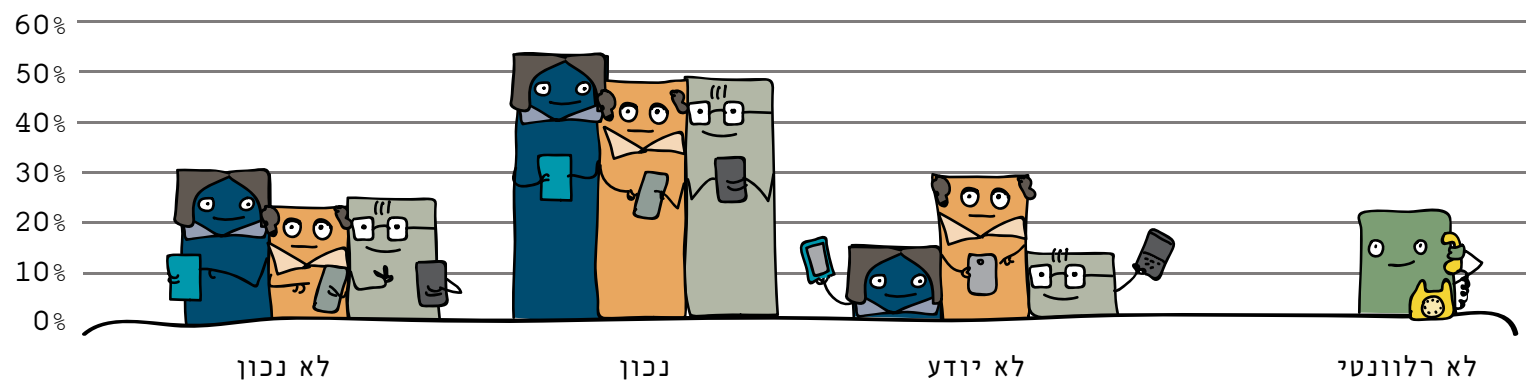
כשאני מדבר עם חברי ברשתות חברתיות לא מדאיג אותי שלמישהו אחר יש גישה לתוכן זה



אני משתמש בהגדרות הפרטיות בפייסבוק על מנת להגביל מי רשאי לקרוא את הפוסטים שלי



כאשר אני משוחח עם חברים בקבוצה בווטסאפ, רק לחברי הקבוצה יש גישה למידע שלי



תלמידים

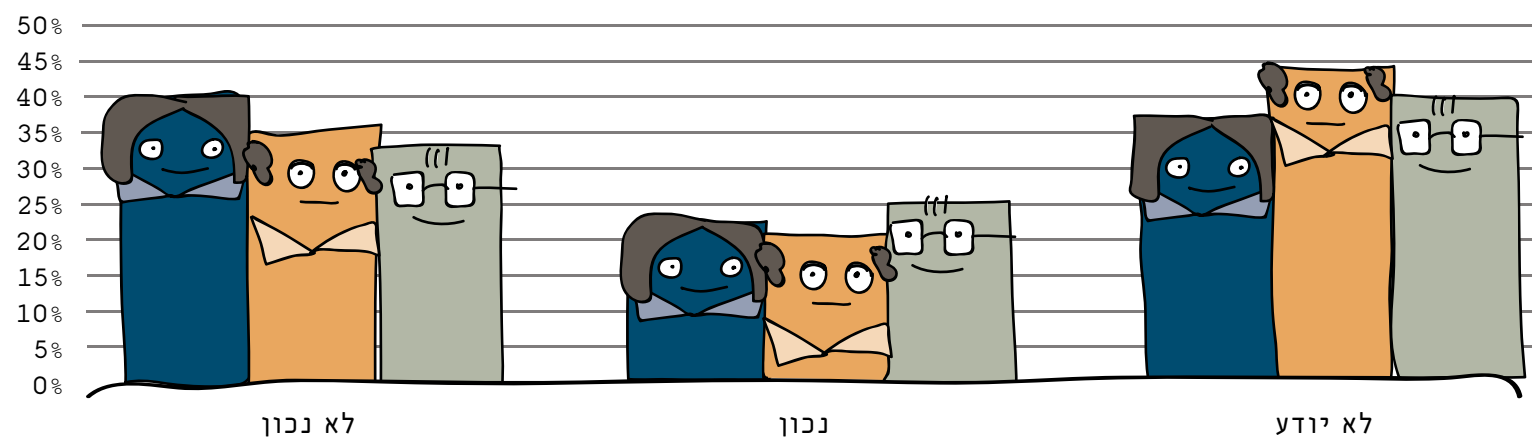


מורים



האוכלוסיה הכללית

מדיניות גוגל כיום לא מאפשרת לחברות אחרות לגשת אל התמונות שלי או למידע האישי שלי



תלמידים

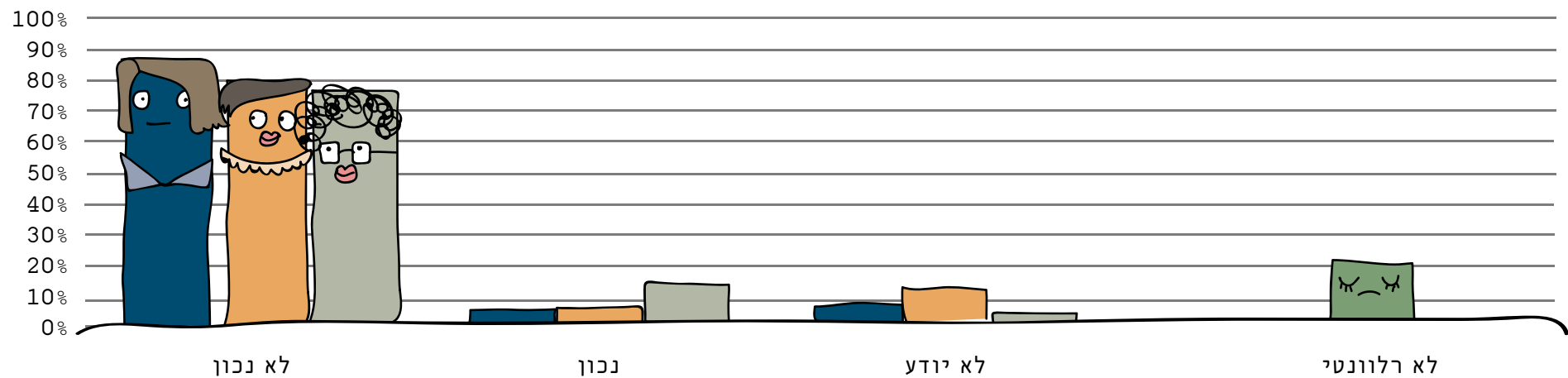


מורים



האוכלוסיה הכללית

כל אחד יכול לראות, לשתף
ולהעביר ברשת החברתית מידע
שאינו מוגדר כפרטי



תלמידים



מורים



האוכלוסיה הכללית

מחקרים קודמים מצביעים כי למשתמשים אכפת פחות משימוש במידע שלהם כאשר מדובר בחברות שהם בוטחים בהן (כמו Google לדוגמה)²⁰. הסקר שלנו הניב תוצאות דומות לתוצאות סקרים אחרים, כגון: מודעות גבוהה להגדרות פרטיות ברשתות חברתיות, ומודעות נמוכה יותר לשימוש בנתונים בטלפונים סלולריים לעומת השימוש שנעשה על ידי רשתות חברתיות. התוצאות אינן מעידות שהדור הצעיר מודע יותר לשימוש בנתונים ברשת, אלא שדווקא המורים מודעים יותר!



הסקר מצביע בבירור על דאגה בקרב האוכלוסייה הישראלית משימוש בנתונים אישיים ברשת. בהשוואה למחקר אחר¹⁹ שהקיף 10,354 מבוגרים מכל העולם, בקרב הישראלים הרבה יותר אנשים מודאגים (55% מהישראלים בהשוואה ל-37% מהנבדקים מכל העולם). נשים באופן עקבי היו מודאגות יותר מגברים, והגיל לא הצביע על מגמה ברורה בכל השאלות הנוגעות לדאגה לפרטיות, תוצאה דומה לתוצאות סקרים אחרים ברחבי העולם.

כל המבוגרים (מעל גיל 18) היו מודאגים יותר מהשימוש בנתונים האישיים על ידי שירותי הבריאות והעברתם לגופים אחרים, גם אם למטרות מחקר, בהשוואה לדאגה משימוש בנתונים במערכת החינוך. רק קבוצת התלמידים (מתחת לגיל 18) הייתה עקבית ולא הסכימה לשימוש חופשי כלשהו של מערכת בנתונים אישיים. ככל שהאדם מעורב ישירות במערכת, כך גדלה הדאגה שלו לגביה. הסקר הצביע גם על היעדר מודעות לשימוש בנתונים אישיים ברשת או על חשד משמעותי (כמות גדולה של תשובות "אינני יודע"). הכותרות הרבות באמצעי התקשורת המצביעות על השימוש לרעה או על ניצול של נתונים אישיים והיעדר מדיניות פרטיות ברורה וקלה לקריאה, יכולים להסביר את "הבלבול" בקרב הנשאלים.

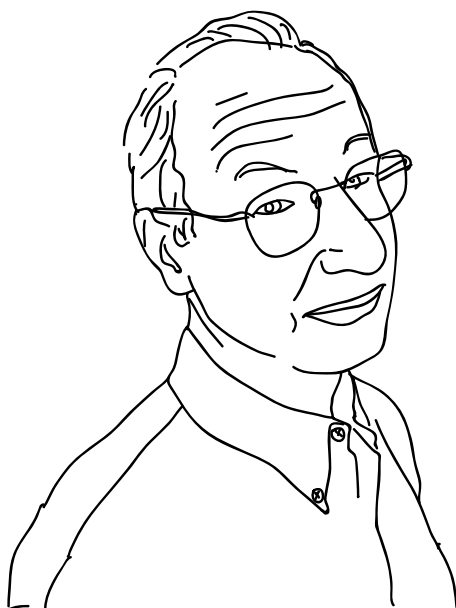
מורים ותלמידים כאחד הביעו עמדה ברורה כנגד שימוש חופשי בנתוני התלמידים על ידי מערכת החינוך, גם אם לשם שיפור הלמידה. תוצאה זו מצביעה, כי ישנה דאגה מובהקת בקהילת החינוך בנוגע למערכות למידה, כגון אלו המוזכרות בדוח זה, הדורשות לאסוף את נתוני התלמידים. יתירה מזאת, התוצאות מעלות חשש בנוגע לגבולות של מערכת החינוך - אילו נתונים של תלמידים היא יכולה להחשיב כנתונים לגיטימיים לאיסוף.

רק המורים היו עקביים כאשר הפגינו מודעות גבוהה להעברת מידע לצד שלישי על ידי חברות אינטרנט שונות. יתר המדגם הפגין אחוז ממוצע של 45%.

1. <http://www.telegraph.co.uk/technology/google/10289798/Google-goes-to-court-over-Gmail-scanning.html>
2. <http://www.wired.co.uk/news/archive/2014-01/03/facebook-private-messaging-lawsuit>
3. <http://gigaom.com/2013/09/21/linkedin-is-breaking-into-user-emails-spamming-contacts-lawsuit/>
4. <http://www.dailymail.co.uk/news/article-2529236/US-whistleblower-Edward-Snowden-delivers-alternative-Christmas-message.html>
5. <http://smartdatacollective.com/bernardmarr/129421/how-big-data-analytics-Facebook-likes-reveal-your-most-intimate-secrets>
6. <http://www.ecampusnews.com/top-news/universities-student-privacy-444/>
7. <http://safegov.org/2014/1/31/google-admits-data-mining-student-emails-in-its-free-education-apps?body=http://safegov.org/2014/1/31/google-admits-data-mining-student-emails-in-its-free-education-apps>
8. <http://www.eschoolnews.com/2014/02/19/resources-education-apps-722/?ps=CeciliaW@cet.ac.il-001300000135NyG-0033000001CbsTP>
9. <http://hackededucation.com/2013/10/17/student-data-is-the-new-oil/>
10. http://blogs.edweek.org/edweek/marketplacek12/2013/12/new_york_battle_over_inBloom_data_privacy_heading_to_court.html
11. <http://www.wnyc.org/story/307074-what-you-need-know-about-inbloom-student-database/>
12. http://www.nytimes.com/2013/10/14/technology/concerns-arise-over-privacy-of-schoolchildrens-data.html?_r=0
13. <http://www.wired.com/wiredscience/2012/10/big-data-is-transforming-healthcare/>
14. <http://www.eschoolnews.com/2013/11/20/data-privacy-cloud-133/4/>
15. <http://techcrunch.com/2013/07/05/att-considers-selling-your-browsing-history-location-and-more-to-advertisers-heres-how-to-opt-out/>
16. Online personal data: the consumer perspective, 2011
17. <http://www.eschoolnews.com/2013/11/20/data-privacy-cloud-133/>
18. To be able to compare the three groups, all teachers were taken out of the original general population sample (N=1000).
19. ComRes polling and research consultancy. March 2013. Big Brother Watch Online Privacy Survey Online Personal Data the Consumer Perspective, UK, 2011

ביג דטה וחינוך -
הזדמנויות או איום?

שואלים את המומחים!



דיוויד ויינברגר,

מנהל מעבדת המדיה בהרווארד

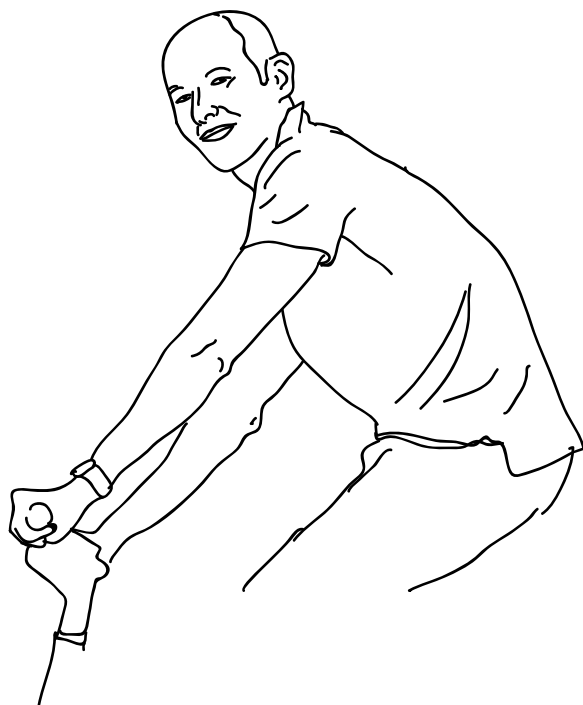
זוהי הזדמנות עצומה עבור הסטודנטים להסתכל מחוץ לקופסה השחורה של רעיונות מחקריים, שכאילו ירדו מהשמים, ולהבין אותם לעומקם. ככל שנתונים הופכים לביג דטה, ובמיוחד כאשר הכלים לאינטראקציה עם ביג דטה הופכים קלים ומוכרים יותר, החומה המפרידה בין למידה למחקר תיטשטש עוד יותר. ככל שביג דטה יהפוך להיות חלק מנחלת הכלל בעולם כולו, הסטודנטים יראו את הכוח הטמון לא רק בו, אלא גם בשיתוף פעולה חברתי שעניינו נתונים, רעיונות וידע.

קובי גל,

מנהל המחלקה לקבלת החלטות אדם-מחשב,

אוניברסיטת בן גוריון

הודות לביג דטה אנו קרובים יותר מאי פעם לחולל מהפיכה בחינוך. זה מתאפשר בשל שתי התפתחויות מרכזיות: מחד, למידה מקוונת או מתווכת על ידי תוכנות נעשית נפוצה יותר ויותר, והיא יוצרת מידע רב בכל הנוגע לתוכן אקדמי ולאינטראקציות התלמידים עם התוכנה, עם המורים וזה עם זה. מאידך, ההתפתחויות בבינה מלאכותית ובלמידת מכונה הניבו אלגוריתמים שמסוגלים להסיק מגמות ודפוסים מורכבים מנתונים בקנה מידה גדול. השילוב של שני ההיבטים הללו יכול לספק כלים חינוכיים מהפכניים לתועלת התלמידים והמורים. טכנולוגיות אלה כוללות: (1) תוכנה חינוכית התומכת בבניית דגם, בחקירה, בניסוי וטעייה, ומספקת סביבה חינוכית עשירה לתלמידים; (2) אלגוריתמים חכמים לניתוח האינטראקציות של התלמידים עם התוכנה בזמן אמת והסקת מסקנות על פעילויותיהם; (3) תמיכה בסיוע למידת מכונה המנחה את אינטראקציות התלמידים כדי להפיק למידה מרבית תוך ההתערבות מועטה ככל האפשר; (4) כלי ויזואליזציה חדשניים שיכולים לספק למורים הערכה בזמן אמת של עבודת התלמידים תוך שיפור הבנת התלמידים את תהליכי הלמידה במינימום מאמץ.



לאון מרקוביץ,

יזם, מייסד שותף ב-WikiBrains

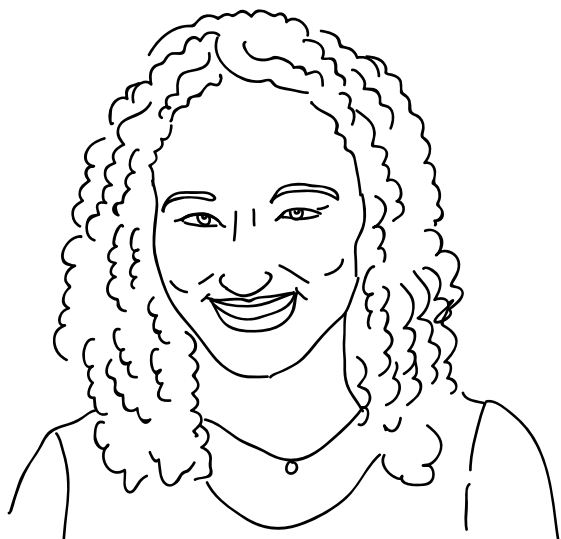


קרוב ל-90% מהנתונים בעולם נוצרו בשנתיים האחרונות. לטוב או לרע, זוהי מהפכה שרק התחילה. כל חברה מאמצת היום, יותר מאי פעם, את הנתונים שמאחורי המספרים והסטטיסטיקות שלה. המחשבים ממשיכים להשתפר בניתוח ביג דטה ובמציאת דפוסים, וככל שיותר אנשים יתוודעו לתהליך הזה, התובנות שנמצא ימשיכו להדהים אותנו. כיום, כל מה שאתם עושים באינטרנט, וגם רוב מה שאתם עושים מחוץ לאינטרנט, נמדד: מה אתם אוהבים, אחרי מה אתם עוקבים, למה אתם מאזינים, וכן הרגלי הקנייה שלכם, שגרת היומיום ובקרב... עוד הרבה יותר. מטרת איסוף הנתונים הזו היא לספק לחברות וללקוחותיהן את השירות הנחוץ להם בזמן המתאים. כיום אפשר להשתמש בכוחו של ניתוח ביג דטה כדי למנוע בעיות ולהתגבר עליהן בזמן אמת! מודעות מעוררת השראה כזועמדה מאחורי הקמת WikiBrains. האלגוריתם מנסה לחקות את הדרך שבה המוח עובד על ידי יצירת דפוסים, המבוססים על חוכמת ההמונים. כל נושא מייצג נידון, וכל קישור לנושא אחר הוא כמו סינפסה. כשלוקחים מיליוני מוחות התורמים אסוציאציות ויוצרים מפות, אנו יכולים לעזור לאחרים לגלות במהירות קשרים הדדיים כדי להגיע לתובנות חדשות.

ישעיהו ברלין אמר ש"הבנה היא הבחנה בדפוסים", ומטרתה של WikiBrains היא לקרב את ההמונים לתפיסה זו. היופי הוא שככל שמשתתפים יותר אנשים, התוצאה הופכת חכמה יותר יותר. אין צורך להסתמך על עצמך, אפשר להשתמש בקשרים שאחרים יוצרים וליהנות מהם. אתר אינטרנט כמו WikiBrains עשוי להגדיר מחדש יצירתיות ולהוביל לתובנות מהירות יותר בקרב יחידים וארגונים. WikiBrains, במהותו הוא תרשים, שמוצא קשרי גומלין בין פרטי מידע שנראה כי אין קשר ביניהם. היא עוזרת לגלות קשרים שייתכן שלא ידעתם שהם קיימים. WikiBrains משמשת חלון ל"מוח" של העולם ועוזרת לזהות אילו אסוציאציות מתעוררות בקרב אוכלוסיות שונות בנושאים שונים במשך זמן. אם אתה ארגון, היא עוזרת לך לראות באופן חזותי כיצד כל דבר וכל נושא מקושר, תוך סילוק הרעש המיותר.

אורלי פורמן,

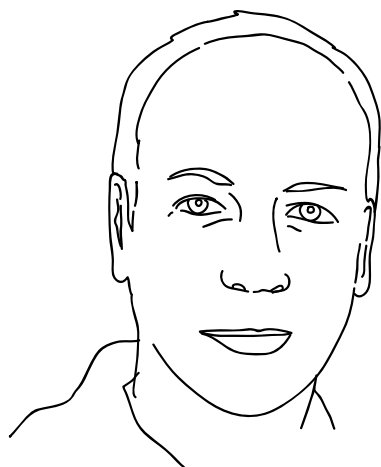
יזמית, מייסדת שותפה ב- **Lingua.ly**



מהפכת הביג דטה נוגעת בעיקר להיבטים בחיי היומיום שלנו וטומנת בחובה הזדמנות גדולה אף יותר כאשר מדובר בדרך שאנו לומדים. רבים מסכימים, כי הגיע הזמן למהפכה אמתית בחינוך: בתי ספר בעולם כולו עדיין פועלים על פי עקרונות אשר עוצבו בעידן המהפכה התעשייתית, תכניות הלימודים סטנדרטיות, והמורים עסוקים רוב הזמן בהעברת תוכן לכיתות צפופות. בסביבה כזו קל להתעלם מההעדפות, מההבדלים ומהצרכים המיוחדים של היחיד. אין פלא שתלמידים רבים מאבדים עניין בלימודיהם - הם קשים מדי, קלים מדי או פשוט משעממים.

ביג דטה טומן בחובו הבטחה גדולה לפתרון בעיה זו ולהשגת למידה מותאמת אישית. אנו יכולים לראות כבר היום כלי ביג דטה במערכות בחינה וקורסים (כגון Desire2Learn, Knewton, Benchprep ועוד), המאפשרים הערכה מהירה ומדויקת יותר של סגנונות הלמידה של תלמידים ושל החוזקות והחולשות שלהם. יש עדיין תכנית לימודים של שיעורים ותוכן הנבחר בקפידה, אולם כל תלמיד יכול לנווט בה בדרך המותאמת לו אישית והמעודדת יותר מעורבות ומוטיבציה מאשר "תכנית אחת המתאימה לכולם".

Lingua.ly לוקחת את הלמידה המותאמת אישית צעד אחד קדימה: היא משתמשת בכלי ביג דטה כדי להוציא את הלמידה מהתחום המוגבל של קורס, ספר לימוד או תכנית לימודים ולהעבירה לאינטרנט; אין מערך שיעור או מטלות עבודה על טקסטים; התלמידים יכולים ללמוד בכל זמן שירצו, ובכל מקום באינטרנט: עיתונים מקוונים, בלוגים, דוא"ל, Facebook ועוד. Lingua.ly אוספת מידע על כל תלמיד, כדי לעקוב אחר אוצר המילים, הרמה וההתקדמות שלו. לאחר מכן היא עוזרת לו לתרגל ומוצאת ברחבי האינטרנט תוכן אמתי המותאם לפרופיל האישי שלו. בכל הנוגע להוראה, גישה זו מאפשרת למורה להתפנות ולהתמקד ברמות הגבוהות יותר של הלימוד (ולא בתהליך איסוף המידע) ובמעקב אחרי ההתקדמות של התלמידים.



נתן אינטרטור,

פרופסור בבית הספר למדעי המחשב ומדעי המוח

באוניברסיטת תל אביב

ביג דטה הוא היכולת ליצור ביעילות מודלים, המתבססים על כמויות גדולות של נתונים, שהינם איכותיים הרבה יותר ממה שיכולנו להשיג לפני מספר שנים. יתרה מזאת, זו היכולת להתאים מודלים מסוימים למקבצי נתונים שונים, או לאוכלוסיות שונות, ובכך להתייחס בצורה נכונה להבדלים בין קבוצות. היבט זה חשוב בחינוך וברפואה מותאמת אישית לשם אבחון וטיפול.



שחר שפיגלמן,

מנהל תכנית, אנליטיקה מתקדמת באינטל

לא זו בלבד שביג דטה אינו מהווה איום, אלא הוא אחת ההזדמנויות הגדולות ביותר בתחום החינוך. ביג דטה יאפשר להתאים את החינוך לצרכים המסוימים של כל תלמיד. הוא גם יספק למורים את הכלים הדרושים לחינוך טוב יותר לכל אחד מתלמידיהם. דמיינו עולם שבו מורה/הורה יכול לראות את כל מסלול הלמידה של התלמיד/הילד שלו מגן הילדים, ולהבין בבירור את הידע ואת הפערים שלו.



עומרי מנדלס, מדען נתונים, אינטל

ביג דטה, בכל תחום, יכול להוביל לקבלת החלטות אנליטית, מבוססת נתונים וטובה יותר. בתחום החינוך, הוא יכול לשפר פלאים את ארגז הכלים של המורה, בעיקר בכל הקשור להבנה מעמיקה של צרכיו של כל תלמיד, שאיפותיו, הצדדים החלשים והחזקים שבו. היתרון המרכזי של השימוש בביג דטה הוא ביכולתו להבדיל את היחיד מהכיתה ולהתאים תכנית לימודים אופטימלית לכל תלמיד.

עם זאת, ללא אינטליגנציה רגשית מתאימה והבנה טובה של המודלים הקוגניטיביים שאנו משתמשים בהם כבני אדם, הוא לעולם לא יוכל לתמוך באופן מלא בתהליך למידה אמתי.

ארנון הרשקוביץ,

מרצה בכיר, המחלקה למדעים ולטכנולוגיה בבית

הספר לחינוך, אוניברסיטת תל אביב



נראה כי אנו נמצאים בתחילתו של עידן חדש, שבו אנו יכולים להשתמש במגוון עצום של נתונים הנאספים במסגרת תהליכי למידה והוראה. נתונים אלה יכולים להפוך תהליכים אלה ליעילים יותר, דהיינו לשפר את ביצועי התלמידים, מערכות החינוך ומקבלי ההחלטות. זוהי הזדמנות נפלאה, שאת ניצניה אפשר כבר לראות במחקר וביישום בעולם כולו. כך, לדוגמה, מערכות למידה אינטיליגנטיות (ITS) יכולות לזהות דפוסי התנהגות של תלמידים - קוגניטיבית, מטא-קוגניטיבית ורגשית - ובמקרים מסוימים להגיב בהתאם. כבר כיום מורים ומרצים יכולים לשפר את הידע שלהם על תלמידיהם באמצעות שימוש במערכות למידה מבוססות מחשב. מוסדות חינוך יכולים לנתח כמויות עצומות של מידע מנתונים מנהליים שהם צוברים באופן שוטף, כדי לזהות הצלחה (או לחלופין, נשירה) של תלמידים. ובעיקר, מחקר המשתמש בשיטות שנלקחו מעולם הביג דטה מרחיב ללא הרף את הידע שלנו על תהליכי הלמידה וההוראה.

אבל - וזהו אבל חשוב - ישנם אתגרים חשובים שעדיין מהווים מכשול למימוש החזון הנפלא הזה של ביג דטה בחינוך. אדון בקצרה במכשולים העיקריים (בסדר שרירותי). זוהי בהחלט אינה רשימה מקיפה, וחשוב לזכור שזהו תחום חדש יחסית, ולכן המכשולים להתפתחותו עדיין רבים.

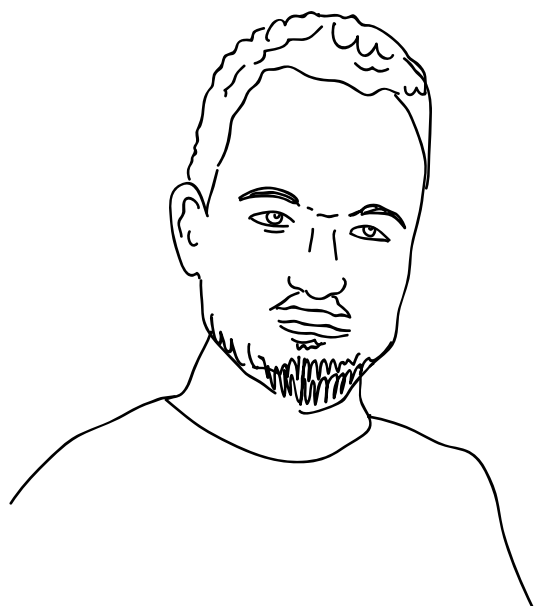
סטנדרטיזציה - תיקנון (מכשול טכני): כיום, אין תקן לאיסוף ולאחסון נתונים בתחום החינוך. במיוחד, אין תקן לאיסוף קבצים מתוכנות חינוכיות. מצב זה מקשה על חוקרים הרוצים לקבל נקודת מבט רחבה וכוללת (השיטות והגישות המפותחות כדי לטפל בנתונים בפורמט ספציפי עשויות לא להתאים לנתונים אחרים). מצב זה גם מערים קשיים על מפתחי אפליקציות גנריות, הרוצים להגיע למספר גדול ביותר של משתמשי קצה (תלמידים, מורים, מנהלי בית ספר, מקבלי החלטות).

בידוד דיסציפלינרי (מכשול מחקרי): רבים מהעוסקים בתחומים שונים של מחקר חינוכי אינם

יודעים ששיטות הלקוחות מעולם הביג דטה יכולות לסייע להם במחקרם. אבל גם אם הם יהיו מודעים לכך - אפשר שלא יוכלו תמיד ליישם שיטות אלה, בשל קשיים אובייקטיביים בהבנתם. כדי ליישם בהצלחה שיטות אלה, דרוש שיתוף פעולה רב-תחומי והפריה הדדית. אף על פי שקהילות המחקר הבינלאומיות העוסקות בביג דטה בחינוך (שתי העיקריות ביניהן Society for Learning Analytics ו-International Educational Data Mining Society Research) הן רב-תחומיות במהותן, כאשר אנו מתבוננים באנשים המרכיבים אותן, מיזמים רב-תחומיים משותפים עדיין אינם בגדר נורמה.

מודעות בשטח (מכשול יישומי): האם התלמידים מודעים לכך שכאשר הם משתמשים בתוכנת למידה, נאספים עליהם נתונים שיכולים לשפר את השימוש שלהם במערכת ואת הלמידה שלהם? האם המורים העובדים עם תוכנת הלמידה יודעים כי נאספים נתונים המתעדים את פעולות התלמידים, וכי נתונים אלה יכולים לסייע להם ולתלמידיהם? האם מנהלי בתי ספר יודעים כי אפשר לנתח את הכמויות העצומות של הנתונים המנהלתיים הנאספים במערכות הממוחשבות של בית הספר שלהם (כגון משוב) כדי להשיג מידע חשוב, אם כי בלתי נראה, הנוגע להיבטים שונים של ניהול בית הספר? הגברת המודעות לאפשרויות אלו עשויה להגדיל את הביקוש לאפליקציות ובכך להאיץ את המחקר והפיתוח בתחום זה. סוגיות אתיות ומשפטיות (נדוש, אבל חשוב להזכיר): על מחקר, פיתוח ויישום בתחום זה להתבצע תוך הגנה על פרטיות המשתמשים במערכות השונות. זוהי לכשעצמה סוגיה כבדת משקל, אולם לעת עתה מספיק רק להזכירה.

רם הדר, יזם, מייסד Quototi



עד לא מזמן, כל החברות היו צריכות להיאבק כדי להשיג מידע על הלקוחות הקיימים או הפוטנציאליים שלהם. כיום כל מה שצריך לעשות הוא לפתוח את הסכרים הווירטואליים (מ - API, ל - Screen scrape וכל מה שביניהם) והמידע פשוט "זורם" פנימה.

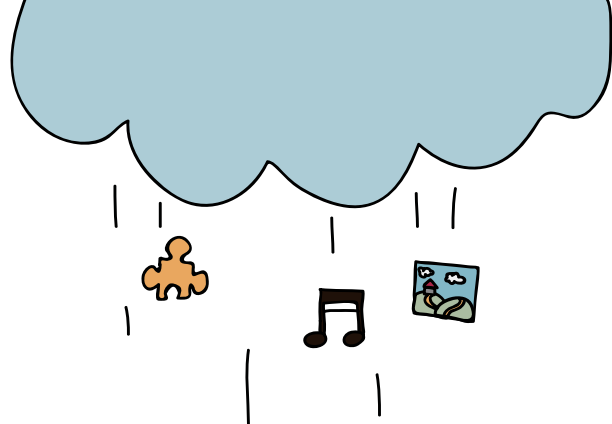
מה שמפתיע הוא, שבזמן שאנו מוסרים את המידע שלנו לחברות שונות - מרצוננו, ולרוב ללא ידיעתנו, אך בהסכמתנו ("אני מסכים"), החל ממשחקים פשוטים המתקיימים גם שירותים המתעדים כל פעילות שלנו, וכלה באפליקציות מיוחדות, מערכות הפעלה המייצרות דוחות ושליחת דוגמאות DNA לניתוח, שעבורו אנו משלמים - בזמן שאנחנו עושים זאת אנו עדיין לא מעוניינים לספק מידע פחות רגיש ליעדים ידועים אחרים. הבעיה השתנתה מ"כיצד להשיג את המידע" ל"לאן הולך כל המידע" הנמצא מאחורי משחק תמים שילדכם התקין.

חשוב לזכור כי"אם מענים מספיק את המידע, הוא יודה בכל דבר". במילים אחרות, יש לתכנן את המידע הרצוי לפני תחילת העבודה. אחרת, התוצאה, כמו המידע, תהיה לוקה בחסר.

אנו ב-Acovado מתמחים בעולם הביג דטה, באסטרטגיה, בפיתוח ובניתוח נתונים, ואנו עוזרים לחברות לעצב את אסטרטגיית המידע שלהם לעתיד הקרוב תוך תכנון לטווח הארוך.

כאשר הקמנו את פלטפורמת Quototi העוזרת למשתמשים לשתף ולאצור טקסט באמצעות מערכת "סמן ושתף", אנו מודעים גם לסכנות. פיתחנו פתרון, אשר מחד, ממנף את פעילותו הטבעית של הקורא כאשר הוא מסמן ומעתיק טקסט וכך הוא מראה לנו מה הוא רואה כתמצית המאמר, ומאידך, נותן למשתמש בעלות מוחלטת על נתונים גולמיים אלה.

המילון החסר



המילון החסר



בינה עסקית Business Intelligence

שורת מתודולוגיות להפיכת נתונים גולמיים למידע שימושי למטרות עסקיות

ענן Cloud

תשתית מחשוב מתכנסת המספקת יכולת להריץ ולשמור מידע של אחרים ממיקום מרוחק

חקר הגנום הצרכני consumer genomics

הגנום הצרכני - תכונות הקובעות את התנהגות המשתמש

אוצר תוכן content curator

מסייע בתהליך חיפוש, איסוף ושיתוף של תוכן קיים

קוקיז cookies

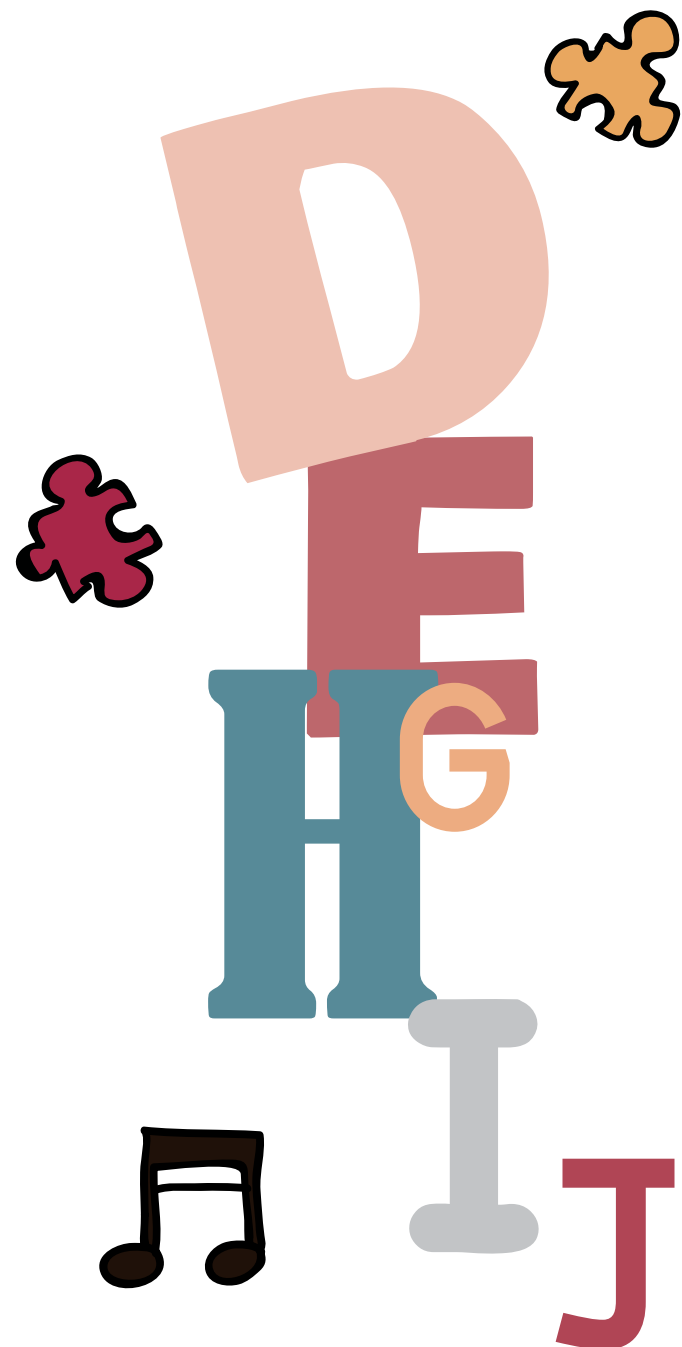
סימנים בלתי נראים המושתלים על הדפדפן שלך כדי לעקוב אחר מידע

ניהול קשרי לקוחות CRM

מודל שיווק המבוסס על מעקב אחר פעולות הצרכן לשיפור ההתאמה האישית

ניתוח מגמות תרבות Culturomics

תחום מחקר המקשר בין מגמות תרבותיות לניתוח איכותי של שימוש במילים על פני תקופת זמן ספציפית



ברוקר נתונים Data Broker

אדם או ארגון החוקר מידע

data dives ("מפגשי נתונים")

אירועים המתקיימים בסופי שבוע המפגישים אנשי מקצוע שונים עם מדעני נתונים

מיצוי נתונים Data exhaust

תוצרי לוואי של מידע, דהיינו פעילות מקוונת

אגרני נתונים data hoarders

אנשים השומרים יותר נתונים ממה שהם צריכים

מונטיזציה של נתונים data monetization

תהליך שבו נתונים מומרים להכנסה, דהיינו שימוש בנתונים למטרות שיווק

מסד נתונים Database

אוסף מאורגן של נתונים

דטפיקציה Datafication

תלות גוברת בנתונים לפעילות שוטפת

("חנות נתונים") datashop

מאגר גדול ופתוח של נתונים המשקפים אינטראקציה עם טכנולוגיה חינוכית

כריית נתוני חינוך Educational Data Mining

תחום מחקר העוסק בנתונים מתחום החינוך

גיגה בייט Gigabytes

יחידה של זיכרון מחשב או קיבולת אחסון נתונים השווה ל- 1,024 מגה בייט או 2^{30} בייט (או רשימה של 2^{33} אחדות ואפסים)

NLP RTP TS

Hadoop

מסגרת קוד פתוח לניתוח מקבצים גדלים של נתונים

חישוב-על HPC

פעילות חישובית הדורשת יותר ממחשב אחד לביצוע משימה

ממגורת מידע information silo

מערכת ניהול מידע שאינה יכולה לתקשר עם מערכות ניהול מידע אחרות

intelligent lampposts

רשתות חיישנים למערכות תעבורה יעילות מבחינה אנרגטית

JSON

ציון אובייקטים ב-javascript

ניתוח למידה Learning Analytics

תחום מחקר העוסק בנתונים מתחום החינוך

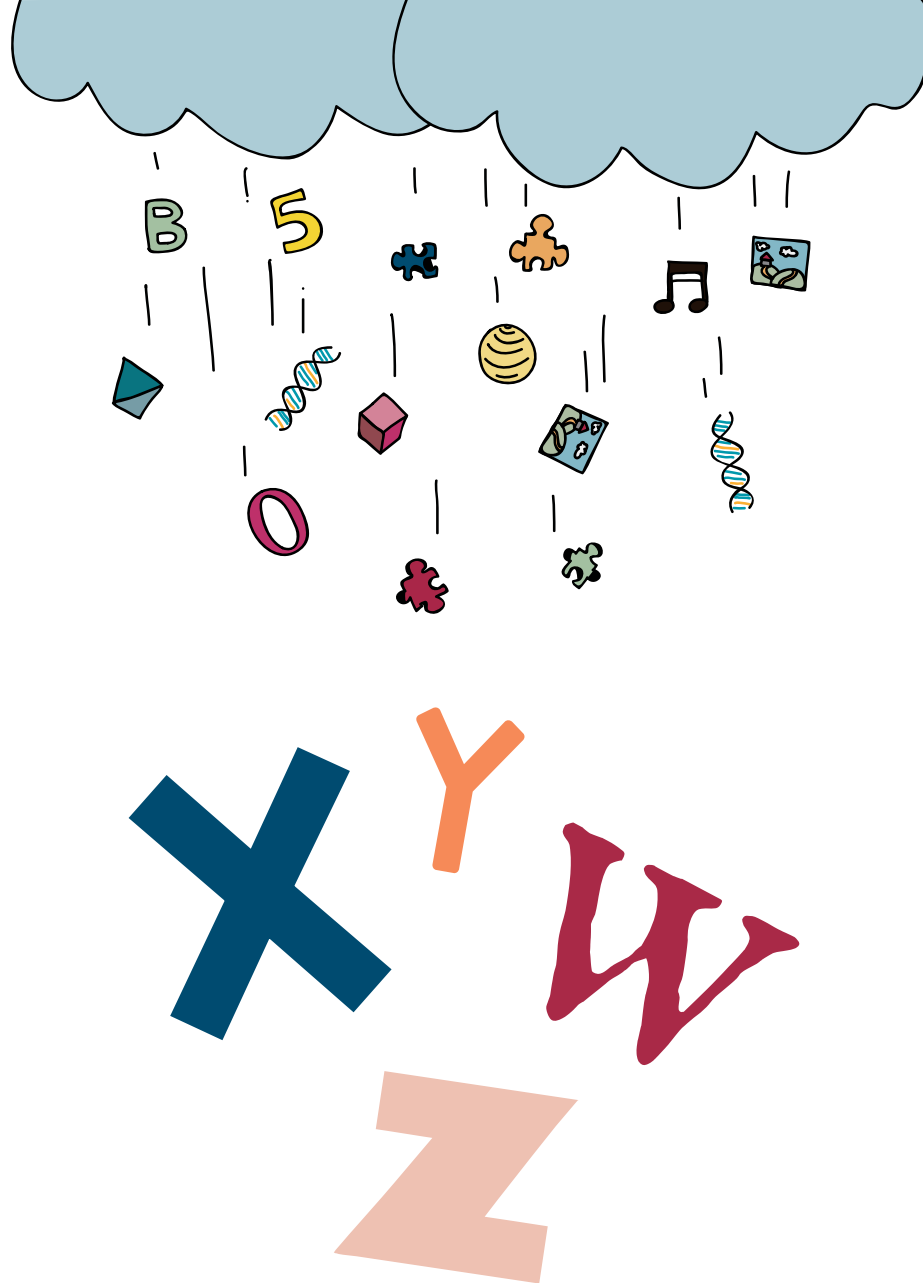
נחונים נזילים liquid data

נתונים פתוחים, זמינים באופן נרחב וניתנים לשיתוף

NoSQL

פורמט של מסד נתונים המספק מנגנון לאחסון ולשליפת מידע





פטה ביט Petabytes

יחידה של זיכרון מחשב או קיבולת אחסון נתונים השווה ל- 2^{15} ביט (או רשימה של 2^{17} אחדות ואפסים)

בסיס נתונים יחסי RDBMS

פורמט למסד נתונים שבו המידע מקובץ ביחס למפתח מוגדר מראש

אגירה siloing

תהליך או מצב שבו מידע אינו מועבר בין חלקים שונים של מערכת או ארגון

("תלמיד מדומה" simstudent)

סימולציה של תלמיד

נתונים זורמים Streaming Data

תהליך העברת נתונים לא מעובדים בקצב מהיר וקבוע

TDA

ניתוח נתונים טופולוגי

טרה ביט Terabytes

יחידה של זיכרון מחשב או קיבולת אחסון נתונים השווה ל- 1,024 גיגה ביט או 2^{40} ביט (או רשימה של 2^{43} אחדות ואפסים)

במבט מסכם:

במבט מסכם:

דבר MindCET:

ערכו האמיתי של ביג דטה הוא במגוון האפשרויות
ההולך וגדל שהוא מספק לנו, כמשתמשים דיגיטליים,
והפיכת הנתונים למצרך זמין לכולם. כולנו משתמשים
בנתונים דיגיטליים, כולנו תורמים להם, עלינו רק
להגביר את המודעות לשימוש הולם בהם.

הפיתוח הנוכחי של מערכות חכמות, המבינות את תהליך הלמידה ומתאימות את תגובתן בזמן אמת, על בסיס מספר עצום של נושאים ובאמצעות שלל ערוצי תקשורת ומידע - פיתוח כזה יכול לגרום מהפכה בחינוך. מערכות למידה המתאימות את ההוראה לצורכי היחיד, מאפשרות לבסס מדיניות על מידע מקרו בזמן אמת, והיכולת לאפשר לכלל קהילת החינוך גישה למידע רלוונטי, שומנות בחובן הזדמנויות ייחודיות לפיתוח מערכת חינוך רלוונטית הרבה יותר.

הציפיות שביג דטה יחשוף את האמת על ההתנהגות האנושית עשויות להוביל לתפיסות מוטעות ולמעשים מוטעים. יש לחשוב ברצינות על אמצעים מגבילים ומונעים, בעיקר בכל הקשור לחיזוי או לניבוי של עתיד התלמידים. יש לתת מענה לחששות בנוגע לפרטיות באמצעות קביעת מגבלות על הגדרת נתוני החינוך הפתוחים לכול. מערכות חכמות צריכות לאפשר יצירתיות של תלמידים ולטפח פעילות קוגניטיבית חופשית המובילה לאפשרויות חדשות לפתרון בעיות. ההתאמה האישית צריכה לשפר את יכולת המערכת לתת מענה לצרכיו האמיתיים של המשתמש.



אנו סבורים, כי הגברת המודעות באשר למשמעות הביג דטה היא המפתח להפיכתו לכלי יישומי אמיתי. הבנת השימושים והאפשרויות הגלומות בביג דטה יכולה להעניק תפקיד פעיל לכל חברי קהילת החינוך לקראת בנייה מחדש של מערכת חינוך טובה יותר.

תלמידים והורים יכולים להיות שותפים להחלטות על גבולות הנתונים של התלמידים שאפשר לגייס לטובת מערכת החינוך, וכך יהיה אפשר לקדם פיתוח מערכות למידה משמעותיות ובה בעת להגן על פרטיות היחיד. היעדר מודעות לסכנות הקיימות, עלול להוביל להערכות מוטעות ולעיכוב ביישומן של מערכות חכמות היכולות להועיל בצורה משמעותית לתלמידים.

מודעות של מורים ושל מקבלי החלטות בתחום החינוך להכרה שנתונים הם אמצעי עזר להבנת התלמיד, אך הם אינם מציגים את כל רובדי התלמיד - מודעות כזו עשויה לתת משקל מאוזן ויעיל יותר לכוחות החיזוי של ביג דטה. פיתוחים של יזמים בתחום ה- EdTech (Educational Technology) יכולים לסלול את הדרך לצורות חדשניות של חקר נתוני תלמידים, אך על המפתחים לשמור על ראש פתוח וסקרנות בתחום העצום הלא מוכר עדיין - הביג דטה.



ראש העיר, הכיתה של בני סיימה את המבחנים הארציים לפי דקה. ניתן לראות את התוצאות המאכזבות בכל המדינה. בוא נחשוב יחד כיצד ניתן לעזור לילדים.



אחרי שקיבלתי את כל המידע לגביך, אני מאוד מסוקרנת להכיר אותך.



תודות

עורכים.....ד"ר ססיליה וייסמן
כתיבה.....ד"ר ססיליה וייסמן
.....רן מגן
מחקר.....ד"ר ססיליה וייסמן
.....אסטלה מלמד
.....רן מגן
.....עידן יצחקי
.....מכון שריד: שרותי מחקר והדרכה
.....נורית וטניק
משתתפים.....אבי ורשבסקי
.....צוות MindCET
.....מחלקת שיווק מטח
.....Shine Little Studio עיצוב גרפי
.....נעמי פיין איור
.....ניר וייס סרט מבוא
.....Inter Office ltd תרגום לעברית

נרצה להודות לכל המומחים ששיתפו עמנו בשמחה את הידע שלהם.



כל הזכויות שמורות © 2013 MindCET

כתובת: צבי בורשטיין 904, ירוחם

קלאוזנר 16, תל אביב ישראל

אתר: www.MindCET.org

טלפון: +972-3-6460924

מייל: MindCET@cet.ac.il