

הערכה אקדמית בסיוע AI: תפיסות מרצים, אתגרים והיבטים אתיים (מאמר קצר)

עדינה משעל
סמינר הקיבוצים
adina.mishal@smkb.ac.il

לאורה ליון
מכללת סמינר הקיבוצים
Laura.levin@smkb.ac.il

תמי זיפרט
מכללת סמינר הקיבוצים
tami.seifert@smkb.ac.il

חן גלעדי
המכללה האקדמית להנדסה על שם סמי
שמעון SCE
chengi1@sce.ac.il

AI-Assisted Assessment in Higher Education: Lecturers' Perceptions, Challenges, and Ethical Concerns (Short paper)

Laura Levin
Kibbutzim College of Education,
Technology and the Arts
laura.levin@smkb.ac.il

Adina Mishal
Kibbutzim College of Education,
Technology and the Arts
adina.mishal@smkb.ac.il

Chen Giladi
SCE Shamoon College of
Engineering
chengi1@sce.ac.il

Tami Seifert
Kibbutzim College of Education,
Technology and the Arts
tami.seifert@smkb.ac.il

Abstract

This study examines lecturers' perceptions of integrating generative artificial intelligence (GenAI) tools in academic assessment. The study employed a quantitative design with 112 lecturers from Israeli higher education institutions, using an online questionnaire. Four perception scales were examined: perceived contribution of GenAI to assessment, perceived enabling conditions, attitudes and self-efficacy regarding the integration of the technology, and ethical concerns. Data analysis included descriptive statistics, correlation tests, and *t*-tests.

The findings indicate extensive use of GenAI for teaching purposes, but limited use in assessment. Even when GenAI is incorporated in the assessment of student work, it primarily supports preparatory stages and the formulation of feedback, rather than final grading decisions. Lecturers with experience in GenAI-based assessment reported more positive perceptions of GenAI's contribution, enabling conditions, and their ability to use the tools, compared to those without such experience. However, ethical concerns,

including fairness, transparency, privacy, and professional responsibility, remained high and consistent across both groups.

The findings highlight a growing gap between the expanding use of GenAI in instructional contexts and the insufficient institutional readiness to support its responsible use in student assessment. Clear and updated institutional policy is required to guide the ethical integration of GenAI in assessment, enabling professional development, structured guidance, and the preservation of human judgment.

Keywords: Artificial Intelligence, Academic Assessment, Faculty Perceptions, Ethics in Higher Education, Technology Adoption.

תקציר

מחקר זה בוחן את תפיסות מרצים בהשכלה הגבוהה כלפי שילוב כלי בינה מלאכותית גנרטיבית (GenAI) בהערכה אקדמית. המחקר נערך בגישה כמותית בקרב 112 מרצים ממוסדות להשכלה גבוהה בישראל, באמצעות שאלון מקוון. נבחנו ארבעה מדדי תפיסה: תפיסת תרומת ה-AI להערכה, קיום תנאים מאפשרים, עמדות ומסוגלות לשילוב הטכנולוגיה, וחששות אתיים. ניתוח הנתונים כלל סטטיסטיקה תיאורית, בדיקות מתאם ומבחני t.

הממצאים מצביעים על שימוש נרחב ב-AI לצורכי הוראה, אך שימוש מוגבל בהערכה. גם כאשר AI משולב בהערכת עבודות, הוא משמש בעיקר לתמיכה בשלבי הכנה ובניסוח משוב, ולא בקבלת החלטות סופיות על ציונים. מרצים בעלי ניסיון בהערכה מבוססת AI דיווחו על תפיסות חיוביות יותר של התרומה, תנאים מאפשרים ויכולת השימוש בכלים, בהשוואה למרצים ללא ניסיון. עם זאת, החששות האתיים, הכוללים סוגיות של הוגנות, שקיפות, פרטיות ואחריות מקצועית, נשארו גבוהים ואחידים בין הקבוצות. הממצאים מדגישים פער מתרחב בין הרחבת השימוש ב-AI בשדה ההוראה לבין היערכות מוסדית מספקת לתמיכה בשימוש בהערכת סטודנטים. נדרשת מדיניות מוסדית ברורה ועדכנית לשילוב אחראי של AI בהערכה, המאפשרת פיתוח מקצועי, הנחיה מוסדרת ושימור שיקול הדעת האנושי.

מילות מפתח: בינה מלאכותית, הערכת עבודות, תפיסות מרצים, אתיקה בהשכלה גבוהה, אימוץ טכנולוגיות.

מבוא

כלי בינה מלאכותית גנרטיבית הולכים ונעשים נוכחים בהשכלה הגבוהה ומשנים תהליכים מרכזיים בהוראה ובלמידה. בתחום ההערכה האקדמית, כלי GenAI מציעים תמיכה ביצירת משוב, בבדיקת התאמה למחונן, ובהענקת הנחיה פורמטיבית (Lim et al., 2023; Merino-Campos, 2025). בהקשר זה, המחקר מתמקד בהערכת תוצרי למידה של סטודנטים – כגון עבודות כתובות, מטלות ופרויקטים – ובתהליכי מתן משוב וציון. יכולות אלה מבטיחות שיפור ביעילות ואיכות המשוב, אך במקביל מעוררות שאלות בנוגע להוגנות, לשקיפות ולשימור שיקול הדעת האנושי בתהליכי הערכה אקדמיים (Holmes et al., 2022). לנוכח התרחבות השימוש בכלי GenAI במוסדות אקדמיים, עולה הצורך להבין כיצד מרצים תופסים ומיישמים בפועל הערכה הנתמכת ב-GenAI. הספרות מצביעה על כך שרוב יישומי GenAI בהשכלה הגבוהה מתמקדים בהוראה, בלמידה ובמשימות תומכות, בעוד שהידע על פרקטיקות המרצים בתחום ההערכה של תוצרי למידה עדיין מוגבל (Kasneci et al., 2023; Zawacki-Richter et al., 2019). מחקרים עדכניים יותר מתחום ההערכה בהשכלה גבוהה מראים כי מרצים נוטים לאמץ GenAI בשלבי הערכה תומכים כגון בעיצוב מחוונים, ניסוח משוב ועריכת לשון, אך מהססים להפקיד בידי ה-GenAI את ההכרעה הסופית בנוגע לציונים. היסוס זה קשור בעיקר לחששות הנוגעים להוגנות, אחריותיות ושקיפות (Bearman et al., 2024; Holmes et al., 2022; Lim et al., 2023). אף שמרצים רבים מזהים פוטנציאל של כלי GenAI להפחתת עומס העבודה ולאפשר מתן משוב אישי ומפורט יותר, חוסר נוחות אתית עשוי לעכב את אימוץ השימוש בהם (Al-Zahrani & Alasmari, 2024; Dabis & Csáki, 2024).

מחקרים אחרים מצביעים על מתח בין אימוץ פרטני מהיר של כלי GenAI על ידי מרצים לבין היעדר הנחיות מוסדיות ברורות ומדיניות מוסכמת. מתח זה בולט גם במערכת ההשכלה הגבוהה בישראל, שבה השימוש ב-GenAI מתרחב

במהירות, בעוד שהנחיות מוסדיות בנוגע לשילובו בהערכה אקדמית נותרות בלתי אחידות בין מוסדות (Council for Higher Education in Israel, 2023).

אף שהנחיות מוסדיות לשילוב GenAI בהערכה הופכות נפוצות יותר (Moorhouse et al., 2023), המחקר הקיים מתמקד בעיקר במדיניות ארגונית, במוכנות מוסדית או בתפיסות סטודנטים, בעוד שתפיסות המרצים עצמם, ובפרט הקשר בין ניסיון מעשי בהערכת עבודות באמצעות GenAI לבין עמדותיהם, תחושת המסוגלות שלהם והחששות האתיים שהם מבטאים, נותרו פחות נחקרות.

מטרת המחקר היא לבחון את תפיסות המרצים כלפי שילוב כלי GenAI בהערכת תוצרי למידה של סטודנטים, ולהעמיק בהבנת הקשר בין ניסיון מעשי בשימוש ב-GenAI להערכה לבין תפיסות אלו. כמו כן, המחקר שואף לחשוף את האופן שבו מרצים תופסים את ההיבטים האתיים של הערכה מבוססת GenAI. מתוך מטרה זו, נגזרו שלוש שאלות המחקר הבאות:

1. לאילו מטרות עושים המרצים שימוש ב-GenAI בהערכת תוצרי למידה של סטודנטים?
2. מהן תפיסות המרצים כלפי שימוש ב-GenAI לצורכי הערכת עבודות סטודנטים?
3. האם קיימים הבדלים בתפיסות בין מרצים בעלי ניסיון בשימוש ב-GenAI להערכה לבין מרצים ללא ניסיון כזה?
4. כיצד מרצים תופסים את ההיבטים האתיים של הערכה מבוססת GenAI?

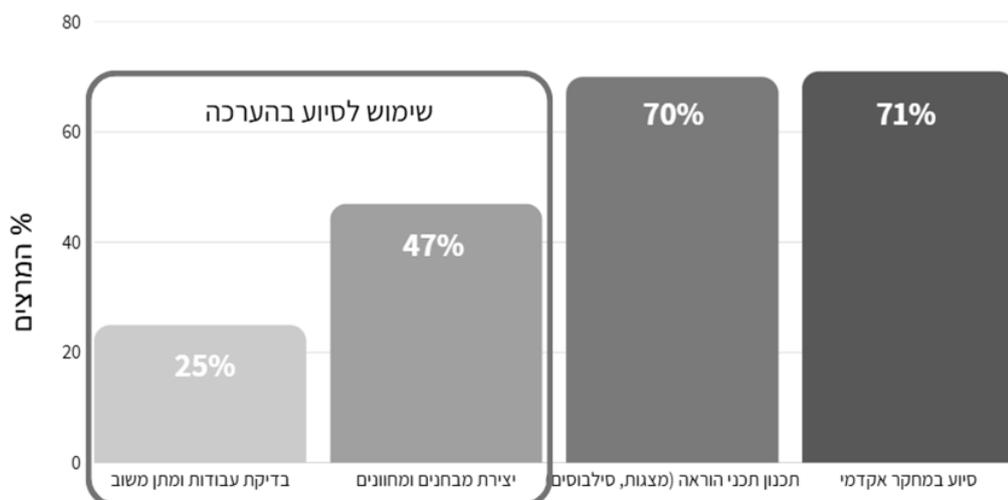
מתודולוגיה

במחקר השתתפו 112 מרצים ממוסדות להשכלה גבוהה בישראל. הנתונים נאספו באמצעות שאלון מקוון שפותח לצורך מחקר זה על בסיס סקירת הספרות וייעוץ עם מומחים בתחום ההערכה האקדמית והטכנולוגיה בחינוך. השאלון כלל 62 פריטים שחולקו לשישה חלקים: פרטים דמוגרפיים ורקע מקצועי; תדירות ומטרות השימוש ב-GenAI; שימושים ספציפיים ב-AI להערכה; ניסיון והכשרה בתחום; וארבע תת-סקאלות של תפיסות ועמדות כלפי הערכה באמצעות GenAI. השאלון הופץ ביולי-אוגוסט 2025. הנתונים נותחו באמצעות סטטיסטיקה תיאורית, מתאמי פירסון ומבחני t להשוואת תפיסות בין מרצים עם וללא ניסיון ב-GenAI להערכה.

ממצאים

תדירות ומטרות השימוש של מרצים בכלי AI בכלל ובהערכה אקדמית בפרט

נמצא שמרבית המרצים (92%) משתמשים ב-GenAI בהוראה לפחות פעם בשבוע, 60% מתוכם מדווחים על שימוש יומיומי. אך כשהתבקשו לדווח על השימוש בכלי GenAI לצורכי הערכה, 37% דיווחו כי מעולם לא השתמשו ב-GenAI למטרה זו.



איור 1. שימוש ב-GenAI בקרב מרצים לפי סוג הפעילות

בנוסף, כ-70% מהמרצים דיווחו כי הם עושים שימוש שוטף בכלים מבוססי GenAI להכנת מצגות, סילבוסים וחומרי למידה. לעומת זאת, כאשר מתקיים שימוש בהערכה, הוא מתמקד בעיקר (47%) בשלבים תומכים, להכנת כלי הערכה כגון יצירת מחוונים וניסוח שאלונים. רק 25% מהמרצים דיווחו על ניסיון בשימוש ב-GenAI להערכת עבודות הסטודנטים ומתן משוב (איור 1).

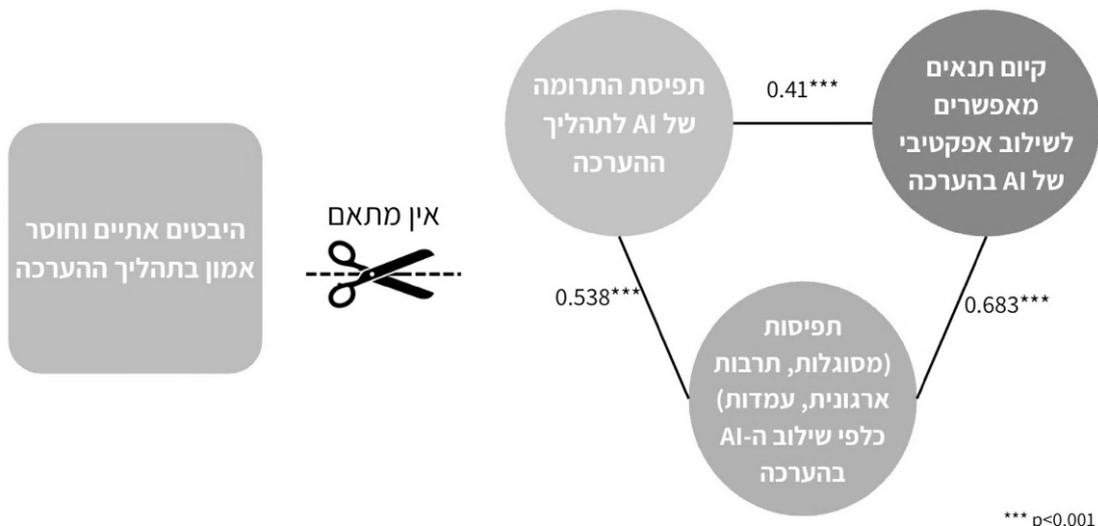
תפיסות המרצים כלפי הערכה באמצעות GenAI

טבלה 1. תפיסות מרצים אודות השימוש ב AI לצרכי הערכת עבודות סטודנטים, ממוצעים וסטיית תקן

ממד תפיסה	α של קורנבאך	תיאור	ממוצע (M)	סטיית תקן (S.D)
תרומה של AI לתהליך ההערכה	0.91	יעילות, עקביות, בהירות המשוב, חסכון בזמן	2.75	1.17
קיום תנאים מאפשרים לשילוב אפקטיבי	0.85	ידע, כלים, מיומנויות, זמן, תמיכה טכנית ופדגוגית מהמוסד האקדמי	3.09	1.02
תפיסת מסוגלות ועמדות כלליות	0.80	בטחון בשימוש, תרבות ארגונית תומכת, פתיחות לשילוב AI בהערכה	3.18	0.76
חששות אתיים	0.80	היבטי פרטיות, שקיפות, הוגנות ואחריות מקצועית	3.72	0.72

הממצאים בטבלה 1 מצביעים כי רמת התרומה שנתפשה על ידי המרצים נמוכה ($M=2.75$), לצד שונות רבה בין המשתתפים ($SD=1.17$). המרצים מעריכים כי קיימים תנאים בסיסיים סבירים לשילוב GenAI בתהליכים אלו, הכוללים ידע טכנולוגי ראשוני ויכולת בניית פרומפטים ($M=3.09$, $SD=1.02$). כמו כן נמצא כי תחושת המסוגלות לשילוב GenAI בהערכה נתפסת כבינונית ($M=3.18$, $SD=0.76$).

ממד החששות האתיים רשם את הערכים הגבוהים ביותר בכל המדדים שנבדקו ($M=3.72$), והציג אחידות גבוהה בתשובות ($SD=0.72$).



איור 2. מתאמי פירסון בין ממדי התפיסות השונים

ניתוח הקשרים בין המדדים (איור 2) מצביע על כך שקיים מתאם בינוני בין תפיסת התרומה לקיום תנאים מאפשרים ($r=0.41$, $p<0.001$) ומתאם גבוה בין תפיסת המסוגלות לתפיסת התרומה ($r=0.538$, $p<0.001$) ובין תפיסת המסוגלות

לקיום תנאים מאפשרים ($r=0.68$, $p<0.001$). לעומת זאת, לא נמצא מתאם מובהק בין החששות האתיים לשלושת המדדים האחרים ($r=0.11$, $p>0.05$).

הבדלים בתפיסות בין מרצים בעלי ניסיון בהערכת עבודות באמצעות AI לבין מרצים ללא ניסיון

ההשוואה בין מרצים בעלי ניסיון קודם בהערכת עבודות באמצעות GenAI לבין מרצים ללא ניסיון התבססה על ארבעת מדדי השאלון, ונבחנה באמצעות מבחני t (טבלה 2). נמצא כי מרצים בעלי ניסיון העריכו את תרומת ה-GenAI לתהליך ההערכה כגבוהה יותר ($M=3.11$, $SD=0.97$) בהשוואה למרצים ללא ניסיון ($M=2.12$, $SD=1.22$), הבדל מובהק סטטיסטית ובעל גודל אפקט גבוה ($t(110)=4.45$, $p<0.001$, $d=0.93$). עם זאת, יש לציין כי גם בקרב מרצים בעלי ניסיון, תפיסת התרומה נותרה בטווח הבינוני והלא גבוה.

בדומה לכך, תפיסת התנאים המאפשרים לשילוב GenAI בהערכה הייתה גבוהה יותר בקרב מרצים בעלי ניסיון ($M=3.47$, $SD=0.84$) לעומת מרצים ללא ניסיון ($M=2.45$, $SD=0.97$), עם הבדל מובהק וגודל אפקט גבוה במיוחד ($t(110)=5.81$, $p<0.001$, $d=1.14$). גם בממד המסוגלות והעמדות הכלליות נמצא יתרון מובהק למרצים בעלי ניסיון ($M=3.43$, $SD=0.69$) בהשוואה למרצים ללא ניסיון ($M=2.74$, $SD=0.68$), מלווה בגודל אפקט גבוה ($t(110)=5.15$, $p<0.001$, $d=1.01$). לעומת זאת, בממד החששות האתיים לא נמצא הבדל בין הקבוצות ($t(110)=-0.02$, $p>0.001$) ושתי הקבוצות דיווחו על רמות גבוהות ודומות של חששות אתיים.

טבלה 2. תפיסות כלפי שימוש בכלי GenAI להערכת עבודות סטודנטים. הבדלים בין בעלי ניסיון לבין חסרי ניסיון ($p<0.001$ ***)

מדד תפיסה	בעלי ניסיון M (n=71)	חסרי ניסיון M (n=41)	DF	t	d
תרומת ה-GenAI להערכה	3.11	2.12	110	***4.45	0.93
קיום תנאים מאפשרים	3.47	2.45	110	***5.81	1.14
תפיסות ועמדות כלפי אימוץ GenAI	3.43	2.74	110	***5.15	1.01
חששות אתיים	3.72	3.71	110	0.062	0.01

דיון ומסקנות

הפער שנמצא בין השימוש הנרחב ב-GenAI בהוראה לבין האימוץ המוגבל בהערכה משקף זהירות מוצדקת מצד המרצים. החלטות הערכה נושאות משקל גבוה ומשפיעות באופן ישיר על עתיד הסטודנטים, ולכן הן דורשות רמת אחריות גבוהה יותר מאשר משימות הוראה שגרתיות. זהירות זו אינה נובעת מהתנגדות לטכנולוגיה או מחוסר נוחות בשימוש בכלים, אלא מהבנה עמוקה של המשמעות המקצועית והאתית של תהליכי ההערכה (Dabis & Csáki, 2024; Holmes et al., 2022). ממצאי המחקר מאשרים את ההנחה המרכזית של תיאוריית הפצת החדשנות (Rogers, 2003), לפיה ניסיון מעשי מגביר את תפיסת התועלת ואת תחושת המסוגלות העצמית. מרצים שרכשו ניסיון בשימוש ב-GenAI להערכה מעריכים באופן משמעותי יותר את תרומתה הפוטנציאלית ומרגישים בטוחים יותר ביכולתם לשלב אותה בעבודתם. עם זאת, ממצא מרכזי מעלה תובנה חשובה: בעוד שהניסיון משפיע לטובה על תפיסות המסוגלות והתועלת, החששות האתיים נותרים גבוהים ויציבים בקרב שתי הקבוצות כאחד. תוצאה זו מצביעה על כך שהדאגות האתיות אינן נובעות מחוסר היכרות עם הטכנולוגיה או מחוסר ניסיון בשימוש בה, אלא מושרשות בהבנה מעמיקה של המורכבות המוסרית הטמונה בהערכה אקדמית. ככל שמרצים צוברים ניסיון, הם נחשפים בפועל למגבלות המודלים, לסיכון להטיות ולקושי להסביר ולהצדיק החלטות הערכה מול סטודנטים (Lim et al., 2023; Holmes et al., 2022). הדאגות האתיות, אם כן, קשורות לערכים המקצועיים ולזהות המקצועית של המרצים, ולכן אינן ניתנות לפתרון באמצעות רכישת מיומנות טכנית בלבד (Holmes et al., 2022). מהממצאים עולה כי טיפול אפקטיבי בחששות האתיים דורש יותר ממענה טכנולוגי או הכשרה. נדרשת מדיניות מוסדית ברורה ומפורטת המגדירה שימושים מותרים ואסורים, קובעת חובות שקיפות כלפי הסטודנטים, מבהירה דרכים לשמירה על פרטיות המידע האקדמי, וקובעת גבולות ברורים לאחריות המרצה (Council for Higher Education in Israel, 2023; Dabis & Csáki, 2024; Moorhouse et al., 2023).

תוכניות הכשרה למרצים צריכות לעבור מגישה טכנית לגישה יישומית-אתית. ההכשרה צריכה לכלול לא רק הכרת הכלים, אלא גם מיומנויות הערכת איכות התשובות שמתקבלות, שיטות תיעוד פרומפטים ותהליכי קבלת החלטות, ועקרונות בקרה ומשוב (Zawacki-Richter et al., 2019). בנוסף, יש להבטיח שהמרצה שומר על אחריות מלאה ובקרה אנושית לאורך כל תהליך ההערכה, תוך בחינה ביקורתית של תפוקות ה-GenAI הכוללת הפעלת שיקול דעת עצמאי ותיעוד שקוף של השיקולים והפרומפטים. צעדים אלה עשויים לחזק את אמון הסטודנטים בתהליך ההערכה, להפחית ערעורים ולבסס תרבות של שימוש אחראי ומבוקר ב-GenAI בהערכה אקדמית (Holmes et al., 2022; Moorhouse et al., 2023). לסיכום, המחקר מצביע על הצורך בגישה כוללת לשילוב GenAI בהערכה – גישה המכירה בפוטנציאל הטכנולוגי אך מציבה במרכז את הערכים האקדמיים, האחריות המקצועית ואת הסטודנט.

מקורות

- Al-Zahrani, A. M., & Alasmari, T. M. (2024). Exploring the impact of artificial intelligence on higher education: The dynamics of ethical, social, and educational implications. *Humanities and Social Sciences Communications*, 11(1), 1-12. <https://doi.org/10.1057/s41599-024-03432-4>
- Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 49(6), 893-905. <https://doi.org/10.1080/02602938.2024.2335321>
- Council for Higher Education in Israel. (2023). *Digital transformation in Israeli higher education: Strategic plan 2022-2025*. Council for Higher Education in Israel.
- Dabis, A., & Csáki, C. (2024). AI and ethics: Investigating the first policy responses of higher education institutions to the challenge of generative AI. *Humanities and Social Sciences Communications*, 11(1), 1-13. <https://doi.org/10.1057/s41599-024-03526-z>
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S.B., Santos, O.C., Rodrigo, M.T., Cukurova, M., Bittencourt, I.I. and Koedinger, K.R., 2022. Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, 32(3), pp.504-526. <https://doi.org/10.1007/s40593-021-00239-1>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Köhl, N., Lang, J., Maheshwari, S., Müller, J., Pfeffer, J., Reinecke, K., Röpke, J., Rußwurm, M., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Lim, T., Gottipati, S., & Cheong, M. L. F. (2023). Artificial intelligence in today's education landscape: Understanding and managing ethical issues for educational assessment [Research Square]. <https://doi.org/10.21203/rs.3.rs-2696273/v1>
- Merino-Campos, C. (2025). The impact of artificial intelligence on personalized learning in higher education: A systematic review. *Trends in Higher Education*, 4(2), 17. <https://doi.org/10.3390/higheredu4020017>
- Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, 5, 100151. <https://doi.org/10.1016/j.caeo.2023.100151>
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>