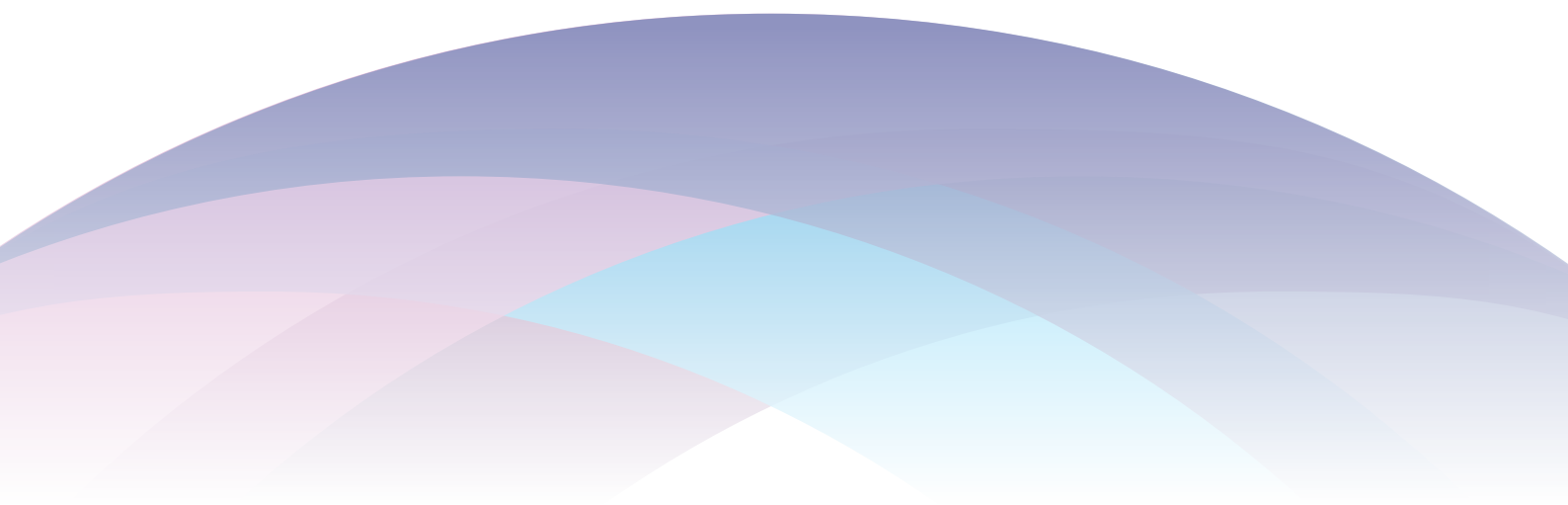

בינה מלאכותית בחינוך - היבטים של הוגנות ושוויון

סקירת ספרות מוגשת למדען הראשי, משרד החינוך

ד"ר חמי רמיאל | יוני 2023



תוכן עניינים



3	תקציר
4	הקדמה
5	פרק א' בינה מלאכותית בחינוך - מבוא כללי
9	פרק ב' הוגנות, הכלה ושוויון
13	פרק ג' תרומות אפשריות של בינה מלאכותית לקידום שוויון, הכלה והוגנות
17	פרק ד' טכנולוגיה חינוכית כמקדמת הוגנות בחינוך - ההבטחה שלא התממשה
20	פרק ה' הוגנות צרה בבינה מלאכותית בחינוך - חששות והתמודדויות
24	פרק ו' הוגנות רחבה - ביקורות על שימוש במערכות בינה מלאכותית
27	סיכום מבט מפוכח על השלכות בינה מלאכותית בחינוך על היבטים של הוגנות ושוויון
29	מקורות

תקציר

בינה מלאכותית (AI) הינה שם כולל למערכות מחשוב מתקדמות המסוגלות ללמוד את המציאות ולהמליץ על החלטות שונות בהתאם לכך. בינה מלאכותית בחינוך, כמו גם בשדות אחרים, מתפתחת במהירות לשלל כלים ויישומים מבטיחה להביא לשיפור וייעול תהליכים חינוכיים אך גם נושאת בתוכה מגוון אתגרים חברתיים, חינוכיים, פדגוגיים ואתיים. סקירה זו מתמקדת בפוטנציאל ובאתגרים של בינה מלאכותית ביחס לקידום הוגנות, הגינות, שוויון והכלה במערכות חינוך. לצד סקירה של התרומות האפשריות של בינה מלאכותית בחינוך לקידום ערכים אלו, הרי שהמסמך מציג שתי פרספקטיבות ביחס להיבטי הוגנות ושוויון בבינה מלאכותית בחינוך: (1) הוגנות צרה, המבטאת את ההשלכות המיידיות של מערכות קבלת החלטות והמלצות מבוססות בינה מלאכותית על קידום או פגיעה בהוגנות ושוויון ו(2) הוגנות רחבה, הדנה בהשלכות של ייצור, הפקה והשפעה של בינה מלאכותית על ערכי הוגנות ושוויון. הסקירה גם מתעכבת על בחינה של ההבטחה שלא התממשה של טכנולוגיות דיגיטליות אשר קדמו לבינה מלאכותית לקידום הוגנות ושוויון בחינוך ומצביעה על ההיבטים המקומיים וההקשריים של יישום והטמעה של טכנולוגיות בחינוך והנטייה של טכנולוגיות אלו לחזק ולהעצים אי-שוויון, אפלויות והטיות הקיימות במערכת. כלל הסקירה מצביעה כי על אף הפוטנציאל הקיים בינה מלאכותית לא יכולה להוות פתרון בפני עצמה לסוגיות של הוגנות ושוויון בחינוך אלא צריכה להשתלב כחלק ממדיניות חינוכית כוללת. זאת ועוד מהסקירה עולה, כי באם בינה מלאכותית תיושם בחינוך תחת מדיניות שאינה מקדמת הוגנות ושוויון, הרי יש חשש סביר שהיא תביא דווקא לפגיעה בערכים אלו ותחזק הטיות ואפלויות.

הקדמה

מסמך זה יתמקד בשאלה האם וכיצד שימוש בבינה מלאכותית במערכות חינוך (AIED) יכול לקדם הוגנות, הכלה ושוויון או לחלופין עלול להוות אמצעי הפוגע במגמות ובשאיפות אלו. המחקר והיישומים של בינה מלאכותית בכלל תחומי החיים והמדע ומתפתחים במהירות גבוהה ומשנים את הדרך שבה בני אדם צורכים מידע ומקבלים החלטות. לצד זאת, הרי שההשפעות החברתיות של התקדמות זו בפיתוח ושימוש בבינה מלאכותית עדיין לוטות בערפל ולצד תקווה ואופטימיות גדולה הרי שישנם רבים החוששים משימושים לרעה בבינה מלאכותית, בהעצמתה של אפליה ופגיעה בקבוצות מיעוט, בפגיעה בזכויות אדם כמו פרטיות וחופש תנועה, בריכוז כוח והון בקרב מספר מועט של תאגידי ענק, בפגיעה בסוכנות האנושית ועוד. לשאלות אלו יש תוקף מיוחד לגבי שדה החינוך ובעיקר לחינוך הבית הספרי (k-12) אשר יש לו תפקיד ציבורי וחברתי מהמעלה הראשונה ואשר עוסק בילדים, אשר פגיעים במיוחד להטיות ואפליות ואשר נזקקים ליותר מוגנות מצד המבוגרים.

העיסוק בבינה מלאכותית בחינוך הוא לא חדש ובינה מלאכותית בחינוך מתקיים כשדה מחקר פעיל לפחות משנות ה-80 של המאה הקודמת (Williamson and Eynon, 2020). יחד עם זאת בשנים האחרונות פריצות דרך טכנולוגיות משמעותיות בתחום הבינה המלאכותית, עליית תעשיית ה-EdTech ודחיפה מצד מדינות ומצד ארגונים בינלאומיים, הביאו לעליה משמעותית ביישומים של בינה מלאכותית במערכות חינוך, בתהליכי למידה והוראה ובהטמעתם בסביבות שונות. יישומים אלו עדיין לא הגיעו לכדי בשלות ולצד הצלחות ישנם לא מעט חסמים והתנגדויות. גם תחום המחקר וההערכה של יישומי בינה מלאכותית במערכות החינוך עדיין לא הגיע לשלב המאפשר להעריך את היישומים השונים מהיבטים שונים ובהקשרים מגוונים (Miao et al., 2021).

ספרות המחקר העוסקת בשאלות של בינה מלאכותית בחינוך נחלקת לשתי זירות וקהילות מחקר שכמעט ולא מדברות זו עם זו: אחת בוחנת בעיקר את היעילות של יישומים שונים ושותפה בפיתוחם של יישומים אלו והשנייה מתבוננת מנקודות מבט שונות על יישומי בינה מלאכותית בחינוך ומדגישה שאלות והקשרים רחבים, תוך שילוב של תיאוריות תרבותיות, משפטיות וחברתיות (Selwyn, 2021). יחד עם זאת ישנה הסכמה בין שתי קהילות המחקר על כך שבינה מלאכותית בחינוך אינה מעמידה רק שאלות טכנולוגיות ופדגוגיות, אלא גם משולבת עם שאלות רחבות יותר של אתיקה ופרטיות, של תפקידי המורות והמורים ושל יחסי חינוך ציבורי עם חברות וארגונים פרטיים המקדמים את השימוש בבינה מלאכותית, מפתחים ומפעילים מערכות כאלה, ופועלים על פי אינטרסים כלכליים שונים.

מסמך זה מתמקד בשאלות ספציפיות של הוגנות, הכלה ושוויון. יחד עם זאת הרי שהנחת היסוד של סקירה זו שלמעשה לא ניתן לנתק בין השאלות השונות ושהוגנות, הכלה ושוויון אינה מתבטאת רק בתוצאות צרות של הישגים או בהיבטים של נגישות, אלא גם בהוגנות חברתית רחבה ובשמירה על זכויות אדם.

מבנה הסקירה הוא כדלהלן: בפרק הראשון יוצג מבוא כללי לבינה מלאכותית בחינוך, כולל הגדרה של בינה מלאכותית וההקשרים החינוכיים שלה. בפרק השני נדון בהגדרות קצרות של מושגי ההוגנות והשוויון והדרך שבה הם מתבטאים ורלוונטיים להקשר של בינה מלאכותית בחינוך. הפרק השלישי יציג וידון בתרומות האפשריות של בינה מלאכותית בחינוך להיבטים של הוגנות ושוויון. בפרק הרביעי נבקש ללמוד מההיסטוריה של שימוש בטכנולוגיות חינוכיות דיגיטליות והשפעתן על קידום או פגיעה בערכי הוגנות ושוויון במערכת החינוך. בפרק החמישי נחזור לעסוק בבינה מלאכותית בחינוך ונציג את החששות המרכזיים הקיימים בקשר לפגיעה מתוך שימוש בכלים ובמערכות אלו בערכי הוגנות ושוויון וכן נפרט מספר עקרונות התמודדות העולים בספרות המחקרית. בפרק השישי תורחב היריעה ויוצגו גם חששות לפגיעה בהוגנות מתוך מבט רחב על הייצור, ההפקה והשימוש בבינה מלאכותית בחינוך.

פרק א' | בינה מלאכותית בחינוך - מבוא כללי

מהי בינה מלאכותית

ישנן הגדרות ופרספקטיבות שונות למה היא בינה מלאכותית (Artificial Intelligence - AI) ובמה היא מתייחדת מטכנולוגיות אלגוריתמיות אחרות. הגדרות אלו הן מורכבות ומציגות היבטים שונים. אחת ההגדרות המפורטות מציינת כי "בינה מלאכותית היא מערכת תוכנה (Software), המוטבעת בחומרה (Hardware) (כלומר היא מעורבת במעבדים, סנסורים ודשבורדים) שתוכננה על ידי בני אדם, שבהינתן מטרה מורכבת, היא מסוגלת לקבל החלטה על סמך תהליך של תפיסה, פרשנות והיגיון המבוסס על נתונים שנאספו על הסביבה" (Dignum, 2021, p.2). הגדרה זו מוסיפה שני מאפיינים, שבמשותף מייחדים מערכות בינה מלאכותית (שם):

- **אוטונומיה** - כלומר המערכת מסוגלת לפעול ולהגיע להחלטה, הכרעה או המלצה ללא שליטה חיצונית וזאת על בסיס נתונים ומטרות שהוזנו לה.
- **אדפטיביות** - כלומר המערכת אוספת נתונים על סביבתה ועל המשתמשים בה ומעדכנת על פיהם את הפרמטרים השונים בקבלת ההחלטות שלה.

בהגדרה צרה יותר אפשר לראות בינה מלאכותית כמערכת ממחושבת המתייחדת ממערכות אלגוריתמיות אחרות בכך שהיא מסוגלת לבצע פעולות מתקדמות שהיא לא תוכנתה אליהם באופן ישיר. בעיקרן מערכות אלגוריתמיות הן מערכות שפועלות על סמך תנאים ותוצאות - If... then.... גם כאשר מדובר במערכות מורכבות ביותר הרי שהתוצאה או ההחלטה שמערכות אלו מקבלות היא תוצאה של התנאים וההעדפות שתוכנתו אליה מראש. לעומתן מערכות בינה מלאכותית הן מערכות המסוגלות בדרכים שונות "ללמוד את המציאות" ולקבל החלטות על סמך למידה זו (ברק-מדינה, 2020).

מכאן שחלק מההגדרות של בינה מלאכותית מתייחסות לדרכים שבהן מערכות בינה מלאכותית מדמות חשיבה "כמו אנושית" שאינה ידועה מראש. לפי הגדרה זו בינה מלאכותית היא מערכת ממוחשבת שעוצבה על מנת להגיב לעולם באמצעות יכולות שנחשבות לאנושיות (Luckin et al., 2016) או "מכונה המסוגלת לחקות חלק מהתפקודים של אינטליגנציה אנושית, כולל רכיבים כמו תפיסה, למידה, הסקה, פתרון בעיות, אינטראקציה לשונית ואפילו יצירת עבודות יצירתיות" (Miao et al., 2021). חשוב לציין, כי גם הגדרות אלה לא טוענות שמדובר בחשיבה אנושית ממש אלא בחיקוי או דימוי של חשיבה אנושית נתונים הפועלים בתוך תחום ספציפי. למרות שזה עשוי להיראות "אינטליגנטי", אינטליגנציה זו קיימת רק בתוך נושא ספציפי או בשילוב עם מערכת כללים והיא בסופו של דבר מבוססת על זיהוי דפוסים ולא על חשיבה או הבנה (Holland, 2020). בסופו של דבר בינה מלאכותית היא פעולה סטטיסטית (מורכבת ומתוחכמת) של זיהוי דפוסים וסבירויות (Leaton Gray and Kucirkova, 2021; Miao et al., 2021). כך, גם כאשר מערכת בינה מלאכותית מייצרת פעולה שנראית מפתיעה, יצירתית או מתוחכמת שהמתכנתים של המערכת לא יכלו לחזות אותה מראש הרי שלפחות בטכנולוגיית הבינה המלאכותית הקיימת כיום, פעולה זו היא ביטוי של זיהוי דפוסים ולא מבטאת "הבנה אנושית" (Hosnagar, 2019; Marcus, 2022).

אפשר לחלק את הבינה המלאכותית לארבעה סוגים של פעולות (או תמות), לרוב משולבות זו בזו (Holland, 2020):

1. **זיהוי** - מערכות טכנולוגיות שמסוגלות לזהות ולהבין את העולם באמצעות חיישנים, לדוגמה זיהוי של חפצים, מרחקים, תווי פנים או צלילים.
2. **למידה** - מחשבים לומדים מנתונים המבוססים על מסקנות סטטיסטיות, לדוגמה, לומדים לזהות סוגים שונים של חפצים או מילים.
3. **ייצוג והנמקה** - מערכות בינה מלאכותית או סוכנים חכמים מייצרים ומשתמשים במודלים או בייצוגים שונים ולאחר מכן מיישמים אלגוריתמים על המודלים הללו כצורה של "היגיון". לדוגמה, אחרי שהמערכת זיהתה מילים ולמדה את השימושים השונים שנעשים במילים שונות במשפט, היא יכולה ליצור מודלים של משפטים מסוגים שונים.
4. **יצירת אינטראקציה כמו-טבעית (כלומר שמורגשת כטבעית) עם בני אדם**, לדוגמה, בוטים המתקשרים עם בני אדם באופן שמדמה תקשורת אנושית.

יש להבחין בין היכולות והפעולות של בינה מלאכותית ובין הטכנולוגיות המאפשרות אותן. בינה מלאכותית היא אוסף של טכנולוגיות שונות שהמרכזיות שבהן הן למידת מכונה (Machine learning) ולמידה עמוקה (Deep learning). למידת מכונה היא קבוצה של אלגוריתמים המאפשרים למערכת מחשב לבצע אופטימיזציה ההולכת ומשתפרת על בסיס הצטברות של נתונים חדשים וכלים סטטיסטיים. למידת מכונה מתייחסת לתהליך אימון בינה מלאכותית על ידי פיתוח תגובות אלגוריתמיות המבוססות על קבוצה של כללים מתוכנתים בהקשר של מערך נתונים (Holland, 2020). בלמידת מכונה המכונה לומדת מנתונים קיימים ומנתוני אימון, מייצרת מודל לניבוי ומפעילה אותו על נתונים חדשים. למידה עמוקה היא מחלקה של אלגוריתמים אשר מזהה דפוסים ותכונות באמצעות שכבות שונות של רשתות נוירונים (העומק מתייחס לא למידת ההבנה של המערכת אלא לשכבות השונות של הרשת). טכנולוגיות אלו בשילוב עם טכנולוגיות אחרות מאפשרות מגוון של יכולות מורכבות כמו זיהוי שפה טבעית (NLP - Natural language processing), זיהוי דיבור (Speech recognition), זיהוי ועיבוד תמונה (Image recognition and processing), סוכנים אוטומטיים, זיהוי רגשות, יצירה מלאכותית ועוד.

אחת התכונות המרכזיות המשותפת לכלל כלי הבינה המלאכותית היא ההתבססות והתלות בנתונים ובמסדי נתונים קיימים על מנת לבצע את פעולתה. במקרה של רוב כלי למידת המכונה נעשה שימוש בנתוני אימון שמאפשרים לאלגוריתמים ללמוד את הנתונים הקיימים ולייצר קלסיפיקציות, ניתוחים ותחזיות. נתוני האימון מערבים, בדרך כלל, פעולה אנושית של תיוג המאפשרת למכונה ללמוד. במקרה של כלי למידה עמוקה וכלים אחרים, ריבוי הנתונים מאפשר לבצע ניתוחים שונים. ההתבססות על נתוני-עתק (Big data) מחייבת השגה של נתונים אלו אם באמצעות שימוש במסדי נתונים קיימים ואם באמצעות איסוף נתונים ומעקב (Pedro et al, 2019). בתוך כך, ישנם כלי בינה מלאכותית כמו כריית נתונים (Data mining) שכל מהותם היא תהליך של גילוי דפוסים ויחסים שימושיים ומעניינים בכמויות גדולות של נתונים (Fischer et al., 2020).

התלות של בינה מלאכותית בנתונים מסמנת את היותה של בינה מלאכותית מערכת שמבוססת על מעקב ואיסוף נתונים אינטנסיבי ומתמשך על בני אדם והתנהגותם. מערכות בינה מלאכותית שואפות להרחיב ככל הניתן את איסוף הנתונים שלהן, על מנת לייצר את התחזיות המדויקות ביותר. ההנחה שקיימת בקרב חלק ממפתחי וחוקרי הבינה המלאכותית היא שככל שמסדי הנתונים יגדלו כך מערכות הבינה המלאכותית יוכלו לייצר החלטות והמלצות מדויקות יותר, לעומתם יש הטוענים כי מנקודה מסוימת תוספת של נתונים היא שולית לשיפור תוצאות כלי הבינה המלאכותית (Marcus, 2022).

בסופו של דבר בינה מלאכותית אינה תהליך אוטומטי או מבוסס מכונה בלבד והוא מערב בני אדם והחלטות אנושיות בכל שלב שלב. המעורבות האנושית אינה רק בשלב היישום והשימוש בפועל, אלא

גם כחלק מתהליך "יצירת הבינה" עצמה. בני אדם מגדירים את הבעיה ואת התוצאות הרצויות, בוחרים את סט הנתונים ומבצעים מניפולציות שונות עבור סט הנתונים הנבחר, פעמים רבות בני אדם מעורבים בסימון הנתונים ובהבחנות בדבר נתוני האימון, וכמובן בני אדם מעורבים בעיצוב ובניית האלגוריתמים. פעמים רבות בשיח על בינה מלאכותית יש נטייה להתעלם מהעבודה האנושית לאורך הדרך ומההשלכות הן של ההטיות האנושיות שבה, מהיבטים של ניצול שיכולים להיות בה ומאינטרסים אנושיים, כלכליים וחברתיים שונים שמעורבים בהטמעה ושימוש של מערכות אלו (Crawford, 2019).

יישומים והשלכות של בינה מלאכותית בחינוך

בהסתמך על שלל הדוחות, הפרסומים והמחקרים אפשר לסמן שלוש קבוצות מרכזיות של שימוש קיים, מתפתח ועתידי בבינה מלאכותית במערכות חינוך (בדגש על מערכות בתי ספר) ברחבי העולם (החלוקה בוצעה על פי Miao et al., 2021):

1. בינה מלאכותית **מרוכזת למידה (או מרוכזת לומד)** - תחת קבוצה זו נכלול מערכות שעניינן ייעול ושיפור הלמידה של התלמיד האינדיבידואלי ומתרכזות בתהליכי הלמידה בתחומי ידע ספציפיים ובמיומנויות למידה ספציפיות כדוגמת צ'טבוטים המסייעים ללמידה ומערכות למידה אדפטיביות.
2. בינה מלאכותית **מרוכזת הוראה (או מרוכזת מורה)** - תחת קבוצה זו נכלול מערכות שאמורות לסייע למורה/ים בתהליכי ההוראה, כדוגמת כלי הערכה אוטומטיים, מערכות לניטור קשב ומצב רגשי של תלמידים או מערכות לזיהוי העתקות.
3. בינה מלאכותית **מרוכזת ניהול וארגון (או מרוכזת מערכת)** שעניינה זיהוי דפוסים וסיוע בקבלת החלטות לגבי תלמידים ומשאבים כדוגמת מערכות לניהול למידה או מערכות לסיוע במיון ובחירה.

היישומים הם רבים ומגוונים והצפי הוא שיישומי בינה מלאכותית בחינוך יתפתחו בכיוונים שונים במהלך השנים הקרובות, אך קשה להעריך מה יהיו השימושים הללו וכיצד הם ייושמו וכמובן עוד יותר קשה להעריך מה תהיה ההשפעה של כלים אלו על מורות ותלמידים. אפשר לראות במערכות בינה מלאכותית עוד שלב ביישום של טכנולוגיה דיגיטלית במערכת החינוך, אבל רבים טוענים כי בינה מלאכותית היא תופעה חדשה ונפרדת, בין השאר מפני שהיא מבטאת רכיב משמעותי בגדילה הולכת וגוברת של קבלת החלטות אוטומטית (Automated decision-making), כלומר קבלת החלטות שאינה מבוצעת על ידי בני אדם, אלא על ידי מערכות ממוחשבות (Selwyn, 2022). בהקשר החינוכי קבלת החלטות אוטומטית יכולה להתקיים בהחלטה מהו התרגיל במתמטיקה שיינתן לתלמיד ספציפי ברגע נתון, כיצד בוט יגיב ויתקשר עם תלמידה החווה קושי או בהחלטה רחבה יותר הנוגעת למיון תלמידים. גם אם קבלת החלטות אוטומטית אפשרית גם על בסיס מערכת אלגוריתמית ללא בינה מלאכותית, הרי שההבדל טמון ביכולת של בינה מלאכותית לא להסתמך רק על קלסיפיקציות וקטגוריות ידועות מראש שהוגדרו על ידי סוכן אנושי, אלא לייצר בעצמה את הקריטריונים והאיזון שבין המשתנים השונים ביחס לכל בחירה ובחירה.

העיסוק בשאלת ההשפעה של בינה מלאכותית בכלל, ובתוך זה שאלות של אתיקה, ושל הוגנות ושוויון מעסיקה ארגונים וממשלות ברחבי העולם. בתוך כך מוצעות מסגרות ברמת המדיניות המדינתית והגלובלית וברמה הרגולטיבית כדרך להתמודד עם האתגרים שמציבות מערכות בינה מלאכותית (לדוגמה: Hosanagar, 2019). השדה החינוכי מציג כמה רמות של מורכבות ביחס לשאלות אלו. ראשית, שדה החינוך הוא שדה ייחודי מכיוון שהוא עוסק בילדים, שלכאורה היכולת שלהם להבין ולהסכים לסוגים שונים של יישומים של מערכות אלו היא מוגבלת, במיוחד לאור זאת שמערכת החינוך היא מערכת ציבורית המחייבת את כלל התלמידים. שנית, מערכות בינה מלאכותית שאוספות נתונים כל הזמן, עלולות לפגוע בזכות של תלמידים ותלמידות לפרטיות ובזכות להישכח (Right to be forgotten) או להתחיל מדף

חלק בשלבים שונים של ההתבגרות שלהם. שלישית, מערכות בינה מלאכותית מייצרות תחזיות על בסיס נתונים קיימים, כך הן עלולות להסליל תלמידים למסלולים ידועים מראש ולא לאפשר רמות שונות של בחירה ואוטונומיה (Bristol and Shawn, 2022).

שאלות אלו באות לידי ביטוי כשלווקחים בחשבון את המטרות הרבות של המערכת החינוכית, שלעיתים הן מטרות סותרות. מערכות בינה מלאכותית יודעות לייצר יעילות ביחס למטרה מסוימת - לדוגמה לייצר תחזית באשר ליכולת של תלמיד להתמודד עם ידע מסוים במתמטיקה. אבל לצד רכישת ידע המטרות של מערכת החינוך מורכבות גם מפיתוח מיומנויות שונות, ביניהן גם כישורים חברתיים-רגשיים, טיפוח זהות עצמית, טיפוח אזרחים תורמים לחברה ועוד. כיצד כל אלה יבואו לידי ביטוי במערכי קבלת החלטה טכנו-אנושיים (כלומר המשלבים אלגוריתמים, בינה מלאכותית ובני אדם) היא שאלה רחבה ומאתגרת. בנוסף מערכות חינוך נתפסות כבעלות תפקיד מהותי ומשמעותי בחברה. אלו מערכות ציבוריות, שבאים בהן לידי ביטוי אינטרסים רבים של שחקנים חברתיים שונים, ושעוברת עליהן ביקורת תמידית מכיוונים שונים. לאור זאת שאלת היישום והטמעה של מערכות בינה מלאכותית תהיה תמיד שאלה מורכבת הנוגעת לאיזון בין ערכים ומטרות שונים. לכך יש להוסיף שמסיבות שונות מערכות חינוך נוטות להיות שמרניות ותהליכי הטמעה של טכנולוגיות חינוכיות הם בדרך כלל איטיים, וכמו שיפורט בהמשך גם כאשר טכנולוגיות חינוכיות שונות הוטמעו במערכות חינוך הן לא נטו להשפיע השפעה דרמטית, אלא להשתלב במבנים הקיימים. במובן זה, כפי שקובע רייך (Reich, 2020): "על אף ההתלהבות מבינה מלאכותית, מערכות אלגוריתמיות אינן פתרונות קסם לפתרון בעיות בחינוך". במובן זה חשוב לראות בכלי בינה מלאכותית כלים שמשפיעים באופן מכריע על אנשים רבים, אבל לצד זאת פועלים באופן שהוא לא שקוף ובוודאי לא מובן לרוב השחקנים האנושיים ולצד היתרונות והיכולות שלהם עלולים גם לגרום לפגיעה בתלמידות ותלמידים (Kizilcec and Lee, 2020).

פרק ב' | הוגנות, הכלה ושוויון

בהקשר של בינה מלאכותית בחינוך

מערכת החינוך והמערכת הבית ספרית נושאת בתוכה מטרות סותרות ומתחים שהם חלק בלתי נפרד ממדיניות חינוך ומעשייה חינוכית. אחד מהמתחים המרכזיים הללו הוא המתח שבין היחיד והחברה. בעוד הפעולה החינוכית מבקשת לפתח ולטפח כל אינדיבידואל הרי שהיא גם מחזיקה בתוכה שאיפה לתיקון חברתי המתבטאת במערכת הוגנת לכלל המשתתפים בה, המספקת הזדמנויות שוות לקבוצות מיעוט, קבוצות שסובלות או סבלו מאפליה ומהדרה, ליחידים בעלי מוגבלויות ועוד. לאור ערכים של דמוקרטיה, חופש ושוויון מערכות חינוכיות ברחבי העולם חרטו על דגלם הוגנות, הגינות, הכלה וצמצום פערים של אי-שוויון.

לצורך מסמך זה נשתמש ונבחין בארבעה מושגים מרכזיים על מנת להציג היבטים אלו:

- **הוגנות (Fairness)** מתייחסת לאופן שבו שונות בין תלמידים בתוצאותיהם היא שיקוף מדויק של הכישרים והידע ולא היבטים אחרים (סביבתיים, תרבותיים וחברתיים). במובן הזה הוגנות עוסקת באופן שבו המערכת החינוכית מייצרת תנאים שבהם כל תלמיד ותלמידה נבחן ומשיג את תוצאותיו ללא הטיות חיצוניות (Loukina et al., 2019).
- **שוויון/אי-שוויון (Equality/Inequality)** מתייחס לשאיפה לשוויון הזדמנויות בין כלל התלמידים והשותפים לחברה. מדדים של שוויון ואי-שוויון בוחנים את האופן שבו המשאבים השונים מחולקים בין קבוצות שונות באוכלוסייה את הדרך שבה חלוקת משאבים זו משפיעה על ההזדמנויות הפרוסות בפני אינדיבידואלים והנגישות שלהם להזדמנויות אלו.
- **הגינות (Equity)** צועדת צעד נוסף מעבר להוגנות ומשלבת תפיסות מורכבות של שוויון ואי-שוויון חברתי. כך בעוד יעדים של הוגנות ושוויון יסתפקו בשיפור בתוצאות ובהישגים בקרב כל שכבות האוכלוסייה, הרי שההגינות מתעניינת לא רק בשיפור עצמו אלא גם בשיפור שהוא יחסי המאפשר לצמצם את הפערים בין קבוצות שונות.
- **הכלה (Inclusion)** הוא מושג רחב יותר, המתפרש בהקשרים מדינתיים שונים באופנים שונים, ומעיקרו מתייחס לאופן שבו המערכת החינוכית תומכת בגיוון ורואה בכל אחד ממשתתפיה - בעלי מוגבלות, תלמידות מאוכלוסיות מיעוט, מהגרים וכדו' - כמי שיכול לתרום תרומה חשובה לסביבת הלמידה (Ydesen et al., 2021).

חשוב לציין כי מושגים אלה אינם מושגים מוחלטים (או מצבים מוחלטים), אלא תמיד מושגים יחסיים המתקיימים ופועלים בתוך הקשרים ספציפיים. במובן זה כל הערכה של שאלות של הוגנות, הגינות, שוויון והכלה בחינוך עומדת ביחס למטרות חינוכיות שונות ועומדת להכרעה של קובעי מדיניות ושל מישמה. הכרעות אלו הינן הכרעות פוליטיות, במובן שהן מבוססות על תפיסות עולם וסדרי עדיפויות שונים. שאלות של הוגנות, הגינות, הכלה ושוויון הינן תמיד שאלות של איזון בין יתרונות וחסרונות שונים. בהקשר זה מסבירים קיזילצ' ולי (Kizilcec and Lee, 2020) כי כל המשגה של הוגנות (fairness) היא מטבעה מתבססת על השוואה חברתית. כך, גם אם נמצא שטכנולוגיה חינוכית הביאה לשיפור בהישגים בקרב כל הקבוצות החברתיות, כולל קבוצות מיעוט וקבוצות מוחלשות, הרי שמה שקובע בעיניהם את מידת ההגינות (equity) של הטכנולוגיה החינוכית הוא היחסיות בשיפור של קבוצות מיעוט וקבוצות מוחלשות או מודרות. לפי תפיסה זו הערכה של הטמעת טכנולוגיה חינוכית בבתי הספר ובמערכות חינוך צריכה להתברר באמצעות השאלה האם הטכנולוגיה הובילה לצמצום הפער שהיה קיים לפני, להרחבת הפער או לשימורו.

בהקשר של טכנולוגיה חינוכית שאלות אלו של הוגנות והגינות מתבטאות ומתכנסות בדרך כלל סביב שלושה היבטים: גישה, שימוש ותוצאות (Warschauer and Matuchniak, 2010).

גישה (Access) מציינת את הדרכים שבהן הטכנולוגיה עצמה נגישה עבור התלמידים, הגישה מתארת לא רק נגישות למכשירים דיגיטליים, לפס אינטרנט רחב או לתשתיות טכנולוגיות בית ספריות, אלא גם להיבטים חברתיים ותרבותיים של נגישות כמו השכלת הורים, הכשרת מורים ועוד.

שימוש (Use) מתייחס לדרכים השונות שבהן בתי הספר ומורות עושים שימוש בכלים הדיגיטליים.

תוצאות (Outcomes) מתייחסות לאופן שבו טכנולוגיה חינוכית מקדמת תוצאות והישגים חינוכיים אך גם הוגנות וצמצום פערים.

מעבר לכך, נשאלת השאלה האם שאלת ההוגנות יכולה להיות מנותקת משאלות של צדק ואתיקה. לדוגמה, האם מוצדק להפעיל פרקטיקות שפוגעות בזכויות היסוד של תלמידים, גם אם מטרתה היא קידום הוגנות לתלמידים מוחלשים? כמובן שאין בכוונת מסמך זה להכריע בשאלות הללו, אלא להצביע על המורכבות של המושגים הללו והאופן שבו הם מתממשים במציאות המורכבת של מערכת החינוך המדינתית ובתי הספר.

הוגנות והגינות במערכות אלגוריתמיות ובמערכות בינה מלאכותית - בין הוגנות צרה להוגנות רחבה

להוגנות, הגינות, הכלה ושוויון יש כמה ממדים בהקשר של בינה מלאכותית: ברמה אחת הוגנות בבינה מלאכותית בחינוך מתבטאת בתוצאות ובהחלטות שמייצרת המערכת. ברמה אחרת הוגנות בבינה מלאכותית נבחנת בהקשרים רחבים יותר העוקבים אחר ההשלכות של הייצור, ההפקה והשימוש במערכות בינה מלאכותית על היבטים שונים של הוגנות, הגינות, הכלה ושוויון. במסמך זה נבחין בין הממדים ברמה הראשונה אותם נכנה **"הוגנות צרה"** ובין הממדים ברמה השנייה אותם נכנה **"הוגנות רחבה"**.

רבים רואים בבינה מלאכותית פוטנציאל חיובי רב בכל מה שקשור לממדים של "הוגנות צרה" (כפי שיפורט בהרחבה בפרק השלישי) וזאת לא רק מתוך היעילות שמערכות אלו מסוגלות לייצר, אלא גם מתוך היכולת שלהן לעקוף הטיות ואפילו אנושיות ולהביא לתוצאות ולהחלטות שבהן ההטיות והתפיסות המוקדמות של בני האדם מוחלשות. בנוסף, לבינה מלאכותית יכול להיות שימוש משמעותי בסיוע ותיווך לאנשים עם מוגבלות או חוסר במשאבים. בנוסף לכלים של בינה מלאכותית יש את הפוטנציאל לחשוף הטיות ואפילו שאולי לא ניתן למצוא אותם ללא שימוש בכלים אלו.

יחד עם זאת, בשנים האחרונות ספרות רבה מתמודדת עם האתגרים שמערכות בינה מלאכותית מציבות ביחס לסוגיות של הוגנות ובעיקר בהטיות ובאפילו שמערכות אלו מייצרות או משקפות (Kizilcec and Lee, 2020). אחד האתגרים המרכזיים הוא ההטיות שמערכות בינה מלאכותית עלולות לייצר, לשמר ולהעצים. חוקרים של הוגנות אלגוריתמית ושל הוגנות בבינה מלאכותית מבחינים בין צורות סטטיסטיות של הטיות וצורות חברתיות של הטיות. הטיות סטטיסטיות כוללת הטיית דגימה ושגיאות מדידה והטיות חברתיות מתייחסות למבנים חברתיים בעייתיים שמיוצגים בנתונים (Mitchell et al., 2021). במילים אחרות בעוד הטיות סטטיסטיות מתייחסות למערכים סטטיסטיים פגומים או לא מספקים, הרי שהטיות חברתיות קשורה לאופן שבו המערכת הסטטיסטית מתבססת על נתונים שמשקפים אפליה או הטיות ומהדהדת אותה בקבלת ההחלטות. באופן יותר מפורט אפשר לחלק את ההטיות הללו לעוד מספר תתי קבוצות המתרחשות בשלבים שונים של הפעולה של הבינה המלאכותית: הטיות היסטוריות, הטיית ייצוג, הטיית מדידה, הטיית הערכה, הטיית צבירה ועוד (Baker and Hawn, 2021).

הטיות אלגוריתמיות עלולות ליצור שני סוגים של פגיעות חברתיות: פגיעות מהקצאת משאבים הנובעות ממניעת הזדמנות או משאב כלשהו מקבוצות ספציפיות או חלוקה לא הוגנת של מוצר בין קבוצות ופגיעות ייצוגיות המתבטאות בייצוג שיטתי של קבוצה כלשהי באור שלילי, או בהיעדר ייצוג חיובי (Baker and Hawn, 2021). דוגמה מצוטטת רבות בהקשר זה היא של תוכנה מבוססת בינה מלאכותית אשר הייתה בשימוש מערכת המשפט במדינת פלורידה בארה"ב ומטרתה הייתה לסייע בקביעת סיכון חזרתיות אצל מי שביצעו פשע. התוכנה נטתה בסבירות גבוהה לסמן פי שניים נאשמים לבנים כ"סיכון נמוך" מאשר נאשמים שחורים, וכמעט פי שניים יותר לחזות פליליות עתידית בקרב נאשמים שחורים מאשר בנאשמים לבנים. הסיבה להטיה של המערכת נבעה בעיקר מהתבססות על נתוני עבר ומההטיות המבניות והחברתיות שבוטאו בנתונים אלו (Hosnagar, 2019).

ענף זה של דיון על הוגנות של מערכות אלגוריתמיות ומערכות בינה מלאכותית מתמקד בהוגנות כמושג סטטיסטי. בהקשר זה ישנם הגדרות שונות ורבות של הוגנות (Fairness) במדעי המחשב ובסטטיסטיקה (Narayanan, 2018; Loukina et al., 2019; Verma and Rubin, 2018). הוגנות סטטיסטית מהסוג הזה ניתנת לתקנון או להתמודדות על ידי התערבות סטטיסטית, על ידי שיפור באלגוריתמים ובמערכים הסטטיסטיים כמו גם על ידי שיפור ותקנון במסדי הנתונים או בנתוני האימון (ליתר פירוט ראו פרק 5). תיאורטית, בהנחה שכלל הגדרות ההוגנות השונות מושגות אפשר להגיע למצב של הוגנות כוללת (כלומר מושלמת מבחינה סטטיסטית) שבה אף הטיה לא התקיימה, יחד עם זאת, במערכות אלגוריתמיות מורכבות, ביניהן מערכות בינה מלאכותית, הוגנות כוללת לא באמת יכולה להיות מושגת, לאור ריבוי התנאים וההטיות הסטטיסטיות האפשריות (Verma and Rubin, 2018).

הוגנות אלגוריתמית זו המבקשת למנוע הטיות או אפילו על ידי התערבויות סטטיסטיות ואלגוריתמיות אינה עוסקת בשאלות רחבות יותר כמו האם התהליך עצמו תקין והוגן, אלו הטיות, אפילו ופגיעות עלולות להתרחש במהלך הייצור, ההפקה, ההטמעה והשימוש במערכות בינה מלאכותית (Kizilcec and Lee, 2020) ואינה מתייחסת לסוגים אחרים של התערבויות כמו קביעת מדיניות טובה יותר, רפורמה במוסדות חברתיים או שינויים חברתיים עמוקים (Barocas et al., 2017). סוג הוגנות זו מתמקד במובן הצר של קבלת החלטות אוטומטית הוגנת, ללא הטיה שנוצרת מהפעולה האלגוריתמית או מהטיה בנתונים המשמשים לפעולת המערכת (Barocas et al., 2017).

לעומת סוג הוגנות זו הרי שבשנים האחרונות התפתחה כתיבה רבה המתבוננת על היבטים רחבים של הוגנות בהקשר של מערכות אלגוריתמיות ומערכות בינה מלאכותית ועל הנזקים והפגיעות החברתיות שהם עלולים לשאת איתם (ושתוצג ביתר הרחבה בפרק השישי). ספרות זו מדגישה את ההיבטים הרחבים יותר של הוגנות, הגינות, הכלה ושוויון בבינה מלאכותית בכלל ובחינוך בפרט הנוגעים לא רק להטיות סטטיסטיות אלא למכלול הטכנו-אנושי המרכיב את הבינה המלאכותית: שאלות של הוגנות ושוויון הקשורות לתשתית החומרית והאנושית הדרושה עבור מערכות בינה מלאכותית, למערכות ולתוצאות הכלכליות השלובות בבינה המלאכותית, להשלכות של בינה מלאכותית על פיקוח ומעקב אחר אזרחים, למקום של הסוכנות והעצמאות האנושית בעולם רווי בינה מלאכותית, ולחלוקה של משאבים כתוצאה ועל בסיס בינה מלאכותית.

מבט על הוגנות, הגינות, הכלה ושוויון מעבר לזווית הראייה הצרה המתמקדת בהטיות שעלולות להיווצר על ידי מערכות בינה מלאכותית היא בעלת חשיבות, מכיוון שאלו יעדים חברתיים ופוליטיים שאינם יכולים להיות מצומצמים רק לבחינה של תוצאות והישגים אלא מכילים בתוכם יחסים וערכים חברתיים וזאת במיוחד בהקשר של המעשה החינוכי המהווה חלק מהותי מהמחויבות של המדינה כלפי אזרחיה ובפיתוח וטיפוח של אזרחים פעילים בחברה דמוקרטית, צודקת ובת קיימא. בתוך כך חוקרים מבקשים לחשוף את האופן שבו בינה מלאכותית בחינוך משולבת בתוך יחסים חברתיים ותפיסות חינוכיות (Williamson and Eynon, 2021), מעמידים בספק את התפיסה הרואה בבינה מלאכותית כפעולה אובייקטיבית וניטרלית ומדגישים כיצד כל הטמעה של מערכות בינה מלאכותית במערכת החינוך משולבת באינטרסים אנושיים

שונים וביחסי כוח חברתיים (Perota and Selwyn, 2019, Selwyn, 2020). כך על אף הפוטנציאל של בינה מלאכותית עבור מערכות חינוך גם בתחום של הוגנות ושוויון, הרי ששאלה זו תמיד צריכה להישאל ביחס להשלכות על זכויות האדם הבסיסיות של מורים ותלמידים ופגיעות אפשריות בהם (Berendt et al., 2020). במילים אחרות, האם שימוש במערכת בינה מלאכותית - שלא רק שמקדמת יעילות חינוכית אלא גם מאפשרת מוביליות חברתית ומהווה כלי להגינות - מוצדק גם כאשר היא פוגעת בדרך בזכויותיהם של מורים או תלמידים?

פרק ג' | תרומות אפשריות של בינה מלאכותית לקידום שוויון, הכלה והוגנות בחינוך

בינה מלאכותית בחינוך נושאת איתה הבטחה לשיפור מערכות חינוך ושיפור הלמידה על בסיס התאמה אישית ותהליכי פרסונליזציה, שיפור ויעול בקבלת ההחלטות והבנה וחיזוי של תהליכי הלמידה והחינוך. ההבטחה הזו המפורטת במספר רב של דוחות, ארגונים ומדינות כוללת בתוכה גם את ההבטחה (או את הטענה) בדבר היכולת והפוטנציאל של מערכות בינה מלאכותית בחינוך לקדם את ערכי ההוגנות, ההגינות והשוויון (ברק -מדינה, 2019, משרד החינוך, 2022; Luckin et al., 2016; Pedro et al., 2019). כפי שמצוין, במספר דוחות (לדוגמה הדוח של אונסקו - Miao et al., 2021) על אף הפוטנציאל הקיים של בינה מלאכותית בחינוך ישנם מכשולים רבים במימושה, ובלתי אפשרי עדיין לקבוע את השפעתה על תלמידים, ולבחון שיפור של יעילות חינוכית ופדגוגית (ראה גם Bulathwela et al., 2021).

ההבטחה על קידום ערכי הוגנות ושוויון באמצעות בינה מלאכותית בחינוך מבוססת מעיקרה על טענה מרכזית אחת המתמקדת ביכולת של מערכות בינה מלאכותית לספק תהליכי הוראה ולמידה מותאמים אישית ויעילים, מה שקרוי פרסונליזציה. יכולת זו כוללת גם הערכה אפקטיבית וזיהוי מוקדם של קשיים שונים, נקודות חוזקה וחולשה. פרסונליזציה, כך נטען, תקדם ערכים אלו מכמה סיבות: ראשית, מכיוון שהיכולת של הוראה מותאמת אישית מבוססת בינה מלאכותית מובילה או טוביל לשיפור בהישגים של תלמידים (Hamilton and Hatie, 2021). שנית, כי היא תאפשר תמיכה מותאמת אישית וגמישה בתלמידים עם צרכים שונים. היתרון השלישי נוגע לסוגיית המשאבים: הן ביכולת להציע שירותים חינוכיים במחיר נמוך ובר השגה תוך התפתחות בקנה מידה רחב (scaling) והן ביכולת להתגבר על העדרם של משאבים מקומיים (כדוגמת חוסר במורים או חוסר בכיתות ראויות) ולבסוף גם בהורדת עומס העבודה ממורים ופינוי זמן ומשאבי הוראה, בין השאר לצורך תמיכה בתלמידים חלשים.

באופן יותר מפורט אפשר לראות את היכולות, היתרונות והפוטנציאל של בינה מלאכותית בחינוך בקידום ערכי הוגנות ושוויון דרך שלושת היבטים של יישום טכנולוגיה חינוכית - גישה, שימוש ותוצאות:

גישה - מערכות בינה מלאכותית, במיוחד כאלה המבוססות על התאמה אישית, מאפשרות ומקדמות למידה מרחוק, בכך הן מסייעות בהגעה למקומות מרוחקים, בהם יש חוסר במשאבים. באופן כללי היא יכולה לספק גישה להזדמנויות למידה מתאימות לאנשים וקהילות בשוליים (marginalized), לאנשים עם מוגבלות, לפליטים, לתלמידים מחוץ לבתי ספר ולתלמידים שחיים בקהילות מבודדות (Pedro et al., 2019). היבט הגישה נוגע גם ליכולת לצבור וליצור משאבי למידה חוצי שפה (בעזרת כלים של NLP ואחרים) המאפשרים להנגיש ולהפיץ משאבי למידה לקבוצות מוחלשות או פריפריאליות (Bulathwela et al., 2021).

שימוש - בהקשר של שימוש, הדוחות השונים מציינים בעיקר היבטים של יעילות ושיפור חוויית הלמידה, תוך הערכה אפקטיבית ובזמן אמת, המאפשרת תמיכה בתלמידים מסוגים שונים. היבט חשוב שמודגש בדוחות הוא קידום למידה מבוססת נתונים. היבט זה כולל גם שימוש בידע מדעי בכלי בינה מלאכותית על מנת לאפיין קשיים בלמידה ולתמוך בהתמודדות עם קשיים אלו. וכן לחזות תהליכי חשיבה של לומדים בודדים על מנת לזהות את קשיי הלמידה שלהם ולספק להם פתרונות למידה מותאמים ומדויקים מסוגלות לזהות תלמידים מתקשים, להציע להם סיוע אוטומטי או לחילופין לסמן עבור סוכנים אנושיים את הקשיים הללו. המערכות הללו יכולות גם לחזות, באופן שמאפשר פעולה מקדימה, אלו תלמידים

צפויים להתקשות. היבט נוסף שמפורט בהקשר של שימוש הוא סיוע למורים הן באמצעות ההתאמה האישית לתלמידים, הנגישות של משאבי הוראה, הורדת עומס העבודה על ידי העברה של משימות רוטיניות למערכות אוטומטיות ופינוי זמן למשימות הוראה בעלות ערך מוסף גבוה, בדגש על תלמידים מתקשים או בעלי צרכים מיוחדים (Chen et al., 2020).

תוצאות והערכה - מעבר לפוטנציאל של שיפור ההישגים המבוסס על התאמה אישית ויעילות, הרי אחת ההבטחות של מערכות בינה מלאכותית היא שיפור ההערכה באופן שמקדם הערכה מעצבת המובילה בתורה לשיפור בהישגים. בתוך כך מערכות בינה מלאכותית יכולות לספק כלים אוטומטיים להערכה וציון, הערכה אוטומטית של מטלות ומשימות ובעיקר הערכה בזמן אמת.

בשולי הדברים ניתן להוסיף את התרומה האפשרית של מערכות בינה מלאכותית לזהות הטיות ואפליות מערכתיות וממסדיות, כאשר הן מכוונות לכך. כך לדוגמה, השתמשו חוקרים בלמידת מכונה על מנת לנבון במאגרי מידע רפואיים (תיקים רפואיים דיגיטליים - EHR) של כ- 20,000 חולים בשיקגו וגילו הטיות שליליות מצד הצוות הרפואי כלפי שחורים, אנשים עם ביטוח ציבורי ורווקים (Sun et al., 2022).

אפשר לראות שישנן מספר הנחות שעומדות בבסיס ההבטחה של בינה מלאכותית בחינוך לקדם הוגנות במערכות חינוך. ההנחה הראשונה קשורה להרחבה ונגישות של הוראה אחד על אחד (עם סוכן אינטליגנטי או עם מערכות הוראה מותאמות אישית) לקהלים רבים ומגוונים (Holstein and Doroudi, 2021). לפי הנחה זו הוראה אחד על אחד, מותאמת אישית מבוססת על כלי בינה מלאכותית, לא רק שתהיה יעילה יותר ומותאמת יותר ברמת הפרט, אלא גם תאפשר להתגבר על צוואר הבקבוק של משאבים אנושיים ואחרים המונעים ממערכת החינוך לספק הוראה עדיפה שכזו לכלל התלמידים ובמיוחד לתלמידים מוחלשים. ההנחה השנייה, למעשה ממשיכה קו זה, ורואה את הפוטנציאל של בינה מלאכותית בחינוך לצמצם פערים קיימים בשירותי החינוך הקיימים, בין אם על ידי כלי בינה מלאכותית שיספקו שירותים אלו באופן ישיר ובין אם על ידי כלי בינה מלאכותית שיקלו על עבודת מורים ואנשי חינוך ויאפשרו להם לספק שירותים אלו, בין השאר באמצעות שחרור המורים מביצוע פעולות בירוקרטיות (Holstein and Doroudi, 2021; Selwyn, 2022). במקרי המבחן המצורפים יש דיון הבוחן שני סוגי מימושים של הבטחות אלו (הראשונה והשנייה) ומדגים את המורכבות שלהם.

חשוב להדגיש, כי מה שמשותף לכלל ההנחות הללו היא שההבטחה של בינה מלאכותית בחינוך לקידום הוגנות במערכות חינוך לא נובעת מטיפול נקודתי בשאלות של אי-שוויון, חלוקת משאבים והתמודדות עם הטיות, אפליות והדרות המתקיימות במערכת, אלא בעיקר מבוססות על ההבטחה הכללית יותר של בינה מלאכותית בחינוך לשיפור וייעול בתהליכי הלמידה, בהתאמת הלמידה וההתאמה לתלמידים - פרסונליזציה, הבנה מעמיקה של תהליכי הלמידה האישיים ובכלל וייעול קבלת ההחלטות ברמת המורה וברמת המערכת. כלומר, ההבטחה הזו מתבססת על כך ששיפור וייעול הלמידה וקבלת ההחלטות החינוכיות תאפשר ליותר תלמידים להגיע להישגים, להתגבר על קשיים של משאבים וגישה, ולקבל החלטות חינוכיות מדויקות, נכונות ומתאימות יותר.

העובדה שלמעשה עיקר ההבטחה של בינה מלאכותית לקידום הוגנות בחינוך מתבססת על היבטים של שיפור וייעול השירותים החינוכיים והרחבת הנגישות אליהם מעלה כמה שאלות: ראשית, ההיסטוריה של טכנולוגיה חינוכית, כפי שנראה בפרק הבא, מלמדת אותנו שהבטחה של אוטומציה של פעולות חינוכיות שמצד אחד מייעלת את הפעולה החינוכית ומצד שני משחררת את המורות מפעולות מסדר נמוך היא הבטחה שלא מומשה בטכנולוגיות קודמות ובעיקר לא קידמה הוגנות ושוויון. אם כן, בהקשר זה, השאלה היא באיזה אופן מתייחדת הבינה מלאכותית מטכנולוגיות אוטומטיות קודמות? ומדוע ניתן להניח שהיא תשפיע באופנים אלו על מערכת החינוך? שנית, יעילות אינה בהכרח משקפת פעולה חינוכית מיטבית. בהקשר של בינה מלאכותית מובעים חששות רבים מכך שהיעילות שמערכות אלו מייצרות בהקשר החינוכי יכולה להוביל לפגיעה בסוכנות הלומדת של תלמידים ותלמידות ובפיתוחם וטיפולם כלומדים עצמאיים (ראו בהמשך וכן לדוגמה: Liu, 2020). אבל מעבר לכך השאלה היותר עמוקה נוגעת לסוג

הראייה החינוכית שכלי הבינה המלאכותית מייצרים וסוגי הידע והאמת שהם מייצרים ומקטלגים. כפי שמסבירה קייט קורפורד "האיפשרים של כלי [הבינה המלאכותית] הופכים להיות האופק של האמת" (Crawford, 2019, P.118). "The affordances of the tools become the horizon of truth". האם בתהליך הייעול הנובע משימוש בבינה מלאכותית תהליכים של למידה, הוראה והערכה יעברו רידוד וצמצום? והאם תהליך זה עלול גם לפגוע בתלמידים שונים, במטרות אחרות של הוראה ובערכים אחרים, כמו שוויון וצדק חברתי?

מקרה מבחן א' - האם מערכת למידה אדפטיבית יכולה לצמצם פערים?

בכמה דוחות ומסמכים (ברק-מדינה, 2019; משרד החינוך, 2022; Pedro et al., 2019) מוצגת סין כמדינה שבה יש את ההתקדמות הרבה ביותר בכל מה שקשור להטמעה ושימוש במערכות בינה מלאכותית בבתי הספר ובמערכת החינוך. בדוחות אלו מוזכרים כמה מערכים טכנולוגיים ספציפיים - כמו מערכות לזיהוי פנים לצורך ניהול כיתה, מערכות אוטומטיות לציון והערכה, מערכות אדפטיביות ללימוד מתמטיקה ואנגלית ועוד - ונאמר כי פרויקטים טכנולוגיים אלו הם בעלי פוטנציאל לא רק לשיפור מערכת החינוך אלא גם לקידום הוגנות והגינות בה. המקרה הסיני מעניין במיוחד לאור המאמץ המדינתי הכולל לקדם טכנולוגיות בינה מלאכותית בחינוך, תוך שילוב בין המגזר הפרטי לציבורי ופתיחת דלת לחברות טכנולוגיה לפתח ולהטמיע מערכות אלו בשדה החינוך (Pedro et al., 2019; Liu, 2020; Knox, 2020).

במיוחד בולטת דוגמה של Liulishou שהיא פלטפורמה אדפטיבית ללימוד אנגלית המאפשרת הוראה-למידה מותאמת אישית (Pedro et al., 2019). חוקרים מסוימים (Aby et al., 2019) נוקטים בעמדה הרואה בפרויקט לא רק פרויקט חינוכי חשוב המעלה את הישגי ויכולות התלמידים במקומות שונים בסין אלא גם כלי לשוויון וצמצום פערים בעקבות היכולת של הפלטפורמה לספק הוראה איכותית באנגלית לתלמידים באזורים פריפריאליים. אבל על אף הטון החיובי של המאמר גם הם מודים שההצלחה של הפרויקט פחות מבוססת על נתונים ברורים ויותר על אופטימיות ותקווה. למולם המאמר של ליו (Liu, 2020) מציג עמדה אחרת וטוען שמערכות אלו מהוות אולי פתרון יעיל ללמידה המונית גם בקרב קהילות ואוכלוסיות עם נגישות נמוכה לחינוך איכותי, אבל למעשה אוכלוסיות חזקות נהנות מחינוך איכותי ואישי (המבוסס בין השאר על טכנולוגיות דיגיטליות אחרות) המחזק את כישורי ההכרה העצמית והכישורים המטה-קוגניטיביים כמו יצירתיות וחשיבה ביקורתית בעוד אוכלוסיות מוחלשות המתבססות על למידה בפלטפורמה האדפטיבית, לומדות באופן שאינו מפתח כישורים אלו ובכך פוגעות בסוכנות העצמית שלהם (Liu, 2020). כך על אף שיתכן שהפלטפורמה הנגישה את הלמידה לאוכלוסיות מוחלשות ופריפריאליות, הרי שבפועל מדובר בלמידה נחותה ביחס למשאבים, טכנולוגיות ושיטות למידה בקרב אוכלוסיות חזקות.

מעבר לכך עמדה זו ממקמת את השימוש בפלטפורמה הזו גם בתוך הקשר רחב יותר של הטמעת בינה מלאכותית בסין בהקשרים של פרטיות ואתיקה (ראו גם Knox, 2020). כך בסין פותחו והוטמעו כלי בינה מלאכותית המבוססים על חיישנים כדוגמת זיהוי תמונה וקול העוקבים אחרי התלמידים ותגובותיהם בזמן אמת. מערכות אלו פעלו בבתי ספר ובכיתות בסין ושימשו לניהול כיתה באמצעות מעקב אחרי הבעות הפנים של תלמידים (עניין, עייפות, ריכוז וכדו'). במהלך שנת 2019 הממשלה הסינית נסוגה מהתמיכה בסוג חודרני שכזה של טכנולוגיות, והודיעה על אמצעים להגבלה ולרגולציה של טכנולוגיות מעקב וזיהוי פנים בבתי הספר בעיקר בעקבות ביקורות של הורים ואנשי חינוך (Zhao, 2021).

מקרה מבחן ב' - מערכות זיהוי פנים וההבטחה להפחתת עבודה שגרתית ממורים

ניל סלווין, מחוקרי הטכנולוגיה החינוכית החשובים בעולם, בוחן מקרה מבחן של פיתוח והטמעה בית ספרית של מערכת זיהוי פנים מבוססת בינה מלאכותית בבתי ספר באוסטרליה. מערכות זיהוי פנים הן

מערכות שפועלות על ידי חילוץ חישובי של תכונות פנים מתמונה דיגיטלית או מסגרת וידאו, והשוואת 'הדפסי פנים' ייחודיים אלה עם תכונות שנשמרו בעבר במסד נתונים כדי לנתח תמונות פנים (Zhao, 2021). מערכות אלו שנויות במחלוקת ונטען שהן מבוססות ומקדמות אפליה על רקע גזע ומראה וסוגים אחרים של פגיעה חברתית, מכיוון שהן נוטות להטות שונות, ומחייבות מעקב דיגיטלי מתמיד וחדירה בלתי פוסקת לפרטיות (Zhao, 2021; Oneil et al., 2022). כפי שמתארים קבוצת חוקרים מאוסטרליה בהובלת סלווין (Selwyn et al., 2022; O'neil et al., 2022) התעשייה מקדמת כלי זיהוי פנים כפתרון לבעיות שונות במערכת החינוך על בסיס ההנחה של אוטומציה ושל מעקב זהיר ('careful' surveillance) תוך הבטחה לטיפול בבעיות כמו אבטחת הקמפוס, רישום אוטומטי של נוכחות וניהול כיתה בהתבסס על זיהוי ותגובות רגשות של סטודנטים (Andrejevic and Selwyn, 2020, Zhao, 2021).

במאמריו בוחן סלווין (Selwyn, 2022; Selwyn et al., 2022) את ההבטחה של מערכות בינה מלאכותית של קבלת החלטות אוטומטיות להפחית עבודה יומיומית וביורוקרטיה ממורים. הוא עושה זאת באמצעות בחינה של אופני שימוש והטמעה של מערכת לזיהוי פנים לבדיקת נוכחות בבית ספר באוסטרליה. סלווין מציין את האופן שבו המערכת מחייבת התקנה והפעלה של תשתיות מסוימות בבית הספר ואת הדרך שבה היא פועלת כנגד תפיסה מקומית והקשרית של העשייה החינוכית כמו גם את הממדים העדינים, כדבריו, של הפחתת המקצועיות (de-professionalizing) של המורים (Selwyn, 2022). בין השאר, הטענה שעולה ממחקריו היא שבעוד המערכת הטכנולוגית ביקשה להתגבר על החיכוך האנושי הרי בפועל "לחיכוכים" הללו יש תפקיד משמעותי בעבודת היחסים הכיתתית ובתפקוד היומיומי.

במבט רחב יותר על טכנולוגיות לזיהוי פנים טוענים אנדריוויץ' וסלווין שאין לראות בטכנולוגיות אלו רק תוספות שגרתיות למערכות בית ספר עם תרבויות נרחבות של ניטור ומעקב אלא מערכות שיש להן השפעה על הפדגוגיה ומערכות היחסים בבתי הספר והסכנה שטכנולוגיית זיהוי פנים תשנה את אופיים של בתי ספר בכיוונים מפלגים, סמכותיים ודכאניים (Andrejevic and Selwyn, 2020). לאור זאת, טוענת קבוצת חוקרים אלו, כי גם אם ימצאו יתרונות פדגוגיים ומנהליים לשימוש במערכות זיהוי פנים בבתי הספר, דבר שלא הוכח עדיין כלל וכלל, הרי שצריך לשקול את עצם האפשרות להשתמש בטכנולוגיות אלו, לאור המשמעויות שהן נושאות בתוכן של מעקב ושליטה קיצוניים (ראה גם מסקנה דומה אצל Zhao, 2021).

שני מקרי המבחן המובאים לעיל מדגימים את הראייה המורכבת הנדרשת כאשר ניגשים לבחון את התרומה הממשית של מערכות בינה מלאכותית להיבטים של הוגנות, הגינות, הכלה ושוויון במערכת החינוך ואת האופן שבו השאלות הללו כרוכות בשאלות רחבות של מדיניות חינוך, של שאלות אתיות ורגולטיביות. לצד זאת, ההבטחה שתוארה בתחילת פרק זה מתבססת על היתרונות הממשיים של מערכות בינה מלאכותית כמערכות טכנולוגיות רבות עוצמה, המבוססות על נתוני עתק (Big data) והמאפשרות לספק שירותים חינוכיים מותאמים, מהירים, ניתנים למעקב וזולים יחסית. העדר ניסיון ממשי לאורך זמן אינו מאפשר להעריך עדיין באופן ברור את התרומה של מערכות אלו לסוגיות שבמרכז סקירה זו. לאור זאת, הפרק הבא מבקש להתבונן על טכנולוגיות דיגיטליות בחינוך (שלא מבוססות על בינה מלאכותית) ועל הידע שנצבר במחקר בדבר התרומה שלהם להיבטים של הוגנות ושוויון. מבט היסטורי זה מאפשר להתבונן על הפער שבין ההבטחה הגלומה בטכנולוגיות חינוכיות ובין הדרכים שבהם הבטחה זו מתממשת במציאות הבית ספרית והחינוכית.

פרק ד' | טכנולוגיה חינוכית כמקדמת הוגנות בחינוך - ההבטחה שלא התממשה

לאור ההבטחה הסוחפת של בינה מלאכותית בחינוך לקדם הוגנות במערכת החינוך מחד והעובדה שבמידה רבה היישום של בינה מלאכותית בחינוך הוא בחיתוליו מאידך, הרי שאחת הדרכים האפשריות על מנת להעריך את התרומה (החיובית או השלילית) של בינה מלאכותית לסוגיות של הוגנות ושוויון היא באמצעות מבט על טכנולוגיות חינוכיות קודמות ולבחון את הפער בין ההבטחה שהן נשאו איתן ובין מה שהתרחש בפועל. במובן הזה, פרק זה מבקש לתת הקשר רחב, היסטורי וחינוכי, בנוגע לאופן שבו טכנולוגיות חינוכיות שונות ומתחדשות מוטמעות ומשפיעות על היבטים של הוגנות ושוויון במערכות חינוך.

חוקרים רבים (לדוגמה: Cuban, 2003; Chan, 2019; Reich, 2020) תיעדו את מחזורי ההתלהבות והתקווה (Chan, 2019) סביב טכנולוגיות חינוכיות עוד מהמצאת הרדיו והטלוויזיה ועד התפתחות טכנולוגיות דיגיטליות ואינטרנטיות. המאפיינים של מחזורים אלו הם התלהבות, תקווה וציפייה גדולה שטכנולוגיות שונות אלו יביאו לשיפור ולשינוי (בלתי נמנע והכרחי לכאורה) בדרכי הלמידה ובמוסדות החינוך. כאשר הציפייה החיובית מסתיימת בקול ענות חלושה - במימוש חלקי בבתי הספר, בהשתלבות של הטכנולוגיות במערכים הקיימים בבתי הספר ועוד. אין זה אומר שטכנולוגיות חינוכיות ובמיוחד טכנולוגיות דיגיטליות בעשורים האחרונים לא נכנסו לשימושים נרחבים במערכות חינוך או שאינן משפיעות בדרכים שונות על דרכי הלמידה ועל מוסדות החינוך, אלא שלרוב יש פער נרחב בין הציפייה והתקווה ובין המימוש בפועל. מבחינה זו, מבט היסטורי על טכנולוגיות חינוכיות, על ההבטחות שהן נשאו איתן והאופנים שבהן הן התממשו, הופעלו והשפיעו על מערכות חינוך הוא הכרחי על מנת להבין לאשורו את הפוטנציאל של בינה מלאכותית בחינוך כמו גם את המגבלות והדינמיקות השונות בניסיונות העכשוויים והעתידיים בהטמעה ובשימוש בבינה מלאכותית בבתי הספר. לבחינה זו חשיבות מיוחדת לאור העובדה שתחום הבינה המלאכותית נוטה במיוחד ל"הייפ" ולהבטחות רחבות ומרחיקות לכת (Marcus, 2022). לענייננו, בפרק זה נשאל האם טכנולוגיות דיגיטליות אחרות שהופעלו והתממשו במערכות חינוך בית ספריות ברחבי העולם בשני העשורים האחרונים אכן תרמו לקידום ההוגנות ההכלה והשוויון?

כנקודת מוצא לדיון זה הרי שהמחקר לגבי טכנולוגיה חינוכית מצביע על הנטייה של טכנולוגיה חינוכית להשתלב במערכים החינוכיים ובמבנים הקיימים (Reich, 2020). כמו שמסביר לארי קובן, היסטוריון של רפורמות חינוכיות בכלל ושל טכנולוגיה חינוכית בפרט, ללא שינוי בשיטות ההוראה והלמידה טכנולוגיות חינוכיות ייטו לחזק את שיטות ההוראה והלמידה הקיימות ולא לשנותן גם אם יש בהן פוטנציאל לשינוי (לדוגמה: Cuban, 2003). אם להשתמש במטבע הלשון שטבע חתן פרס ישראל, חוקר החינוך גבריאל סולומון (2000), הרי שהטכנולוגיה החינוכית היא הזנב של הכלב שהוא מערכת החינוך או הפדגוגיה, ובדרך כלל הכלב מקשקש בזנב ולא להיפך. כך כפי שמראים חוקרי טכנולוגיה חינוכית רבים האקלים הלימודי הכללי בבתי הספר מעצב את הדרך שבה משתמשים במדיה דיגיטלית, והטכנולוגיה משרתת, כאשר היא מופעלת באופן יעיל, להשיג את המטרות הקודמות של בית הספר ולא לשנות את המטרות עצמן (Warschauer and Matuchniak, 2010). גם שאלת ההשפעה של טכנולוגיה חינוכית על הישגים לימודיים מתווכת דרך שאלות רבות של הקשרים ותנאים מקומיים. בסקירת מחקרים רבים (Warschauer and Matuchniak, 2010) עולה, כי הקשר בין שימוש בטכנולוגיה בבית הספר ובין תוצאות והישגים אינו מובהק ותלוי, בין השאר, בתנאים ובהקשרים שונים כמו הרמה הסוציו-אקונומית של התלמידים ובתי הספר. כמו כן עולה, כי לעיתים לא זוהתה כלל תרומה חיובית לטכנולוגיה על התוצאות, כאשר הסיבות לכך יכולות להיות שונות ומגוונות ובעיקר נטועות בהקשר המקומי.

המילטון והטי (Hamilton and Hattie, 2021) ביצעו את אחת הסקירות המעמיקות והרחבות בתחום וטוענים, כי לצד העובדה שרוב המחקרים עד כה לא מראים השפעה משמעותית של טכנולוגיה חינוכית על תוצאות והישגים, הרי שאחת הבעיות היא העובדה שהמושג טכנולוגיה חינוכית הוא מושג רחב מאד ומתבטא במגוון כלים ותוכניות - חלקן שקופות ולא מורגשות, חלקן משולבות זו עם זו. לאור זאת, הסקירה המטא-אנליטית והניתוח שהם מבצעים מבקשת להצביע על תחומים או קבוצות של טכנולוגיות ושימושים טכנולוגיים בהם יש השפעה משמעותית יותר לעומת תחומים בהם יש השפעה פחותה. לדוגמה, הם מראים כי למערכות הוראה חכמות מותאמות אישית ומערכות שמבוססות על וידאו אינטראקטיבי יש השפעה חיובית על הישגים בעיקר במתמטיקה ולטכנולוגיות אחרות יש השפעה חיובית יותר אצל תלמידים עם צרכי למידה מיוחדים לעומת תלמידים ללא צרכים מיוחדים. יחד עם זאת, ממצאים אלו מסויגים לאור העובדה שהמטא-אנליזות הללו מכילות בתוכם מחקרים בהקשרים, בגדלים, ובתנאים שונים (ראו גם Tamim et al., 2021).

מהניתוח של המילטון והטי אפשר לראות כיצד טכנולוגיה, מוצלחת ככל שתהיה, יעילה יותר או פחות בתוך הקשרים מקומיים ובתנאים שונים כמו תחומי דעת, סוגים שונים של תלמידים, צורות שונות של למידה ובחינת הישגים ועוד. לאור זאת, עולה שיש קושי להעריך את התרומה וההשפעה האפשרית של טכנולוגיה חינוכית בכללה ואפילו טכנולוגיות ספציפיות מכיוון שטכנולוגיות חינוכיות תמיד פועלות בתוך הקשר (Contextual) ומתייחסות לגורמים ותנאים שונים בשדה (Relational) (לדוגמה: Castañeda and Williamson, 2021). לפיכך, טוענים פוקט ורפלאו (Puckett and Rafalow, 2020) כי הדיון בנוגע לטכנולוגיה חינוכית צריך לנוע משאלות של השפעה (Impact) לשאלות של משא ומתן (Negotiation). בעוד שאלות של השפעה מבקשות להעריך ולכמת כיצד שימוש בטכנולוגיה מסוימת משפרת תוצאות והישגים חינוכיים שונים, הרי שאלות משא ומתן מבקשות לכוון את תשומת הלב לאופנים שבהם טכנולוגיה תפעל בהקשרים שונים, כיצד היא תעמוד ביחס למטרות חינוכיות שונות, יישומים שונים של מדיניות, למבנים מוסדיים וארגוניים ועוד. שאלות של משא ומתן מתמקדות ביתרונות והחסרונות של טכנולוגיה ביחס להקשרים ספציפיים של שימוש חינוכי.

בהמשך לתפיסה זו מסבירים חוקרים (Warschauer et al., 2004) שטכנולוגיה בבית ספר וההשפעה שלה הינה תלויה הקשר במיוחד בסוגיות ספציפיות של חלוקת משאבים ואי-שוויון. לדוגמה, מכיוון ששימוש בטכנולוגיה בבתי הספר דורש ותלוי בסוגים שונים של משאבים וגישה פנים בית ספריים, כמו מכשירים ותשתיות חומריות ומקצועיות, וחוף בית ספריים, כמו תמיכה ביתית, הרי שהוא יכול להוות גורם שמגדיל אי-שוויון וחוסר הוגנות מכיוון שהוא יוביל לשימוש חלקי או נחות בהקשרים מוחלשים לעומת שימוש טוב יותר בהקשרים חברתיים חזקים (Warschauer et al., 2004; Puckett and Rafalow, 2020).

רפלאו ופוקט (Rafalow and Puckett, 2021) מבקשים להרחיב את נקודת המבט על הדיון בנוגע לקשר שבין טכנולוגיות בבית ספר ואי-שוויון מתוך כך שהם מצביעים על ההקשרים שבהם בתי הספר עצמם ומערכות חינוך בכללותן הם מוסדות שמבנים "אי-שוויון קטגוריאלי". כלומר הם ארגונים שבנויים על יצירת קטגוריות וקלסיפיקציות מבדילות המשמרות הבחנות בין יחידים, קבוצות ואוכלוסיות. בהקשר זה הם מבקשים להפנות את תשומת הלב לאופן שבו הקטגוריזציה של מוסדות החינוך מתממשת, מתווכת ואף מועצמת על ידי טכנולוגיות שונות. בהקשר זה, חוקרים אלו קוראים להתבונן גם על בחירות בתכנון ובעיצוב הכלים הטכנולוגיים - כמו על אילו נתונים הם מתבססים, בחירות שונות בתכנון הכלי, הפיצ'רים והאפשרויות השונות המוצעות ומופעלות בכלי הטכנולוגי - והאופן שבו הן מתממשות בהקשרים בית ספריים ספציפיים, מתייחסות לחלוקות הדמוגרפיות הבית ספריות ולהנחות של בתי ספר לגבי החלוקות הללו. לדוגמה, המבוססת על מחקרים רבים, עלול להתרחש מצב כי כלפי אוכלוסיות מוחלשות יופעלו יותר תוכנות השמות דגש על מעקב ופיקוח וכלפי אוכלוסיות "חזקות" יופעלו יותר תוכנות שמעודדות למידה עצמאית ובחירה. ואפילו אם יופעלו כלים טכנולוגיים זהים כלפי אוכלוסיות שונות, הרי הדגשים בשימוש יהיו שונים (לדוגמה, בצורת פיצ'רים שבית הספר יבחר להפעיל או דרך שימת לב לפיצ'רים שונים בהקשרים שונים) (ראה גם Gilliard and Selwyn, 2022).

הדוח של רייך ואיטו (Reich and Ito, 2017) שכותרתו: From good intentions to real outcomes: Equity by design in learning technologies בוחן ומסכם את הקשר שבין טכנולוגיות חינוכיות ושאלות של אי-שוויון. הדוח מראה שעל אף ההבטחות, העדויות מצטברות לכך שטכנולוגיות חדשות אלו נוטות להיות משומשות ונגישות באופן לא שוויוני, ובכך הן נוטות להגביר ולהגדיל את אי-השוויון. גם קיזילצ'ק' ולי (Kizilcec and Lee, 2020) קובעים כי רוב המחקרים שבוצעו על ההשפעה של טכנולוגיות חינוכיות על תוצאות תלמידים מצאו כי בדרך כלל יש הרחבה בפער שבין קבוצות מוחלשות וחזקות, גם אם יש מחקרים יוצאי דופן המדווחים על שמירת הפער ואפילו צמצום מסוים שלו.

בהמשך לתפיסת המשא ומתן ולאופן שבו טכנולוגיה חינוכית צריכה להיות מובנת בתוך הקשר ויחסים, הרי שרייך ואיטו מציינים שלושה ממצאים מרכזיים המסבירים את הנטייה של טכנולוגיות דיגיטליות חדשות לשמר ואף להגדיל את אי השוויון: ראשית, הם מציינים שבתי ספר המשרתים אוכלוסיות מחוזקות נוטים להשתמש באותן טכנולוגיות בדרכים פרוגרסיביות ויעילות יותר מאלו שמשרתים אוכלוסיות מוחלשות. שנית, הם מצביעים על כך שבדרך כלל (ובניגוד להבטחה שלרוב מלווה אותם) חומרים וכלים פתוחים ונגישים לא מובילים לדמוקרטיזציה או לקידום הוגנות, אלא מביאים יתרון לא פרופורציונאלי לילדים להורים בעלי השכלה גבוהה. ולבסוף הם מראים שמה שקובע תוצאות והישגים אינה הטכנולוגיה עצמה - גישה ושימוש - אלא מרכיבים מערכתיים וחברתיים שבתוכם הטכנולוגיה מוטמעת. במילים אחרות, הטכנולוגיה אינה מצליחה להתגבר על ההטיות המבניות, המערכתיות והחברתיות המובילות לאי-שוויון. במובן זה הנסיונות של מערכות חינוך להשתמש בטכנולוגיה על מנת לקדם הוגנות ושוויון נכשלו בדיוק מכיוון שבדרך כלל כוחות חברתיים ושאלות מבניות לא נלקחו בחשבון על ידי קובעי המדיניות, יצרני הטכנולוגיה ומיישמייה. זאת ועוד, הם טוענים כי העדר התייחסות לכוחות החברתיים, המוסדיים והתרבותיים הוביל ועלול להוביל גם בעתיד לתוצאות לא מתוכננות ופוגעניות בגלל הטיות מוסדיות או הטיות בלתי מודעות שלא נלקחות בחשבון בשלבי הפיתוח וההטמעה של הטכנולוגיה או לחילופין בגלל פער שבין המפתחים וקובעי המדיניות ובין קהל היעד וסביבת היישום.

פרק ה' | הוגנות צרה בבינה מלאכותית בחינוך - חששות והתמודדויות

בפרק זה נעמוד על החששות הספציפיים שעולים לגבי שימוש בבינה מלאכותית בחינוך. הדגש בפרק זה יהיה על חששות הנוגעים למה שכינינו הוגנות צרה, כלומר להיבטים הנוגעים למערכת בינה המלאכותית עצמה ולדרכי פעולתה ולא להיבטים רחבים יותר - אתיים, תרבותיים וחברתיים. כמו כן פרק זה יפרוש מספר כלי עבודה קיימים להתמודדות עם חששות אלו.

כאמור, מערכות בינה מלאכותית תלויות במידע המוזן אליהן ובנתונים שמשמשים אותן לאימון לבניית מודלים ולקבלת החלטות אוטומטית. כך חששות שקשורים לאיסוף נתונים שואלים: מהם הנתונים שהוזנו למערכת? מה מידת הייצוגיות שלהם? וכיצד אפשר להכליל מהם (Kizilcec and Lee, 2020)? ברמה הראשונית, יכולות להופיע הטיות סטטיסטיות כמו הטיית דגימה או שגיאות מדידה, אם כי אלו נדירות יותר במערכות בינה מלאכותית (Mitchell et al., 2021). אך מעבר לכך, בין אם מדובר בניחות של נתוני עתק ובין אם נתוני אימון, הרי ההסתמכות על נתונים קיימים העלולים להכיל בתוכם הטיות חברתיות, מעלים את הסיכון של הנצחת ואף הגברת אי-השוויון חברתי שמשתקף בנתונים אלו. כך, אם הנתונים המוזנים למערכת מכילים בתוכם הטיות או משקפים עיוותים ואי-שוויון בין קבוצות חברתיות הקיימות במציאות, המערכת תלמד הטיות וביטויי אי-שוויון אלו, תעתיק ותהדהד (ויתכן אף תעצים) אותם (נהון, 2019; Baker et al., 2019; Holstein and Doroudi, 2021). דעות קדומות והטיות חברתיות יכולות לזחול לאלגוריתמים בדרכים שונות. לדוגמה, למידת מכונה עשויה ללמוד מהנתונים שעל פיהם היא מאומנת על מצב חברתי כלשהו שמכיל הדרה ואפליה מסיבות היסטוריות וחברתיות - כמו גברים שמועסקים במשרות מסוימות יותר מנשים או שמרוויחים יותר על אותן משרות - ועל פיהם לתת עצות או הכוונות המשמרות הדרות ואפלויות אלו. מקרים דומים של שימור הטיות יכולים להיות כאשר הבינה המלאכותית מזהה בעצמה קורלציות בין תכונות - כמו מוצא אתני ומשכורת - ומשמרת אותם באמצעות הפיכתן לקריטריונים בקבלת ההחלטות. לעיתים הקורלציות הללו יכולות להיות נסתרות יותר ועוברות דרך משתני ביניים שונים (Baker et al., 2019).

הטיות אלו עלולות להיות חמורות יותר כאשר בני אדם המושפעים מגורמים חברתיים ותרבותיים שונים מעורבים בתיוג הנתונים, במיוחד כאשר זהותם של המתייגים אינה מייצגת שונות חברתית (Baker and Hawn, 2021). תת בעיה נוספת לבעיה זו היא הטיית חזרה של מאמצים מוקדמים, כלומר האופן שבו החלטות מוטות של מורים ובעלי תפקידים אחרים הופכות להיות הבסיס לנתונים להמשך וכך אותן ההטיות משוכפלות. במובן הזה הסיכון של הטיות מוקדמות (בנתוני הלמידה וההחלטה של הבינה המלאכותית) יכול להיות בעיה מצטברת כאשר ההטיות הראשוניות מייצרות תוצאות והחלטות שבתורן מייצרות ומחזקות ומגבירות את ההטיות דרך מעגלי היזון חוזר (Feedback loop) (Holstein and Doroudi, 2021).

בנוסף, ישנם סוגי הטיות הקשורים לגורמים אינהרנטיים לאלגוריתמים שמניעים את הבינה המלאכותית ואת קבלת ההחלטות האוטומטית. ביניהן: הטיות נגד קבוצות מיעוט הנובעות מהדומיננטיות של הנתונים של קבוצת הרוב בתהליכי האופטימיזציה של מודלים, וכן הטיות שיצמחו מתוך העדר-ספציפיקציה במודל, הנובעים מהפשטה של מורכבות או בהתעלמות מפרמטרים שונים במודל. לדוגמה, מודלים של למידה מהווים הפשטה של מורכבות תהליך הלמידה וכך עשויים להתעלם או לא להתחשב מספיק בפרמטרים שונים. (Holstein and Doroudi, 2021).

בנוסף להטיות העלולות להיגרם מהסתמכות על נתונים או מגורמים הנובעים מאופי המערכות האלגוריתמיות יש גם חשש להטיות הנובעות מהפעולה של מערכות הבינה מלאכותית בעולם. אלו הטיות שעלולות לנבוע מהתכנון הסוציו-טכנולוגי של המערכות ומיחסי הגומלין שבין קבלת ההחלטות האוטומטית וקבלת ההחלטות האנושית (Holstein and Doroudi, 2021). כלומר גם כאשר ההמלצות עצמן של מערכות הבינה המלאכותית לא מייצרות הטיות ופערים הרי שהשימוש שלהן בפועל יכול לתרום לתוצאות של אי-שוויון או הטיות ואפילו שונות.

אחת הנקודות המרכזיות בהקשר זה היא העובדה שבינה מלאכותית היא מערכת מורכבת שקשה להבנה למשתמש האנושי וקבלת ההחלטות האוטומטית שלה בדרך כלל לא מוסברת ואינה שקופה. מערכות הבינה המלאכותית הן למעשה "קופסה שחורה" (Black box) שפעולותיהן ותוצאותיהן בדרך כלל אינן מובנות ואינן ניתנות לבקרה (Baker et al., 2019). עובדה זו מקשה על המשתמשים להבין את המלצות המערכת כמו גם את השיקולים מאחוריהן ולפתח עודף אמון במערכת או לחלופין חוסר אמון במערכת. כלומר מצד אחד עלולה להתפתח נטייה מוטעית להתייחס לתוצר של מערכות בינה מלאכותית, המנתחות נתוני-עתק ובמהירות כאמת מדעית, ולכן לא להפעיל מנגנוני בקרה הולמים או לא לזהות הטיות ועיוותים (נהון, 2019) ומצד שני שימוש במערכות בינה מלאכותית באופן חשדני עלול להוביל לפגיעות ולהטיות גם כן. הספרות (לדוגמה Hosnagar, 2019) דנה בהרחבה בשאלת האמון ובאופן שבו הטיות אנושיות ותפיסות של בני אדם את מערכות הבינה המלאכותית עלולות להוביל לשימוש מוטה בבינה מלאכותית (ראו בדיון תחת הכותרת שאלת האמון). אחד החששות הקשורים לכך הוא הבלבול האפשרי בין קורלציות והסברים. מערכות בינה מלאכותית לא מתיימרות להציג הסברים לתופעות אלא קורלציות. אלא שלמשתמשים אנושיים יש נטייה לראות בקורלציות הסבר (או לנסח קורלציות כהסברים) וכך להצדיק פגיעות או הטיות (Kizilcec and Lee, 2020). בהקשר זה, אחת הטענות היא שדווקא הקושי של חלק ממקבלי ההחלטות האנושיים לפרש ולהבין את ההתוצאות של מערכת הבינה המלאכותית עלול לגרום לאותם מקבלי החלטות להעדיף ולהדגיש תוצאות מוטות במקום לנסות לתקן אותן, בין השאר מתוך העדפה של הסברים וקורלציות פשוטות (Parapadakis, 2020).

הממשק בין קבלת החלטות אנושיות וקבלת החלטות אוטומטית הוא ממשק מורכב המערב בתוכו שאלות של אמון, של פסיכולוגיה אנושית, של נורמות ותפיסות חברתיות ותרבותיות מצד אחד ומצד שני מדגימות את ההתמודדות האנושית עם מנגנון חשיבה אינטליגנטי המדמה חשיבה אנושית. במובן הזה, ככל שהבינה המלאכותית תתפתח ותיושם במערכת החינוך שאלת הממשק ואופני העבודה המשותפים שבין אנשי חינוך ובין קבלת החלטות אוטומטית מבוססת בינה מלאכותית תהווה אתגר גם לסוגיות של הוגנות ושוויון.

שאלת האמון

בינה מלאכותית ומערכות אלגוריתמיות משולבות עם החלטות וביצועים אנושיים. כאמור אחד החששות להטיות ולאפוליות נובע מהאופן שבו מבצעים אנושיים - קרי אנשי חינוך ומורים - יעבדו ויפעלו עם מערכות של בינה מלאכותית. בהקשר הזה אחד מהנושאים המרכזיים הוא שאלת האמון של בני האדם במערכות בינה מלאכותית. שאלת האמון היא שאלה משמעותית, מכיוון שאמון גדול מדי של שחקנים אנושיים מונע אפשרות לבקרה של המערכת והופך את השחקן האנושי ממחליט לטכנאי, ולעומת זאת חוסר אמון מעלה את האפשרות להטיות אנושיות נוספות, ולשימוש לא מושכל בבינה המלאכותית (Hosnagar, 2019).

מחקר שנערך במכון וייצמן (Nazaretsky et al, 2021) בחן את שאלת האמון במערכות בינה מלאכותית מצד מורים למדעים בישראל שהשתתפו בהכשרה מקצועית לשימוש בסביבת הלמידה PETEL המספקת הוראה מותאמת אישית. חלק מהממצאים מלמדים כי מורים אינם ששים לקבל המלצות מבוססות בינה

מלאכותית כאשר המלצות אלו סותרות את הידע הקודם שלהם על תלמידיהם. כלומר הם נוטים שלא להאמין למערכת כאשר היא סותרת את האינטואיציה והידע והניסיון הקודמים שלהם. לעומת זאת, באופן פרדוקסלי, הם מצפים שבינה מלאכותית תהיה נכונה לחלוטין גם במצבים שאולי אמת מוחלטת לא קיימת. כלומר המורים שהשתתפו במחקר גילו אמון גבוה מאד במערכת הבינה המלאכותית, אבל במצבים מסוימים אמון זה הושהה או בוטל.

ממצאי מחקר זה מדגימים את המורכבות של שאלת האמון ושל היחסים שבין משתמשים שהם אנשי חינוך ובין המערכות הטכנולוגיות. הם גם מסמנים את החשיבות של ההתייחסות לסוכנות האנושית באימוץ של בינה מלאכותית בחינוך ולאופנים שבהם משתמשות - כלומר, מורות, מחנכות, מנהלות ואף תלמידות - תופסות את פעולת הבינה המלאכותית, את התוקף החינוכי שלה וכיצד הן פועלות עם ולצד הטכנולוגיה.

התמודדות עם חששות ומערכי בקרה

חוקרים בתחום של בינה מלאכותית בכלל ובינה מלאכותית בחינוך מנסים להציע מגוון סוגי התמודדות לבעיות ההוגנות כפי שפורטו. אפשר לחלק את דרכי ההתמודדות הללו לשתי קבוצות: 1. שיפור, תקנון ותיקון של האלגוריתמים והמערכים הסטטיסטיים של הבינה המלאכותית עצמם 2. תיקון ושיפור של נקודות הממשק שבין מערכות הבינה המלאכותית והמשתמשים האנושיים.

צעדי הוגנות (Fairness measures) הינן שיטות וגישות סטטיסטיות ואלגוריתמיות שונות המאפשרות לתקן ולתקן הטיית אלגוריתמיות (Baker and Hawn, 2021; Mitchell et al., 2021). תחום זה של טכניקות אלגוריתמיות לשיפור הוגנות הוא תחום מחקר פעיל שצבר תאוצה בשנים האחרונות (Holstein and Doroudi, 2021). הבנה והעמקה בשיטות אלו דורשות ידע אלגוריתמי וסטטיסטי מתקדם ולכן לא נכנס לתיאור מפורט שלהן במסמך זה. קיזילצ'י ולי (Kizilcec and Lee, 2020), לדוגמה מציעים מספר המלצות לטכניקות שניתן לאמץ בכל שלב של תכנון והפעלה של מערכות בינה מלאכותית לצורך שיפור ההגנות האלגוריתמית בתחום החינוך. יחד עם זאת, חוקרים בתחום (Baker and Hawn, 2021) מציינים שהפעלה של צעדי ההוגנות הללו בפועל מעלה סוגים אחרים של בעיות ומכשולים, כמו פגיעה אפשרית בדיוק החיזוי של הכלי, ושכסופו של דבר מדובר באיזון (Trade-off) בין משתנים שונים, כאשר האיזון האופטימלי קשור להקשר ולתחום הספציפי.

שאלת האיזון והפשרה (Trade-off) בין ערכים ותכונות שונות נכונה גם לקבוצה השנייה של צעדי ההוגנות האפשריים המתייחסת להיבטים של הממשק האנושי עם מערכות הבינה המלאכותית (Holstein and Doroudi, 2021). דרכי ההתמודדות השונות שיוצגו בהמשך, מעלות בעיות וסוגיות אחרות ועלולות לפגוע בהיבטים אחרים של הוגנות ושוויון. לדוגמה: חוסר שקיפות של מערכת בינה מלאכותית עלול לפגוע באמון של משתמשים בה, וניסיון למנוע הטיית ספציפיות עלול לפגוע במקרים מסוימים בלומדים אינדיבידואליים.

דיגנום (Dignum, 2021) מציעה שלושה עקרונות לפיתוח בינה מלאכותית בחינוך ושימוש הוגן ואחראי במערכות אלו: אחריותיות (Accountability), אחריות (Responsibility) ושקיפות (Transparency). במידה רבה, עקרונות אלו מקבילים אך לא זהים לעקרונות שפותחו בועדת המשנה לרגולציה ואתיקה (נהון, 2019): של הסברתיות (Explainability), שקיפות ואחריותיות. לצורך פישוט סקירה זו מאחדת בין המושגים (וכן מושגים נוספים מורכבים יותר המופיעים בסקירות והצעות רגולטיביות אחרות - Hosnagar, 2019) על מנת להקל על הקורא/ת:

אחריות/אחריותיות - שאלת ההוגנות (הצרה והרחבה) של בינה מלאכותית בחינוך צריכה להיות חלק מהאחריות של מפתחי, מעצבי ומקבלי ההחלטות לגבי הטמעה ושימוש של מערכות אלו. במובן הזה, כל השחקנים האנושיים הקשורים למערכות אלו הם חלק משרשרת האחריות (Chain of responsibility) של מערכת הבינה המלאכותית (Dignum, 2021). לשיקולי ההוגנות כחלק מהאחריות בשלבים שונים של פיתוח והטמעה של הבינה המלאכותית יש כמה יישומים והשלכות. ראשית, על מנת לצמצם תוצאה מפלה או מוטה של טכנולוגיה, יש ללמוד באופן פעיל על אוכלוסיית היעד, ולזהות מבעוד מועד אוכלוסיות שעלולות להיפגע או להיות מיוצגות באופן לא מספק (נהון, 2019). שנית, חלק מהאחריות היא גם בייצוג של אוכלוסיות שונות בפיתוח עצמו כדרך לפיתוח רגישות ומודעות לסוגיות ולהדרות והטיות אפשריות ולכן חשובה גם שותפות משמעותית של מגוון אוכלוסיות, כולל מקרב אוכלוסיות מוחלשות או מודרות, בעיצוב ופיתוח של הטכנולוגיות עצמן (נהון, 2019; Holstein and Doroudi, 2021). שלישית, המעצבים והמפתחים של טכנולוגיות בינה מלאכותית בחינוך צריכים להיות מודעים גם לדרכים השונות שבהן הטכנולוגיות השונות מוטמעות ומאומצות ולשקול דרכים לפעול באופן שמתייחס ומגיב לשונות הזו (Holstein and Doroudi, 2021). באופן יותר כללי מציעים חוקרים אלו להשקיע בכלים ותהליכים על מנת לתמוך הן בפיתוח של טכנולוגיות בינה מלאכותית הוגנות בחינוך והן בשימוש הוגן בטכנולוגיות אלו. לדוגמה הם מציעים לשלב בפונקציות היעדים של מערכות בינה מלאכותית תוצאות שאינן מתייחסות רק לתוצרי למידה של סטודנטים בודדים אלא גם לתוצאות הקשורות לשוויון, הוגנות והגינות (Holstein and Doroudi, 2021).

שקיפות/הסברות/אחריותיות - כאמור, אחת הבעיות המרכזיות של מערכות בינה מלאכותית היא בעיית ה"קופסא השחורה" כלומר הקושי האנושי - של משתמשים ולעיתים גם של המפתחים עצמם - בלהבין את התוצאות, ההמלצות וההחלטות של מערכות הבינה המלאכותית (Baker et al., 2019). בשנים האחרונות יש עלייה במאמץ להפוך את מערכות הבינה המלאכותית ליותר שקופות וברורות, אבל גם לאחר מאמץ זה שאלת ההסבר וההבנה של החלטות מערכות אלה היא אתגר משמעותי (ובטכנולוגיות מסוימות כגון למידה עמוקה הדבר בלתי אפשרי כלל - Marcus, 2022). עקרון ההסברות מדגיש את הצורך בהסבר של המערכת על הפעולה וההחלטה של הבינה המלאכותית הן ברמת דרכי הפעולה הכלליות שלה והן במתן הסברים על ההכרעות השונות (נהון, 2019; Khosravi et al., 2022). בהקשר זה טוענת דינגום (Dignum, 2021) שחלק מהעיקרון הזה דורש עיצוב מערכות בינה מלאכותית שיכולות לתקשר בצורה ברורה את היכולות שלהן אבל גם את המגבלות - מה הן לא יכולות לעשות, מה יכולות להיות נקודות עיוורון או הטיות אפשריות בתוך המערכת - ובנוסף גם להציג ולהגדיר את המקומות שבהם המערכת מעבירה את השליטה לבני אדם ולאנשי חינוך (Holstein and Doroudi, 2021). במובנים רחבים יותר עקרונות אלו קוראים גם לשקיפות ביחס למסדי הנתונים שבהם משתמשים כבסיס לפעולת הבינה המלאכותית, כמו מסדי הנתונים שמשמשים לנתוני האימון או מסדי הנתונים שנאספים מהתלמידים, כמו גם ביחס לאינטרסים ולמטרות של חברות מסחריות, כמו שקיפות ביחס לשימוש נוסף שנעשה בנתונים, בגופים שיש להם גישה לנתונים ועוד.

פרק ו' | הוגנות רחבה - ביקורות על שימוש במערכות בינה מלאכותית

כאמור, ישנה כתיבה רבה המתבוננת על המחירים והפגיעות הקיימות והאפשריות משימוש במערכות בינה מלאכותית בהקשרים שונים. לדוגמה, אוניל (O'neil, 2016) מכנה מערכות אלגוריתמיות מבוססות נתוני עתק (Big data) "נשק מתמטי להרס" (Weapons of math destruction) וטוענת שהן אינן שקופות, מכילות הטיות לא ידועות רבות ומשמרות ואף מגבירות צורות שונות של אפליה והדרה. קייט קורפורד בספרה Atlas of AI (Crawford, 2019) מפרטת את הדרכים השונות שבהם מערכות בינה מלאכותית משפיעות לרעה על בני אדם ועל חברות שונות, החל משלב הייצור וההפקה של החומרים הפיזיים הנדרשים לשם ייצורה של בינה מלאכותית וכלה בפגיעות באוכלוסיות מוחלשות. הסוציולוגית רוהה בנג'מין (Benjamin, 2019) מדגימה מגוון דרכים שבהן היררכיות חברתיות, בעיקר גזעיות, מוטמעות בשכבות שונות של ההגיון האלגוריתמי וטוענת שלקבלת החלטות אוטומטית יש פוטנציאל להסתיר, להאיץ ולהעמיק אפליה תוך נראות שהיא ניטרלית ואובייקטיבית (ובהקשר החינוכי: Dixon-Román et al., 2019) ושושנה זובוף (Zubuff, 2019) טוענת, כי מערכות בינה מלאכותית הן חלק מרכזי ממה שהיא מכנה "קפיטליזם של מעקב" שבו נתונים אישיים הופכים לסחורה מה שמוביל לפגיעה בזכויות אישיות ובמשטר הדמוקרטי עצמו. התבוננויות אלו על מערכות בינה מלאכותית רלוונטיות גם בקשר לשדה החינוך ומהוות את הרקע להבנת שאלות של הוגנות רחבה.

בחלק זה יפורטו בקצרה ארבעה סוגים של חששות וביקורות ביחס להוגנות רחבה בבינה מלאכותית בחינוך. נושאים אלו מבטאים סוגיות רחבות הנוגעות לשאלות אתיות ועקרונות ביחס לבינה מלאכותית באופן כללי, אך יחד עם זאת יש להן השלכות ישירות על שאלות של הוגנות ושוויון.

1. העדר הוגנות הקשורה לתהליכי ייצור והפקה של בינה מלאכותית

בינה מלאכותית, כמו מערכות מחשוביות אחרות, נדמית כפעולה ללא השלכות חומריות, אבל בפועל להפקה ולייצור של בינה מלאכותית ומערכות אלגוריתמיות מורכבות ישנם מחירים סביבתיים ואנושיים לא מעטים. כך להפקה של חומרים הכרחיים לשימושם של מערכות בינה מלאכותית ישנם ממדים רבים של פגיעה סביבתית וניצול כוח עבודה זול. בנוסף ככל שהמערכת האלגוריתמית מורכבת יותר ופועלת על מסדי נתונים גדולים יותר כך הפקתה דורשת אנרגיה רבה יותר (Cawrford, 2019; Marcus, 2022). בעיה נוספת שקורפורד וחוקרים אחרים מצביעים עליה קשורה לאופן שבו חלק ממסדי הנתונים מופקים ומיוצרים ולא רק להטיות החברתיות שטמונות בהם או לדרך שבה הם מייצגים אוכלוסיות שונות. פעמים רבות מסדי הנתונים הללו נוצרים ללא הסכמה של משתמשים או תוך ניצול של אוכלוסיות "שביות", כמו בעלי תיק פלילי או אוכלוסיות מוחלשות (Cawrford, 2019).

2. העדר הוגנות כלכלית ועליית כוחם של תאגידי ענק

השדה העסקי של הטכנולוגיה הדיגיטלית הוא שדה שאינו שוויוני ולתאגידי ענק כדוגמת אמזון, גוגל ומטה (פייסבוק לשעבר) יש כוח מונופולי בתחומים ספציפיים. התלות בנתונים של הבינה המלאכותית מעצימה יתרון זה. ככל שלחברות יש יותר גישה לנתונים כך הן מסוגלות לשפר את המערכות שלהן בצורה משמעותית. דבר זה יכול להוביל לפגיעה בתחרות וליצירה והעצמה של מונופולים, כאשר שחקנים בעלי כוח שוק ובעלי גישה לנתוני עתק ינצלו את ההסדרים הכלכליים של רשת האינטרנט על מנת לעצב את האופן שבו שחקנים חדשים יכנסו לתחום (נהון, 2019). לכך יש להוסיף את העובדה שרוב כלי הבינה המלאכותית תלויים בתשתיות פרטיות של חברות ענק כמו AWS של אמזון (Williamson and Eynon, 2021).

מבט על הכלכלה הפוליטית של הבינה המלאכותית הוא חשוב במיוחד בשדה החינוך שהוא שירות ציבורי האמור לשרת את האינטרסים החברתיים ולהימנע מהסחרה של נתונים אודות ילדים. בהקשר הזה מפנה קנט סלטמן (Saltman, 2020) את הזרקור לאופנים שבהם הפרויקט של בינה מלאכותית בחינוך מבטא הפניית עורף לחינוך ציבורי ומייצר סוגים שונים של הפרטה. בין השאר, הוא מצייין את הדרכים המורכבות שבהן בינה מלאכותית מכוננת סוגים חדשים, מופרטים ולא ציבוריים של שליטה ובעלות (כדוגמת שליטה על נתונים או על כלים ספציפיים).

3. העדר הוגנות המבוסס על מעקב ופיקוח מתמידים

אחת הביקורות המרכזיות בקשר לבינה מלאכותית נוגעת לאופי הפולשני שלה, לאור זאת שחלק גדול מהיישומים דורשים מעקב מתמשך אחר פעולות התלמידים, מחוותיהם ואף רגשותיהם (Miao et al., 2021). שושנה זובוף (Zuboff, 2019) טבעה את המושג "קפיטליזם של מעקב" המציע להסתכל על כלל המערכות הללו כמופעים של שליטה וסטנדרטיזציה, וכפי שניל סלווין כותב בהקשר של מערכות אלגוריתמיות לניתוח למידה (Learning analytics) הרי שבסופו של דבר מערכות אלו עסוקות במהותן בשליטה ובביטוי של כוח (Selwyn, 2020).

כפי שמציין גיליארד (Gilliard and Selwyn, 2022) כמעט כל טכנולוגיה דיגיטלית, ובמיוחד טכנולוגית בינה מלאכותית המבוססת על ריבוי נתונים, משקיעה ומעורבת במעקב ובניטור תלמידים. גיליארד מצביע על הקו הדק שבין ניטור תלמידים לצורך ביטחון ודאגה עבורם ובין מעקב שעניינו הוא שליטה ופיקוח "דכאניים" בדומה למערכות כליאה ובטחון. הוא (ואחרים) טוען שאמנם בתי ספר תמיד היו מעורבים גם בשליטה ופיקוח מהסוג הזה, אך מערכות אלו מגדילות את הנטיות הללו בגלל יכולות הניטור המועצמות שלהן - ברמת כמות הנתונים ואופי המעקב החודרני - ובכך מנרמלות מעקב ופיקוח תמידי במערכות חינוך. גיליארד גם מדגיש את ההיבטים של אפליות והטיות כנגד אוכלוסיות מוחלשות המעורבות במנגנוני פיקוח ומעקב מבוססי טכנולוגיה דיגיטלית לא רק בהיבטים של הטיות אלגוריתמיות (מה שקראנו "הוגנות צרה"), אלא גם באופן שמערכות פיקוח ומעקב מופעלות באופנים שונים ומועצמים כנגד או ביחס לאוכלוסיות מוחלשות ועל פי קווים של גזע ומוצא. הדברים מועצמים במיוחד כאשר אנו בוחנים מערכות מבוססות חיישנים כגון אלו העוסקות בזיהוי פנים, קול ורגשות המיועדות לעקוב אחרי התלמידים באופן מתמשך ומפורט.

4. העדר הוגנות הנובע מפגיעה בסוכנות האנושית

בפרק שדן בתרומות האפשריות של בינה מלאכותית בחינוך ראינו שאחת ההבטחות המרכזיות של טכנולוגיות אלו היא יצירת תהליכי למידה יעילים מותאמים אישית. להבטחה זו יש גם צד שמעלה חששות מכיוון שמערכות שכאלה עלולות לפגוע בהתפתחות יכולות הלמידה העצמאית ובהתפתחות לומד עצמאי, מכיוון שתהליך ניווט הלמידה נלקח מהלומד והמלמד (ברק-מדינה, 2020). אפשר למסגר את החששות הללו כתהליך של פגיעה בסוכנות האנושית, כ-Deskilling ואפילו דה-הומניזציה מכיוון שלפחות חלק ממערכות הבינה מלאכותית דורשת מהמורים ומהתלמידים להשתלב בשיטות הוראה מחייבות, עם אינטראקציה אנושית מינימלית, בעקבות מסלולים מובנים של תוכן, מה שמפחית את סוכנות (Agency) הלומד והמלמד, ואת יכולת הבחירה והמעורבות שלו (Miao et al., 2021). פרסונליזציה מתווכת בינה מלאכותית עלולה לפגוע גם בתפיסת העצמי בעקבות הירידה בסוכנות העצמית ברמה האינדיבידואלית (Kucirkova and Mackey, 2020).

ההשלכות של פגיעה בסוכנות העצמית של תלמידים על הוגנות הן בדיוק תמונת המראה של ההבטחה של תהליכי למידה אוטומטיים מותאמים אישית. החשש המרכזי נוגע לאופן שבו מערכות שכאלה יופעלו בקרב אוכלוסיות מוחלשות ופריפריאליות כדרך להתגבר על פערי משאבים חינוכיים ואולי יובילו לשיפור מסויים בהישגים, אבל יפגעו בפיתוח וטיפוח של לומד עצמאי המוכן למאה ה-21, כאשר לעומתם תלמידים מאוכלוסיות חזקות יותר יזכו בחינוך השם דגש רב יותר על פיתוח כישורי למידה וחשיבה (לדוגמה Liu, 2020).

מימד נוסף של הפגיעה בסוכנות האנושית יכול לבוא לידי ביטוי בהתעלמות מסויימת של מערכות הבינה המלאכותית והוראה אוטומטית מותאמת אישית מהקשרים חברתיים ותרבותיים של הוראה ולמידה. כאשר הפרופיל שמייצרת המערכת בנוגע ללמידה הוא פרופיל אישי המתעלם מההקשרים השונים (Perota and Selwyn, 2020). החשש פה הוא לפגיעה בתלמידים הבאים מהקשרים תרבותיים וחברתיים שונים מתוך כך שהקשרים אלו לא נלקחים בחשבון.

סיכום - מבט מפוכח על השלכות בינה מלאכותית בחינוך על היבטים של הוגנות ושוויון

על אף הפוטנציאל של בינה מלאכותית בחינוך לקדם הוגנות הכלה ושוויון הרי שחוקרים מבקשים בראש ובראשונה למצוא דרכים למנוע ממערכות אלו לחזק ולהעצים הטיות, אפליה ואי-שוויון (לדוגמה: Holstein and Doroudi, 2021). כפי שראינו עמדה זו מתבססת הן על חלק מהתכונות של מערכות בינה מלאכותית - כמו התלות בנתונים קודמים - והן על הניסיון שנצבר ביחס לטכנולוגיות דיגיטליות אחרות והשפעתן על השדה החינוכי.

בפרק השני דנו במטרות החברתיות השונות של הוגנות, הגינות, הכלה ואי שוויון ובאופן שבו ערכים אלו מתורגמים הן לשאלות של הוגנות צרה, כלומר לתפקוד הספציפי של בינה מלאכותית בחינוך וההשפעות שלה, ולשאלות של הוגנות רחבה, העוסקים בשאלות רחבות יותר של ייצור, הפצה ומערכי כוח המעורבים סביב בינה מלאכותית. שני ההיבטים הללו של הוגנות נסקרו בהרחבה בפרקים חמש ושש. לאחר מכן דנו בתרומות האפשריות של בינה מלאכותית לערכי הוגנות ושוויון בחינוך. הצבענו על כך שהפוטנציאל להנגשה, ייעול, התאמה אישית ושיפור הישגים של מערכות אלו עבור תלמידים ותלמידות שונים, כולל תלמידים מרקע מוחלש או בעלי קשיים יכול להוות תרומה משמעותית לשיפור הוגנות, הגינות ושוויון במערכת החינוך. יחד עם זאת, כפי שנסקר בפרק שלאחריו, הבטחות דומות ופוטנציאל דומה היו נחלתן של טכנולוגיות דיגיטליות חינוכיות קודמות והבטחות אלו ברובן הגדול לא התממשו. ולפחות בכל מה שקשור להיבטים של הוגנות, הגינות, הכלה ושוויון לטכנולוגיות חינוכיות יש נטייה לשמר, ולעיתים אף לחזק, אי-שוויון, הטיות ואפליות, וזאת מכיוון שטכנולוגיות אלו נוטות להשתבץ במערכים הקיימים. בכך, שאלת ההשפעה של טכנולוגיות שונות ובתוכן מערכות בינה מלאכותית צריכה לקחת בחשבון שאלות של הקשר מקומי ומערכתי, של יחסים בין הטכנולוגיות ובין מערכים טכנולוגיים ואנושיים אחרים, הן בתהליכי קבלת החלטות, הטמעה ושימוש של הטכנולוגיות.

אם כן, סקירה זו מבססת את ההבנה שכלים דיגיטליים מתוחכמים ועוצמתיים ככל שיהיו לא יכולים לעמוד בפני עצמם כפתרונות חינוכיים חברתיים אלא משתלבים במערכים חינוכיים וחברתיים קיימים ובמדיניות החינוכית הקיימת (לדוגמה: Bulathwela et al., 2021). ההטמעה והשימוש בטכנולוגיות אלו בהקשר של מערכות חינוך הוא עניין מורכב ובלתי צפוי התלוי בהקשרים מקומיים ובגורמים שונים.

יותר מכך, סקירה זו מצביעה על כך שלצד הפוטנציאל החיובי הנרחב של בינה מלאכותית להביא לשיפור וייעול תהליכים חינוכיים, להנגיש אותם לקהלים רחבים ולקדם שירותי חינוך איכותיים לקבוצות אוכלוסייה שונות וביניהן גם קבוצות אוכלוסייה מוחלשות, הרי שבינה מלאכותית מעלה חששות בנוגע להוגנות, הגינות, הכלה ושוויון: ראשית, בנטייה של מערכות בינה מלאכותית לשכפל ואף להעצים הטיות, אפליות ואי-שוויון הקיימים בחברה מתוך ההישענות של מערכות אלו על הנתונים הקיימים. שנית, קיים חשש להטיות ופגיעות הנובעות מתכונות אינהרנטיות של כלי הבינה המלאכותית, כגון מתוך התבססות על נתוני אוכלוסיית הרוב או יצירת מודלים המפשטים את המורכבות החברתית. שלישית, יכולות לצוץ הטיות ופגיעות שונות הנובעות מהממשק שבין מערכות הבינה המלאכותית והמשתמשים האנושיים. כפי שראינו אחת התכונות שמעוררת דאגה היא העדר השקיפות של בינה מלאכותית והקושי האנושי בהבנת ההחלטות האוטומטיות שמערכות אלו מייצרות. תכונה זו מעוררת לא רק שאלות של אמון אלא גם סוגיות הנוגעות ליכולת לייצר בקרה להחלטות ולהמלצות של כלי הבינה המלאכותית. מנקודת מבט של מה שקראנו לאורך סקירה זו "הוגנות רחבה" עולים חששות בנוגע לפגיעה בערכי ההוגנות והשוויון סביב תהליכי הייצור וההפקה של הבינה המלאכותית, סביב שאלות של שליטה ובעלות ומערכי הכוח הכלכליים המעורבים במערכות אלו וכן בפגיעה בסוכנות האנושית הן של מורות והן של תלמידים.

בסופו של דבר ישנה הסכמה רחבה בקרב ארגוני חינוך בינלאומיים העוסקים בנושא כי סוגיות של הוגנות, הגינות, הכלה ושוויון הינן סוגיות פוליטיות וחברתיות שיכולות להיפתר רק במדיניות חינוכית כוללת ומכוונת. טכנולוגיות בינה מלאכותית עשויות לעזור באם הן ישתבצו במערכי מדיניות שכזו, אבל לא סביר שהן יהוו פתרון לבעיות אלו (Miao et al., 2021). ואף יותר מכך כל עוד לא קיימת מדיניות שכזו או שהמערכות הטכנולוגיות לא מהוות חלק ממדיניות כזו, הרי שיש סיכוי סביר שכלי הבינה המלאכותית בחינוך יגבירו את הפגיעה בסוגיות אלו, כפי שמסוכם בדו"ח של אונסקו: "ללא התערבות אפקטיבית של מדיניות, פריסת הבינה המלאכותית בחינוך צפויה לשקף את תהליך יציב ובלתי נמנע של הגדלה ולא של שיפור של אי השוויון הקיים בלמידה" (Miao et al., 2021 p.21). כלומר הם מזוהרים שאם טכנולוגיות בינה מלאכותית בחינוך לא יהיו מתוכננות כראוי ומשולבות עם מדיניות של קידום הוגנות ושוויון הן עלולות להחמיר את אי השוויון החינוכי בהן ברמה המקומית והן ברמה הגלובלית, להסיט משאבים חינוכיים מפתרונות מוכחים להוגנות, הכלה ואי-שוויון ואף ולהפיץ הטיות מסוכנות בקנה מידה רחב (Bulathwela et al., 2021).

לאור זאת, כלל הדוחות והמחקרים בנושא מצביעים על כך שקידום בינה מלאכותית בחינוך צריך לקחת בחשבון הן שאלות של הוגנות, הגינות, הכלה ושוויון במונח הצר והן במונח הרחב, תוך שקילת היתרונות האפשריים של בינה מלאכותית לצד הקשיים שהיא מעוררת ושעלולים לצוץ בעתיד. פתרונות טכנולוגיים-סטטיסטיים המתמודדים עם שאלת ההוגנות, ההכלה והשוויון הצרים, לצד מערכות בקרה הם הכרחיים, אך לא מספיקים (Sahelgren, 2021), ונדרשת פרספקטיבה רחבה יותר שבוחנת את מערכי הכוחות ואת התרבות הארגונית ששימוש בבינה מלאכותית מייצר (Davies et al., 2021).

מקורות

ברק מדינה, ע. (2020). מעבדת בינה מלאכותית בחינוך: דו"ח ביניים כבסיס לחשיבה אסטרטגית וסיכום שנת פעילות תש"פ. אגף מו"פ, משרד החינוך. <https://www.archive.mop.education/wp-content/uploads/-----AI.pdf>

משרד החינוך. (2022). בינה מלאכותית בחינוך: מתווה ההערכות אסטרטגית. מינהל חינוך טכנולוגי, משרד החינוך, ירושלים.

נהון, ק. (עורכת וכותבת ראשית) (2019). דין וחשבון של ועדת משנה של המיזם הלאומי למערכות נבונות בנושא אתיקה ורגולוציה של בינה מלאכותית.

סלומון, גבריאל, 2000. טכנולוגיה וחינוך בעידן המידע. חיפה ותל אביב: זמורה ביתן והוצאת הספרים של אוניברסיטת חיפה.

Andrejevic, M., and Selwyn, N. (2020). Facial recognition technology in schools: Critical questions and concerns. *Learning, Media and Technology*, 45(2), 115-128.

Baker, R. S., and Hawn, A. (2021). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, 1-41

Baker, T., Smith, L., and Anissa, N. (2019). *Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges.* London: Nesta <https://www.nesta.org.uk/report/education-rebooted/>

Barocas, S., Hardt, M., and Narayanan, A. (2017). *Fairness in machine learning.* Nips tutorial, <https://fairmlbook.org/pdf/fairmlbook.pdf>

Benjamin, R. (Ed.) (2019). *Captivating technology.* Durham: Duke University Press.

Berendt, B., Littlejohn a., and Blakemore, M. (2020) AI in education: learner choice and fundamental rights. *Learning, Media and Technology*, 45:3, 312-324, DOI: 10.1080/17439884.2020.1786399

Bulathwela, S., Pérez-Ortiz, M., Holloway, C., and Shawe-Taylor, J. (2021). Could AI Democratise Education? Socio-Technical Imaginaries of an EdTech Revolution. arXiv preprint arXiv:2112.02034.

Bristol, D., and Shawn, Z. (2022). *Artificial intelligence in education: Should students pay the price for Algorithmic bias.* AI and Ethics, <https://doi.org/10.1007/s43681-021-0065096-7>

Castañeda, L., and Williamson, B. (2021). Assembling New Toolboxes of Methods and Theories for Innovative Critical Research on Educational Technology. *Journal of New Approaches in Educational Research*, 1-14.

Chan, A. (2019). Venture ed: Recycling hype, fixing futures, and the temporal order of edtech. *DigitalSTS: A field guide for science and technology studies*, 161-177.

Chen, L., Chen, P., and Lin, Z. (2020). Artificial intelligence in education: A review. IEEE Access, 8, 75264–75278. <https://doi.org/10.1109/ACCESS.2020.2988510>

Crawford, K. (2019). *The Atlas of AI.* Yale University Press.

- Cuban, Larry. 2003. *Oversold and Underused: Computers in the Classroom*. MA: Harvard University Press
- Davies, H. C., Eynon, R., and Salveson, C. (2021). [The mobilisation of AI in education: A Bourdieusean field analysis](#). *Sociology*, 55(3), 539-560.
- Dignum, V. (2021). [The role and challenges of education for responsible AI](#). *London Review of Education*, 19 (1), 1, 1–11. <https://doi.org/10.14324/LRE.19.1.01>
- Dixon-Román, E., Philip N. T. and Nyame-Mensah, A. (2019). [The racializing forces of/in AI educational technologies](#). *Learning, Media and Technology*, DOI: 10.1080/17439884.2020.1667825
- Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., ... and Warschauer, M. (2020). [Mining big data in education: Affordances and challenges](#). *Review of Research in Education*, 44(1), 130-160.
- Gilliard, C. and Selwyn, N. (2022). [Automated Surveillance in Education](#). *Postdigital Science and Education: 1-11*.
- Hamilton, A., and Hattie, J. (2021). [Not All That Glitters Is Gold: Can Education Technology Finally Deliver?](#) Corwin Press.
- Holland, B. (2020). Artificial Intelligence (AI) in K-12. Version 1.0. Consortium for School Networking (CoSN) in <https://www.cosn.org/tools-and-resources/resource/artificial-intelligence-ai-in-k-12/>
- Holstein, K., and Doroudi, S. (2021). [Equity and Artificial Intelligence in Education: Will" AIEd" Amplify or Alleviate Inequities in Education?](#). *arXiv preprint arXiv:2104.12920*.
- Hosnagar, K. (2019). *A human's guide to machine intelligence*. New York: Viking.
- Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y. S., Kay, J., ... & Gašević, D. (2022). [Explainable Artificial Intelligence in education](#). *Computers and Education: Artificial Intelligence*, 3, 100074.
- Kizilcec, R. F. and Lee, H. (2020). [Algorithmic Fairness in Education](#). In W. Holmes and K. Porayska-Pomsta (Eds.), *Ethics in Artificial Intelligence in Education*, Taylor and Francis
- Knox, J. (2020). [Artificial intelligence and education in China](#). *Learning, Media and Technology*, 45(3), 298-311.
- Kucirkova, N., and Mackey, M. (2020). [Digital literacies and children's personalized books: Locating the 'self'](#). *London Review of Education*, 18(2), 151-162.
- Leaton Gray, S. and Kucirkova, N. (2021). [AI and the human in education: Editorial](#). *London Review of Education*, 19 (1), 10, 1–4. <https://doi.org/10.14324/LRE.19.1.10>
- Liu Y. L.(2020) The Future of the Classroom? China's experience of AI in education in Chun, A., Ding, J., Creemers, R., Gal, D., Han, E., Liu, Y. L., and Lewis, D. (2020). [The AI powered state: China's approach to public sector innovation](#).
- Loukina, A., Madnani, N., and Zechner, K. (2019). [The many dimensions of algorithmic fairness in educational applications](#). In *Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 1-10).

- Luckin, R., Holmes, W., Griffiths, M., and Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson. <https://static.googleusercontent.com/media/edu.google.com/en//pdfs/Intelligence-Unleashed-Publication.pdf>
- Marcus, G. (2022). Deep Learning Is Hitting a Wall. *Nautilus*. https://nautil.us/deep-learning-is-hitting-a-wall-14467/?_sp=baa47b0d-796f-4e82-a2ea-a3d30cf85973.1649588017885
- Miao, F., Holmes, W., Huang, R., and Zhang, H. (2021). *AI and education: A guidance for policymakers*. UNESCO Publishing. <https://unesdoc.unesco.org/ark:/48223/pf0000376709>
- Mitchell, S., Potash, E., Barocas, S., D'Amour, A., and Lum, K. (2021). Algorithmic Fairness: Choices, Assumptions, and Definitions. *Annual Review of Statistics and Its Application*, 8. <https://doi.org/10.1146/annurev-statistics-042720-125902>
- Narayanan, A. (2018). "21 Fairness Definitions and Their Politics". Tutorial at the Conference on Fairness, Accountability, and Transparency. <https://www.youtube.com/embed/jiXluYdnyyk>
- Nazaretsky, T., Cukurova, M., Ariely, M., and Alexandron, G. (2021). [Confirmation bias and trust: Human factors that influence teachers' attitudes towards AI-based educational technology](#). In *CEUR Workshop Proceedings* (Vol. 3042).
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.
- O'Neill, C., Selwyn, N., Smith, G., Andrejevic, M., and Gu, X. (2022). The two faces of the child in facial recognition industry discourse: biometric capture between innocence and recalcitrance. *Information, Communication and Society*, 1-16.
- Parapadakis, D. (2020). [Can Artificial Intelligence Help Predict a Learner's Needs? Lessons from Predicting Student Satisfaction](#). *London Review of Education*, 18(2), 178-195.
- Pedro, F., Subosa, M., Rivas, A., and Valverde, P. (2019). *Artificial intelligence in education: Challenges and opportunities for sustainable development*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000366994>
- Perrotta, C. and Selwyn, N. (2019). [Deep learning goes to school: toward a relational understanding of AI in education](#), *Learning, Media and Technology*, DOI: 10.1080/17439884.2020.1686017
- Puckett, C., and Rafalow, M. H. (2020). From "Impact" to "Negotiation". In *The Oxford handbook of sociology and digital media*.
- Rafalow, M. H., and Puckett, C. (2021). [Sorting Machines: Digital Technology and Categorical Inequality in Education](#). *Educational Researcher*, 0013189X211070812.
- Reich, J. (2020). *Failure to Disrupt: Why Technology Alone Can't Transform Education*. Harvard University Press.
- Reich, J. and Ito, M. 2017. *From Good Intentions to Real Outcomes: Equity by Design in Learning Technologies*. Irvine, CA: Digital Media and Learning Research Hub. <https://clalliance.org/publications/good-intentions-real-outcomes-equity-design-learning-technologies/>
- Sahlgren, O. (2021). [The politics and reciprocal \(re\)configuration of accountability and fairness in data-driven education](#). *Learning, Media and Technology*. DOI: 10.1080/17439884.2021.1986065
- Saltman, K.J. (2020). [Artificial intelligence and the technological turn of public education privatization: In defence of democratic education](#). *London Review of Education*, 18 (2): 196–208. <https://doi.org/10.14324/LRE.18.2.04>

- Selwyn, N. (2020). Re-imagining 'Learning Analytics'... a case for starting again?. *The Internet and Higher Education*, 46, 100745.
- Selwyn, N. (2021). AI' in education: two sides of a conversation. In *Critical Studies of Education and Technology*. <https://criticaledtech.com/2021/11/26/ai-in-education-two-sides-of-a-conversation/>
- Selwyn, N. (2022). [Less work for teacher? The ironies of automated decision-making in schools](#). in Pink, S., Berg, M., Lupton, D. and Ruckenstein, M. (Eds). *Everyday automation: experiencing and anticipating automated decision-making*. Routledge
- Selwyn, N., Campbell, L., and Andrejevic, M. (2022). Autoroll: scripting the emergence of classroom facial recognition technology. *Learning, Media and Technology*, 1-14.
- Sun, M., Oliwa, T., Peek, M. E., and Tung, E. L. (2022). [Negative Patient Descriptors: Documenting Racial Bias In The Electronic Health Record: Study examines racial bias in the patient descriptors used in the electronic health record](#). *Health Affairs*, 10-1377
- Tamim, R. M., Borokhovski, E., Bernard, R. M., Schmid, R. F., Abrami, P. C., and Pickup, D. I. (2021). [A study of meta-analyses reporting quality in the large and expanding literature of educational technology](#). *Australasian Journal of Educational Technology*, 37(4), 100–115. <https://doi.org/10.14742/ajet.6322>
- Verma, S., and Rubin, J. (2018). [Fairness definitions explained](#). In *2018 IEEE/ACM International Workshop on Software Fairness (Fairware)* (pp. 1-7). IEEE.
- Warschauer, M., and Matuchniak, T. (2010). [New technology and digital worlds: Analyzing evidence of equity in access, use, and outcomes](#). *Review of research in education*, 34(1), 179-225.
- Warschauer, M., Knobel, M., and Stone, L. (2004). [Technology and equity in schooling: Deconstructing the digital divide](#). *Educational policy*, 18(4), 562-588.
- Williamson, B., and Eynon, R. (2020). [Historical threads, missing links, and future directions in AI in education](#). *Learning, Media and Technology*, 45(3), 223-235
- Ydesen, C., Acosta, F., Milner, A.L., Ruan, Y., Aderet-German, T., Gomez Caride, E. and Hansen, I. S. 2020. [Inclusion in testing times – implications for citizenship and participation](#). Paper commissioned for the UNESCO Futures of Education report.
- Zhao, S. (2021). [Facial Recognition in Educational Context](#). In *2021 International Conference on Public Relations and Social Sciences (ICPRSS 2021)* (pp. 10-17). Atlantis Press.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.