

ביג דאטה בחינוך - פוטנציאל ואתגרים

ג'ורא אלכסנדרון וארנון הרשקוביץ¹

בשנים האחרונות עובר העולם מהפכה טכנולוגית, חברתית ותרבותית שבמרכזה האיסוף והשימוש ב-"ביג דאטה" (נתוני עֵתָק) למטרות שונות. בבסיסה מבוססת מהפכה זו על שילוב של שני גורמים - נתוני עתק וכוח חישוב חזק וזמין - ובעיקר על ההבנה כי ניתן לנצל את אלו למטרות שונות.

1. מבוא

בשנים האחרונות עובר העולם מהפכה טכנולוגית, חברתית ותרבותית שבמרכזה האיסוף והשימוש ב-"ביג דאטה" (נתוני עֵתָק) למטרות שונות. בבסיסה מבוססת מהפכה זו על שילוב של שני גורמים - נתוני עתק וכוח חישוב חזק וזמין - ובעיקר על ההבנה כי ניתן לנצל את אלו למטרות שונות. נתוני העתק הם תולדה של השימוש הנרחב במערכות דיגיטליות מקוונות ומרושתות כמעט בכל אספקט של החיים המודרניים - תקשורת, פיננסים, מסחר, תרבות, ביטחון, תחבורה ועוד; מערכות אשר מייצרות ושומרות בעת פעילותן "היסטוריה" דיגיטלית של המשתמשים בהן. כך, למשל, כל עסקה בכרטיס אשראי שאנו עושים, נשמרת במחשבי חברת האשראי. החברות משתמשות במידע זה בין היתר על מנת להציע ללקוחות מבצעים המותאמים להם באופן אישי; כך הלקוחות מרגישים מרוצים מן היחס האישי שהם זוכים לו, והחברה מגדילה את הסיכוי כי המבצע המוצע אכן יתממש לכדי עסקה חדשה. ההמלצות הללו הן תוצאה של שימוש באלגוריתמים חכמים מבוססי ביג-דאטה, אשר מפותחים הן באוניברסיטאות והן בארגונים מסחריים, והם אלו שהופכים את הנתונים הרבים למידע רב-ערך.

עולם ה-"ביג דאטה" חדר בשנים האחרונות גם לתחום החינוך, וזאת בשל התרחבות השימוש בסביבות דיגיטליות ללמידה ולניהול למידה, כגון מודל (Moodle), לומדות שונות, קורסים המוניים פתוחים (MOOC-ים) ומערכות ניהול בית ספריות. תוכנות

אלו שומרות באופן מתמיד מידע רב ורב-ערך, שבו ניתן לעשות שימוש כדי לייעל תהליכי למידה והוראה. זה כעשור עוסקים בתחום זה חוקרים רבים ברחבי העולם ממגוון דיסציפלינות מדעיות.

למהפכת ה-"ביג דאטה" בחינוך פוטנציאל לשנות את פני החינוך במספר דרכים (Bienkowski, Feng and Means, 2012). ראשית, השימוש בנתוני עתק מאפשר לחקור את תחום הלמידה, ואת הלמידה המקוונת בפרט, בצורות שלא התאפשרו בעבר, מתוך מטרה לשפר שיטות הוראה, להתאים חומרי למידה ולייעל מערכות חינוך (Siemens, 2013). שנית, במערכת החינוך מספר גדול של בעלי עניין - לומדים, מורים, מנהלים ומפקחים, קובעי מדיניות ועוד. לכל אחד מהם צרכים ומטרות שונים, ולכולם היכולת להסתייע במידע המבוסס על הדאטה המדובר (Lee et al., 2015). שלישית, מערכת החינוך עוסקת באופן שוטף בהערכה ומדידה מתוך מטרה להעריך את הלומדים, המורים, בתי הספר והמערכת בכללותה. השימוש ב-"ביג דאטה" מאפשר הערכה עשירה ומדויקת יותר. רביעית, מערכת החינוך מורכבת מלומדים רבים, שאינם זהים בצרכיהם, יכולותיהם, העדפותיהם ומטרותיהם; "ביג דאטה" מאפשר לאפיין את הדומה ואת השונה בין התלמידים וליצור מנגנונים להתאמה אישית של הלמידה ללומדים (EdSurge, 2016). שני הפרקים הבאים עוסקים בפוטנציאל הטמון בשילוב של "ביג דאטה" בתחום החינוך ובאתגרים שונים שאיתם נדרש להתמודד כדי להוציא פוטנציאל זה לפועל.

1 ד"ר ג'ורא אלכסנדרון, המחלקה להוראת המדעים, מכון ויצמן למדע, רחובות

ד"ר ארנון הרשקוביץ, בית הספר לחינוך, אוניברסיטת תל אביב

2. ביג דאטה בחינוך - פוטנציאל

בכיתה ח' בבית הספר "שדות האתמול" המורה למדעים עומדת ומסבירה לכל התלמידים על הפרדת תערובות באמצעות כרומוטוגרפיה, ואף מקרינה סרטון BrainPOP קצר בנושא. חלק מן התלמידים מקשיבים, חלקם לא; חלק מאלו שמקשיבים מבינים, חלקם לא (אך לא נעים להם לשאול). היא נותנת לילדים מטלה לבית, שאותה הם מתבקשים להגיש בזוגות, על פי בחירתם. בתום תקופת הלימוד של הנושא, המורה נותנת מבחן שאותו כותבים התלמידים על דפים בזמן השיעור. לאחר שבועיים היא מחזירה את המבחנים - על הדף הראשון מתנוסס ציון מספרי, ובין דפיו מסתתרות הערות משוב קצרות. חלק מן התלמידים כבר שכחו על מה היה המבחן, כי עכשיו לומדים את הנושא אנרגיה חשמלית. בכיתה ח' בבית הספר "ניצני המחר" באותה העיר ממש, התלמידים עובדים באותן שעות במסגרת פיילוט עם מערכת למידה אדפטיבית מבוססת-נתונים². המורה למדעים מסתובבת בין תלמידיה, אשר עובדים עם תוכנה ייעודית להוראת נושא הפרדת תערובות באמצעות כרומוטוגרפיה. כל אחד מן התלמידים מתקדם בקצב שלו ומקבל מהמערכת באופן מידי משוב מפורט לתשובותיו. בתוך המערכת הם מבצעים סימולציה ובוחנים את ההשפעה של חום המים על הפרדת צבעים של עלים. בתוך כך המורה מקבלת על המסך שלה חיווי ויזואלי ברור לגבי ההתקדמות של כל אחד מן התלמידים, ויודעת אל מי מהם כדאי לה לגשת לצורך מתן סיוע. בשעה שהיא מסתובבת ביניהם, קוראים לה חלק מן התלמידים לשאלות אישיות שלא נעים להם לשאול בקול רם בפני כל הכיתה. הנתונים על התקדמותם מסייעים למורה גם בחלוקת הכיתה לזוגות לצורך הגשת מטלת בית. בסיום הנושא אין צורך במבחן, שכן המערכת משקללת את כלל הפעילות של התלמידים כדי להעריך אותם בצורה מדויקת ועשירה, וכוללת בניתוח שלה התייחסות גם להיבטים כגון יצירתיות או פערים שיש צורך להשלים.

אם התיאור של הכיתה השנייה נשמע לכם דמיוני - חשבו שוב! מערכות כגון המערכת המתוארת נמצאות כיום במחקר ופיתוח באקדמיה ובתעשייה ובכיתות ניסוי. מערכות אלו מתבססות על שיטות מעולם ה"ביג דאטה", שמאפשרות לנתח בזמן אמת כמויות גדולה של נתונים מקיפים ומעמיקים. המידע המופק מנתונים אלו

עשוי לספק מידע רלוונטי לבעלי עניין, כגון תלמידים, צוות בית הספר, הורים, מקבלי החלטות בתחום החינוך, מפתחי סביבות למידה דיגיטליות ומפתחי תכנים. מידע כזה מאפשר לטייב תהליכי קבלת החלטות ולשפר תהליכי למידה.

דוגמה בולטת לחשיבות השימוש בנתונים המתארים את הלמידה היא השינוי האפשרי באופן שבו תלמידים מוערכים. פעולת ההערכה מתבצעת היום בעיקר באמצעות מבחנים בתום פרק למידה תוך שימוש במדדי הישגים בלבד. בעזרת השימוש בנתונים, לעומת זאת, ההערכה יכולה להפוך לחלק בלתי נפרד מן הלמידה עצמה. הערכה כזו תיתן משוב בזמן אמת לתלמידים ולמורים, על מנת לשקף לראשונים את תהליכי הלמידה שלהם, ואילו לאחרונים - לתת תמונת מצב עדכנית על המתרחש בכיתותיהם. הערכה כזו יכולה להביא בחשבון לא רק ידע, אלא גורמים רבים נוספים, כמו ניסיונות שנעשו עד למתן התשובה, תדירות העבודה ומיקום הלמידה, שיתוף פעולה עם עמיתים וכו' (Berland, Baker, & Blikstein, 2014).

באמצעות שימוש ב"ביג דאטה" ניתן להתאים לכל תלמיד נתיב למידה אישי מיטבי, על ידי שימוש באלגוריתמים המנתחים את היסטוריית הלמידה של התלמיד ונעזרים בתובנות שהופקו מהיסטוריית הלמידה של תלמידים ה"דומים" לו. כך למשל, יקבלו התלמידים משימות המתאימות לרמה שלהם ולתחומי העניין שלהם או יופנו למשאבי למידה שנמצאו אפקטיביים בפתרון קשיים דומים אצל תלמידים אחרים (Alexandron, Zhou, & Pritchard, 2015). סוג כזה של התאמה-מבוססת-אלגוריתמים במערכות ללמידה אישית מכונה למידה אדפטיבית (Adaptive Learning). בלמידה מסוג זה, שיש בה מרכיב משמעותי של למידה עצמאית עם מחשב, עולה חשיבותה של האינטראקציה בין הלומד למחשב. כלים המנצלים אלגוריתמים מהתחום של ניתוח שפה טבעית (Natural Language Processing) יכולים לשמש כדי לשפר ולהאניש את האינטראקציה בין המחשב ללומדים, לתת משוב מידי על תשובות פתוחות שעד היום חיכו לבדיקת מורה, לזהות תסכול, חוסר תשומת לב ועוד (Hershkovitz, de Baker, Gobert, Wixon, & Pedro, 2013). התקווה היא שההתאמה האישית גם תגביר את המעורבות של התלמידים בלמידה ותעלה את המוטיבציה שלהם, ועדייות ראשונות על

2 מערכת למידה אדפטיבית היא מערכת המתאימה את הלמידה באופן אוטומטי לצורכי הלומד הספציפי

3. ביג דאטה בחינוך - אתגרים

מהפכת ה"ביג דאטה" בחינוך מציבה אתגרים רבים, הן כדי למצות את הפוטנציאל הגלום בנתונים, והן כדי להבטיח שהשימוש בהם נעשה בצורה נאותה. להלן התייחסותנו למספר אתגרים מרכזיים. סביבות למידה יעילות משלבות בצורה נכונה בין התוכן הלימודי, הפדגוגיה והסביבה הפיזית. כדי לנצל בצורה מיטבית כלים מבוססי דאטה, נדרש עיצוב מחדש של כל אלו. בפרט, על מנת לעשות שימוש בדאטה, יש צורך בהגברת השימוש בסביבות למידה דיגיטליות. כאמור, השימוש בסביבות אלו מצריך חשיבה מחדש על הפדגוגיה ועל הסביבה הפיזית. מכאן נגזרות גם מיומנויות חדשות שעל המורים לרכוש, ובתוכן היכולת לפרש ייצוגי מידע מורכבים ולפעול בעקבות זאת לקידום הלמידה.

אתגר אחר קשור ל"מאחורי הקלעים" של האלגוריתמים המבוססים על "ביג דאטה". למשל, כדי לאפשר לאלגוריתם לקדם תלמידים בקורס פיזיקה ממוחשב, האלגוריתם צריך "להבין" את מבנה הידע בפיזיקה, לדעת להתאים בין השאלות במערכת למבנה הידע הזה, ולדעת אילו נושאים עשויים לסייע בקידום של נושאים אחרים. רק כך תוכל המערכת - אחרי שהיא מזהה קשיים וחוזקות של תלמידים בשאלות מסוימות - להתאים לתלמידים מסלולים מיטביים לקידום הלמידה. אתגר זה קשור בטבורו לתחום הבינה המלאכותית, העוסק במידול וייצוג של ידע ובפיתוח של תהליכי קבלת החלטות 'חכמים' המבוססים עליו.

בישראל, כמו גם ברוב המדינות, כולל המדינות המפותחות, קיים מחסור בתשתיות טכנולוגיות (מכשירי קצה, רשת אלחוטית בפס רחב) במערכת החינוך הציבורית. הדבר מוביל לסינדרום של ביצה ותרנגולת: בהיעדר תשתית מתאימה, קשה לאמץ חידושים פדגוגיים מבוססי טכנולוגיה; הקושי להחדיר טכנולוגיות למידה חדשניות לשוק החינוך מייצר אפקט מצנן על ההיתכנות הכלכלית של פיתוח טכנולוגיות חדשניות, שבהיעדרן אין מוטיבציה להשקיע בתשתיות שיעמדו ללא שימוש, וחוזר חלילה...

בנוסף לתשתית, הטכנולוגיה הנדרשת כדי לאפשר את מהפכת ה"ביג דאטה" בחינוך מציבה גם היא אתגרים משמעותיים למימוש המהפכה. מה שמקשה במיוחד הוא החוסר בסטנדרטיזציה. מערכות שונות אוספות דאטה שונה ושומרות אותו בפורמטים שונים, וכך קשה לפתח כלים ושיטות גנריים שיעבדו עם כל

האפקטיביות של מערכות ללמידה אדפטיבית מתחילות להיצבר. כמובן, השימוש בנתונים לא יאפיין רק את המורים והתלמידים. מנהלי מוסדות חינוכיים יכולים להסתייע ב-"ביג דאטה" כדי לזהות דפוסי נשירה (Kotsiantis, Pierrakeas, & Pintelas, 2003); מפתחי תכנים יכולים לקבל מידע מדויק על ההתמודדות של התלמידים עם התכנים המוצגים בפניהם (למשל, האם הווידאו אכן עזר לתלמידים לענות על השאלות שאחריו? האם השאלות קשות/קלות מדי? וכו'). מעצבי מערכות למידה יכולים לקבל מידע רב-ערך על אופני השימוש בפלטפורמות שפיתחו (למשל, האם התלמידים קוראים את ההסברים לתרגיל או משתמשים בעזרה המוצעת להם? האם יש חשיבות להגדלת הגופן או להצגת הטקסט בצבע אחר? וכו').

לחוקרים בתחום החינוך מעניק עידן ה"ביג דאטה" הזדמנויות נפלאות. באמצעות נתוני העתק ניתן לחקור בשיטות חדשניות שאלות מסורתיות על אודות תהליכי למידה. כך, למשל, ניתן לבנות מודלים מתמטיים המתארים בדיוק רב את התקדמות הלמידה של כל תלמיד בנושא מסוים, ואף להציע למורה דרכים לקדם כל אחד מן התלמידים. דוגמה נוספת: ניתן לחזות מראש מי מן התלמידים עשוי להיקלע למצוקה ולהתערב בעוד מועד כאשר משלבים נתונים ממקורות שונים - למשל, מתוכנה ללימוד מתמטיקה, ממערכת הניהול הבית-ספרית וממערכת שבה מדווחים התלמידים על תחושותיהם לאורך הלמידה. השימוש בנתוני עתק מאפשר לעקוב אחר תהליך הלמידה ברזולוציה גבוהה מאוד ולזהות בדיוק רב את הרגע שבו התלמיד מגיע לידי הבנה. מהפכת ה"ביג דאטה" מאפשרת לחקור שאלות ודילמות חדשות, האופייניות ללמידה מבוססת-מחשב. כך, למשל, למידה באמצעות לומדות המציעות עזרה, מאפשרת לתלמידים להשיג את התשובות הנכונות גם אם לא התאמצו לחשוב על השאלות (על ידי שלילת הרמז האחרון המציע את התשובה), דבר העלול לפגוע בתהליך הלמידה שלהם (Cocca, Hershkovitz, & Baker, 2009); דוגמה אחרת היא למידה מרחוק (כגון בקורסים מקוונים המוונים - MOOCs) שאינה מחייבת כי האדם שמגיש את הפתרונות הוא אכן התלמיד עצמו, אף שהמערכת חושבת שכן (Alexandron, Ruipérez-Valiente, Chen, Muñoz-Merino, & Pritchard, 2017). באמצעות שימוש בנתונים הנאספים, ניתן לזהות התנהגויות אלו ואחרות.

mining and learning analytics: Applications to constructionist research. *Technology, Knowledge and Learning*, 19(1-2), 205-220.

Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. Washington, DC: SRI International, 1-57.

Cocca, M., Hershkowitz, A., & Baker, R. S. J. D. (2009). The impact of off-task and gaming behaviors on learning: Immediate or aggregate. *Frontiers in Artificial Intelligence and Applications*, 200(1), 507-514.

del Blanco, A., Serrano, A., Freire, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2013). E-Learning Standards and Learning Analytics. 2013 IEEE Global Engineering Education Conference, 1255-1261. <https://doi.org/10.1109/EduCon.2013.6530268>

EdSurge (2016). *Decoding Adaptive*. London: Pearson.

Hershkovitz, A., de Baker, R. S. J., Gobert, J., Wixon, M., & Pedro, M. S. (2013). Discovery With Models: A Case Study on Carelessness in Computer-Based Science Inquiry. *American Behavioral Scientist*, 57(10), 1480-1499.

Ho, A. (2017). *Advancing Educational Research and Student Privacy in the "Big Data" Era*. Washington, DC: National Academy of Education.

Lee, H. S., Xie, C., Nourian, S., Cheney, K. R., Moussavi, R., Gobert, J., Gweon, G., Damelin, D., Finzer, W., Lester, J. C., Wiebe, R. N., & Smith, A. (2015). Big data and learning analytics: A new frontier in science and engineering education research. Symposium presented at the annual meeting of National Association for Research in Science Teaching. Chicago, IL.

Kotsiantis S.B., Pierrakeas C.J., Pintelas P.E. (2003) Preventing Student Dropout in Distance Learning Using Machine Learning Techniques. In: Palade V., Howlett R.J., Jain L. (eds) *Knowledge-Based Intelligent Information and Engineering Systems. KES 2003. Lecture Notes in Computer Science*, vol 2774. Springer, Berlin, Heidelberg

Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, 57(10), 1380-1400.

del Blanco, Serrano, Freire, Martinez-) בשוק (Ortiz, & Fernandez-Manjon, 2013).

לאתגרים הללו יש כמובן גם היבטים כלכליים משמעותיים, שכן כרוכה בהם השקעה כספית ניכרת. כיוון שרוב ההשקעה בחינוך ולמידה היא ציבורית, ומבנה העלויות הוא ברובו קשיח, מציאת מודלים כלכליים ישימים למהפכת ה"ביג דאטה" בחינוך היא אתגר משמעותי.

לבסוף, לשימוש ב"ביג דאטה" היבטים אתיים משמעותיים הקשורים לאיסוף המידע ולשימוש בו. בתחום החינוך יש להיבטים אלו מספר מאפיינים ייחודיים (Ho, 2017). שניים כאלו קשורים לגיל המשתמשים ולחובת השימוש במערכת; ראשית, התלמידים בבתי הספר, שעל הדאטה המתאר את פעילותם אנו מדברים, הם קטינים, ועובדה זו הנה משמעותית כאשר בוחנים אם הם רשאים להסכים או להתנגד לאיסוף הנתונים. שנית, השימוש במערכות הלמידה מתבצע בדרך כלל כחלק מתכנית הלימודים, ולכן השאלה שעולה היא מה לעשות במקרה שבו תלמיד - או הוריו - אינו מעוניין שנתוניו ייאספו. לכן נדרשים כללים ברורים וישימים, המגובים בטכנולוגיה ובמתודולוגיות מתאימות, שיוודאו שהנתונים הנאספים מנוצלים בצורה הולמת ומשרתים אך ורק למטרת שיפור החינוך והלמידה.

לסיכום, למהפכת ה"ביג דאטה" יש פוטנציאל לתרום בצורה משמעותית לעולם החינוך והלמידה, אולם ההתקדמות בתחום דורשת התמודדות עם אתגרים מחקריים, פדגוגיים, טכנולוגיים, ואתיים מורכבים.

רשימת ספרות

Alexandron, G., Zhou, Q., & Pritchard, D. (2015). Discovering the Pedagogical Resources that Assist Students in Answering Questions Correctly - A Machine Learning Approach. *Proceedings of the 8th International Conference on Educational Data Mining*, 520-523.

Alexandron, G., Ruipérez-Valiente, J. A., Chen, Z., Muñoz-Merino, P. J., & Pritchard, D. E. (2017). Copying@Scale: Using Harvesting Accounts for Collecting Correct Answers in a MOOC. *Computers and Education*, 108, 96-114.

Berland, M., Baker, R. S., & Blikstein, P. (2014). Educational data