

●●● היררכיה של הסקה: מסגרת I-DIRECT להגדרת מעורבות אנושית בזיהוי ודירוג של אינדיקטורים תפיסתיים בשיח פתוח של פרחי הוראה, בעזרת בינה מלאכותית



הכנס הארצי - שנתי ה-24 של מיט"ל 2026

עדי יעקב עזריה

המעבדה לחקר שילוב טכנולוגיות בלמידה

המרכז לחינוך מדעי וטכנולוגי

בית הספר לחינוך

אוניברסיטת תל אביב



חוקרים



פרופ' ענת כהן

ראשת מעבדת המחקר AL&T
מרכז החינוך למדעים
וטכנולוגיה בית הספר לחינוך
אוניברסיטת תל אביב



גיא כהן

סטודנט לדוקטורט וחוקר,
מעבדת מחקר AL&T
מרכז החינוך למדע
וטכנולוגיה בית הספר לחינוך
אוניברסיטת תל אביב



מרב רוטרי-סבן

סטודנטית לדוקטורט וחוקרת,
מעבדת המחקר AL&T
מרכז החינוך למדע וטכנולוגיה
בית הספר לחינוך אוניברסיטת
תל אביב



ד"ר אלה ברונשטיין

ראשת היחידה לקידום
איכות ההוראה, המחקר
וההערכה, הפקולטה
למדעי הרפואה והבריאות
ע"ש גר"י



עדי יעקב-עזריה

חוקרת, מעבדת מחקר AL&T
מרכז החינוך למדע
וטכנולוגיה בית הספר לחינוך
אוניברסיטת תל אביב



מחקר זה נתמך על ידי פרויקט הממומן על ידי האיחוד האירופי
במסגרת תוכנית ארסמוס (הסכם מענק מס' 1-LV01-2024-
KA220-HED-000250857).



AI-EmpaTe



AI-EmpaTe

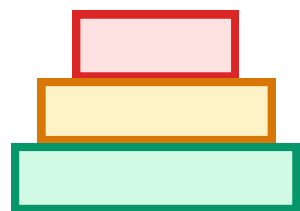
AI-EMPOWERED TEACHING

UNLOCKING THE FUTURE OF DIGITAL EDUCATION FOR TEACHERS



Co-funded by
the European Union





מיטל 2026 | היררכייה של הסקה

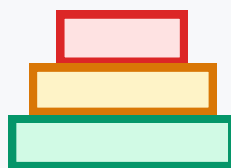
רקע: שימוש ב-TAM על מנת לבחון תפיסות מורים

מודל קבלת הטכנולוגיה (TAM)



Davis, 1989

המגבלה: TAM נמדד בדרך כלל באמצעות שאלונים סגורים, דבר המגביל מאוד את עומק ההסקה.



הזדמנויות ואתגרים בשימוש בפתוח כמרחב מחקר

אתגרים

ריבוי משמעויות ושונות בין נבדקים

נדרש ניתוח מעמיק - דבר הדורש משאבים רבים

אולם

הזדמנויות

משקף את תפיסות המורים

השפה משקפת עמדות וזהות מקצועית

מקור מידע עשיר על האופן שבו מורים תופסים ומתייחסים ל-GenAI



הזדמנויות ואתגרים בשימוש בשיח פתוח כמרחב מחקר

אתגרים

ריבוי משמעויות ושונות בין נבדקים

נדרש ניתוח מעמיק - דבר הדורש משאבים רבים

אולם

הזדמנויות

משקף את תפיסות המורים

השפה משקפת עמדות וזהות מקצועית

מקור מידע עשיר על האופן שבו מורים תופסים ומתייחסים ל-GenAI

הפתרון הפוטנציאלי: שימוש ב-LLM

השאלה המרכזית: מהן המגבלות, ההטיות - ומה התפקיד המדויק של האדם?



רמה 1: ראייה ישירה

"השתמשתי ב-GenAI ליצירת חומרי לימוד מותאמים אישית לתלמידים שלי, כמו דפי עבודה, סיכומים ושאלות חזרה. הכלים עזרו לי לפשט נושאים מורכבים ולהנגיש אותם בצורה יצירתית. לדוגמא, התלמידים קיבלו תרגולים דיגיטליים עם משוב מיידי בהתאם לצרכים שלהם. לעיתים המידע אינו מדויק ודורש אימות. לכן אני בודקת את המידע מול מקורות אמינים ומלמדת את התלמידים לחשוב ביקורתית".

קשה לעיתים להתאים תוכן באופן מלא לצרכים מגוונים. לכן אני משתמשת ב-GenAI כבסיס ומתאימה ידנית לפי רמת התלמידים והמשובים.

לגבי דאגות לגבי פרטיות ושימוש הוגן בטכנולוגיה, אני בוחרת כלים שמכירים בפרטיות נתונים ונמנעת מהעלאת מידע אישי של תלמידים".

שימוש בפועל

מגבלות,
חסמים ודרכי
התמודדות



“השתמשתי ב-GenAI ליצירת חומרי לימוד מותאמים אישית לתלמידים שלי, כמו דפי עבודה, סיכומים ושאלות חזרה. הכלים עזרו לי לפשט נושאים מורכבים ולהנגיש אותם בצורה יצירתית. לדוגמא, התלמידים קיבלו תרגולים דיגיטליים עם משוב מיידי בהתאם לצרכים שלהם.

לעיתים המידע אינו מדויק ודורש אימות. לכן אני בודקת את המידע מול מקורות אמינים ומלמדת את התלמידים לחשוב ביקורתית.

קשה לעיתים להתאים תוכן באופן מלא לצרכים מגוונים. לכן אני משתמשת ב-GenAI כבסיס ומתאימה ידנית לפי רמת התלמידים והמשובים.

לגבי דאגות לגבי פרטיות ושימוש הוגן בטכנולוגיה, אני בוחרת כלים שמכירים בפרטיות נתונים ונמנעת מהעלאת מידע אישי של תלמידים.”



Self
Direction

“השתמשתי ב-GenAI ליצירת חומרי לימוד מותאמים אישית לתלמידים שלי, כמו דפי עבודה, סיכומים ושאלות חזרה. הכלים עזרו לי לפשט נושאים מורכבים ולהנגיש אותם בצורה יצירתית. לדוגמא, התלמידים קיבלו תרגולים דיגיטליים עם משוב מיידי בהתאם לצרכים שלהם.

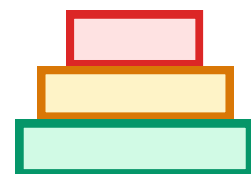
לעיתים המידע אינו מדויק ודורש אימות. לכן אני בודקת את המידע מול מקורות אמינים ומלמדת את התלמידים לחשוב ביקורתית.

קשה לעיתים להתאים תוכן באופן מלא לצרכים מגוונים. לכן אני משתמשת ב-GenAI כבסיס ומתאימה ידנית לפי רמת התלמידים והמשובים.

לגבי דאגות לגבי פרטיות ושימוש הוגן בטכנולוגיה, אני בוחרת כלים שמכירים בפרטיות נתונים ונמנעת מהעלאת מידע אישי של תלמידים.”

תועלת נתפסת

מאמץ נתפס



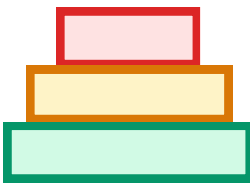
היררכיית ההסקה





מסגרת המחקר: היררכיית ההסקה ועקרונות מתודולוגיים





שאלות המחקר

שאלת על

באילו תנאים ובאיזו מידה ניתן להשתמש ב-GenAI לזיהוי ולדירוג אינדיקטורים תפיסתיים בשיח פתוח של פרחי הוראה - במיוחד כאלה הדורשים רמות גבוהות של היסק ופרשנות?

2

שאלת משנה

Few-shot: האם ההנחיה מסייעת?

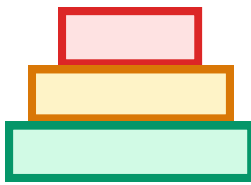
האם ובאיזו מידה משתנה רמת ההסכמה בין מדרגים אנושיים ל-GenAI בעקבות שימוש באסטרטגיית הנחיה מונחית-דוגמאות, על פני הרמות השונות?

1

שאלת משנה

Zero-shot: מהי רמת ההסכמה?

מהי רמת ההסכמה בין מדרגים אנושיים ל-GenAI בדירוג אינדיקטורים תפיסתיים, על פני שלוש רמות היררכיית ההיסקיות - תוך שימוש באסטרטגיית הנחיה ללא דוגמאות?



שיטה: מדגם, הליך ומשתני מחקר

משתתפים ומקור נתונים

ניתוח הנתונים התבצע באמצעות קפא משוקלל של כהן

50 (7 הושמטו) פרחי הוראה השיבו על שלוש שאלות פתוחות

משתני המחקר:

10 משתנים ב-3 רמות הסקה

מידע בלתי מספיק = לא ניתן לניתוח

Few-shot | לאחר אימון המודל

שלב 2

מה עשינו?

- סופקו למודל מחוון אופרטיבי מפורט + 5 דוגמאות שקודדו על ידי מומחים
- תהליך איטרטיבי אישר את הבנת המכונה
- ChatGPT קידד ונימק מחדש את אותן תגובות מהשלב הקודם

מטרה

לבחון אם אימון המודל (מחוון + דוגמאות) משפרים את ההלימה בין המכונה למומחים על פני שלוש רמות ההיסקיות.

Zero-shot | הערכת בסיס

שלב 1

מה עשינו?

- הוזנו ל-ChatGPT5.3 הגדרות וקווים כלליים בלבד - ללא דוגמאות
- תגובות המשתתפים הוזנו, קודדו ונומקו על ידי המכונה
- הושאו לקידוד מומחים אנושיים

מטרה

קביעת נקודת בסיס (Baseline) לביצועי המכונה תחת מינימום התערבות אנושית, לצורך השוואה עם השלב הבא.

מידת ההסכמה לאורך שלוש הרמות לפני האימון

רמה 3: $k=0.200$

GENAI קבע שאין מספיק מידע
עבור 34% מהמקרים (בהשוואה
ל-8% בקידוד האנושי)

רמה 2: $k=0.195$

GENAI קבע שאין מספיק מידע עבור 43%
מהמקרים (בהשוואה ל-20% בקידוד
האנושי)

רמה 1: $k=0.385$



מיטל 2026 | היררכייה של הסקה

מידת ההסכמה לאורך שלוש הרמות לאחר אימון - האם ההנחייה מסייעת?

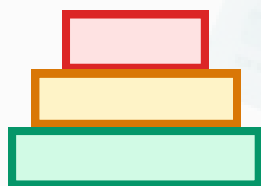
רמה 3: k עלה ל- 0.392
אחוז מקרי 'אין מספיק מידע'
ירד ל-17%

רמה 2: k עלה ל- 0.338
אחוז מקרי 'אין מספיק מידע'
ירד ל-20%

רמה 1: k עלה ל- 0.434



מיטל 2026 | היררכייה של הסקה



איפה החוקר בלולה?

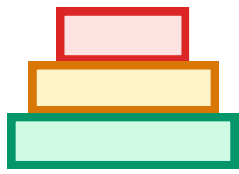
פרדוקס היעילות

1. פיקוח אנושי הכרחי - אך הוא מגביל בהכרח את היעילות ש-AI היה אמור לספק.
2. הצורך במעורבות אנושית עולה ככל שעולה מידת ההסקה הנדרשת

המסקנה: משתנים שונים דורשים מידה שונה של מעורבות אנושית

- ביצועי AI טובים יותר ברמה המפורשת
- הנחייה משפרת ביצועים בשלוש הרמות, הן בהקשר לרמת התאימות, והן מבחינת כיסוי ההסקה

לכן נדרשת מסגרת עבודה מכוילת: כזו שמתאימה את מידת המעורבות האנושית לרמת ההיסקיות, במקום יישימות אחידה של ה-AI.



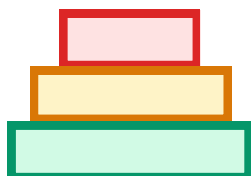
I-DIRECT: מסגרת ממוקדת-החלטות לניתוח איכותני נתמך בינה מלאכותית



חוקר אנושי - מוביל ונוכח בכל השלבים

בינה מלאכותית - פעילה/תומכת בשלבים 2-4, תומכת באופן פוטנציאלי בשלבים 1,5

H = תפקיד אנושי = תפקיד הבינה המלאכותית = לולאת משוב = מרכיב אנושי פעיל = בינה מלאכותית פעילה = תמיכה בבינה מלאכותית



מסקנות

ממצא מרכזי



יכולת הזיהוי של אינדיקטורים תפיסתיים בסיוע AI משתנה בהתאם לדרישות ההיסק ולאסטרטגיית ההנחיה: ביצועים חזקים יותר בעדות ישירה ולאחר הנחייה מפורטת עם דוגמאות.

תרומה תיאורטית



היררכיית היסקיות המקשרת בין רמות עמימות לבין מידת הפרשנות האפשרית של אינדיקטורים המבוססים על שיח: ככל שהקונסטרוקט משתמע יותר - פחות ניתן לסמוך על AI ויותר נדרש שיפוט אנושי.

תרומה יישומית



מסגרת I-DIRECT - מסגרת מכוונת-החלטה ליישום GenAI בניתוח איכותני הקושרת בין דרישות ההיסק לבין חלוקת התפקידים בין אדם ל-AI.

מחקרי המשך



לחדד את הגבולות והתנאים לשימוש ב-AI בניתוח איכותני, ולבחון כיצד ניתן לנבא את רמות המעורבות האנושית הנדרשות על בסיס היררכיית הקונסטרוקטים המוצעת.



שאלות ודיון

ביה"ס לחינוך | אוניברסיטת תל-אביב

adiyaakov@mail.tau.ac.il / meravrotary@mail.tau.ac.il / allab@tauex.tau.ac.il

תודות: מחקר זה נתמך במסגרת פרויקט Al-EmpaTe, במימון האיחוד האירופי (Erasmus)

תוצאות: קפא משוקלל של כהן - לפני ואחרי ההנחיה

קפא משוקלל של כהן (K); % - שיעור התגובות שסווגו כ'מידע בלתי מספיק' | * $p < .001$ ** $p < .01$ *** $p < .05$

עם דוגמאות Few-shot			ללא דוגמאות Zero-shot			משתנה	רמה
אדם %	AI %	K	אדם %	AI %	K		
11	0	0.577***	0	0	0.434***	שימוש בפועל	עדות ישירה
0	0	0.352***	0	0	0.262*	חסמים וצרכים	
0	0	0.338***	0	5	0.464***	מיקוד בסיכונים	
0	0	0.716***	25	0	0.491***	דרכי התמודדות	
0	0	0.184*	0	2	0.276**	צורך בפיתוח מקצועי	
20	32	0.338***	20	43	0.195*	הכוונה עצמית של המורה	עדות מסונתזת
7	9	0.411***	7	20	0.287***	תועלת נתפסת (PU)	פרשנות קונסטרוקטים (TAM)
16	34	0.157	16	84	-0.667	קלות שימוש נתפסת (PEOU)	
2	16	0.522***	2	23	0.577***	כוונה התנהגותית (BI)	
5	9	0.481***	5	7	0.604***	עמדה (ATT)	

גבוהה מאוד | 0.61-0.80 גבוהה | 0.41-0.60 בינונית | 0.21-0.40 נמוכה | 0.00-0.20 נמוכה מאוד 0.81-1.00 K (Landis & Koch, 1977) פירוש ערכי