

אוניברסיטת בר אילן
המחלקה ללימודי מידע

הצעת מחקר לתיזה

תוכן חופשי באקדמיה הישראלית
בחינת ההשפעה של תוכן חופשי על האקדמיה
בישראל כתוצאה מהורדת חסמים הקשורים
בזכויות יוצרים.

Free Content in the Israeli Academia

Testing the effect of free content on the Academia in
Israel as a result of removing boundaries related to
Copyright.

ליאור קפלן

הצעה זו מוגשת כחלק מהדרישות לשם קבלת תואר מוסמך
במחלקה ללימודי מידע של אוניברסיטת בר-אילן

בהנחיית פרופ' יהודית בר אילן

תשע"א

רמת גן

תוכן עניינים

3.....	מבוא.....
7.....	רקע תיאורטי.....
10.....	שאלות המחקר.....
10.....	שיטה.....
10.....	אוכלוסיה ומדגם.....
10.....	הגדרות.....
10.....	כלי המחקר.....
11.....	הליך.....
11.....	מגבלות המחקר.....
12.....	רשימת מקורות.....

מבוא

ההגדרה של אקדמיה המופיעה בויקיפדיה (אקדמיה, 2009) הינה: "אקדמיה היא מוסד שבו פועלים אנשי מדע ורוח להרחבת הידע האנושי ולהפצתו.", אך פעילות האקדמיה למחקר ולהרחבת הידע האנושי מוגבלת כיום במידה רבה כתוצאה מנושאי זכויות יוצרים. תוכן חופשי הוא רעיון שפורח בעשור האחרון ועיקרו הוא שחרור ושיתוף של תוכן ומידע באמצעות הורדת חסמים הקשורים בזכויות יוצרים. לרעיון ישנם מימושים רבים במגוון תחומים, חלקם מתייחסים למידע וידע אקדמי.

בעוד המגמה קיימת בעיקר במוסדות אקדמיים בחוץ לארץ, ניתן לראות ניצנים ראשוניים לשינוי גם במוסדות אקדמיים בישראל שהתחילו לאמץ שימוש בתוכן חופשי. מטרת מחקר זה הינה להבין כיצד הורדת חסמים הנובעים מנושאי זכויות יוצרים באמצעות תוכן חופשי משפיעים כיום ועשויים להשפיע בעתיד על האקדמיה בישראל.

השם העברי של המונח זכויות יוצרים מתאר פרשנות מודרנית של המונח Copyright. כפי שמופיע בחוק המלכה אן (Statute of Anne) משנת 1709 באנגליה, זכויות היוצרים (זכויות העתקה בתרגום ישיר) הראשונות היו מנגנון לשליטה ביכולת העתקה על ידי הדפסה מחדש של ספרים עם הפיכת הדפוס למסחרי וללא המונופול בתחום שהיה קיים עד אז. יש לזכור כי הזכויות היו נתונות למדפיסים (שהיו לרוב גם המוציאים לאור) ולא ליוצרי היצירה כמו שנהוג כיום.

הזכויות המוקנות ליוצרים התפתחו במהלך השנים - נוספו להן "זכויות מוסריות" ובנוסף הוגדרו המגבלות המוטלות על אחרים כחלק מאותן זכויות של בעלי היצירה. מצד שני, התווסף המושג "טיפול הוגן" (אנגליה) ו-"שימוש הוגן" (ארה"ב) על מנת לאזן את האינטרס הציבורי¹. כמו כן, הוכנס לשימוש המושג נחלת הכלל (public domain) שהוצג בחוק המלכה אן, ומתאר את המצב עבור יצירות שזכויות היוצרים שלהן הסתיימו.

אם נחזור לשורשי האקדמיה, החל מתקופת הפילוסופיים היוונים ועד להמצאת הדפוס במאה ה-15, נראה כי מידע וידע היו נדירים וקשים להשגה. חיבור בין אקדמאים שונים היה קשה והדרך העיקרית לשימור והעברת מידע היתה על ידי העתקה ושיכפול יצירות שכללו כתבי קודש, חיבורים מדעיים וחיבורים פילוסופיים. העתקות אלה (שחלקן מהוות תרגום) אפשרו לידע היווני והרומי לשרוד את תקופות ימי הביניים של אירופה על ידי שימורו ופיתוחו על ידי האיטלקים. דוגמה לכך הוא ספרו של אוקלידס "יסודות" (Elements) שתורגם מיוונית לערבית, ולאחר כ-400 שנה תורגם ללטינית וכך הוצג מחדש באירופה לאחר שהידע באירופה אבד. הספר הודפס מחדש בונציה לאחר המצאת הדפוס וכתוצאה מכך זמין לנו עד היום כ-2300 שנה לאחר כתיבתו.

לפני המצאת הדפוס, העתקת ספרים בוצעה באופן ידני ושכפול כל ספר היתה עבודה מפרכת שנמשכה מספר שנים. קושי זה השאיר את תחום העתקת הספרים לאלה שהקדישו לו את חייהם (נזירים לדוגמה) או כאלה שהתפרנסו מכך לרוב על ידי שכפול ספרים עבור השליט או בעל ממון כלשהו. בעולם זה, קושי העתקת הספרים נתפס כמגבלה ולא כיתרון מאחר והוא הקשה גם על הפצת הרעיונות והידע של המחבר.

1 Fair Dealing (אנגליה) ו-Fair Use (ארה"ב) אינם זהים משפטית, אם כי הבסיס הרעיוני והמטרה זהים. בארה"ב ל-Fair Dealing יש משמעות שאינה קשורה לעולם זכויות היוצרים.

בעולם זה, ללא זכויות יוצרים, המחקר המדעי והפילוסופי התקדם ללא מגבלות הקשורות לפרסום, אך סבל רבות ממגבלות אחרות כגון תקשורת וקשיי שפה. ההתקדמות הושגה באמצעות השיטה שקיימת עד היום והיא התבססות על מחקרים של אחרים כדי להתקדם קדימה. כפי שהגדיר זאת אייזק ניוטון ב-1676 "אם הרחקתי לראות, הרי זה משום שעמדתי על כתפיהם של ענקים." (מתוך מכתב של ניוטון לרוברט הוק, פברואר 1676 כפי שמופיע אצל Maury, 1992, תרגום ע"י ויקיפדיה).

אך כיום המצב של האקדמיה שונה באופן מהותי, כאשר היכולת לעמוד על כתפיהם של ענקים מוגבלת כתוצאה מזכויות היוצרים. נחלת הכלל המהווה את הבסיס האיתן של הידע והתרבות האנושיים כוללת כמות ספרים משמעותית. השגת עותקים ומתן גישה לספרים אלה (לצד ספרים שעדיין מוגנים בזכויות יוצרים) דורשת מכל מוסד אקדמי מאמצים ומשאבים כספיים רבים. בעולם הדפוס נכנס גם מימד השימור בהתאם לכמות השימוש, גיל הספר, האם הוא מודפס באופן פעיל ולנדירותו (כתבי יד לדוגמה). פעולות אלה דורשות מכל מוסד אקדמי משאבים רבים כדי להחזיק ספריה מכובדת, משאבים שלרוב הגופים קשה להשקיע. מצד שני, השימוש ההוגן, למרות שמוגדר בחוק זכויות היוצרים, הינו תחום אפור ללא קווים ברורים. חריגה ממגבלות אלה עשויות לגרור עונשים משמעותיים, בעיקר בעידן בו הקניין הרוחני נתפס לעיתים כחשוב או משמעותי יותר ביחס לזכויות אחרות (לרבות קניין שאינו רוחני), וזאת על פי גובה הקנסות הקשורים בהפרת זכויות יוצרים. בישראל ומדינות רבות, הפרת זכויות יוצרים למטרות רווח הינה עבירה פלילית.

על הקושי שבניצול של זכויות השימוש ההוגן יעיד פרויקט "השכלה נגישה: שימוש הוגן בהשכלה הגבוהה" שנוצר על ידי חוקרים ממוסדות האקדמיה בישראל במטרה לנסח כללי שימוש בזכויות במטרה להתגבר על האפקט המצנן של סכנת ההפרה של זכויות היוצרים. מתוך הצהרת המשימה של הפרוייקט (השכלה נגישה. (ח"ת):

דיני זכויות יוצרים מהווים בשנים האחרונות אחד החסמים המרכזיים להנגשת חומרי לימוד (כגון: מחקרים, מאמרים, הרצאות ומערכי שיעור, סיכומים, טבלאות, סרטים, צילומים ונתונים אמפיריים) לצורכי הוראה ומחקר במוסדות להשכלה גבוהה. הדבר נובע, בין היתר, מחוסר הוודאות המשפטי בנוגע לגבולות השימושים המותרים בחומרים אשר מוגנים בזכויות יוצרים.

[...]

האקדמיה מהווה קהילה של יוצרי תוכן ומשתמשי תוכן לצורכי הוראה, חינוך ודעת. למוסדות להשכלה גבוהה תפקיד חשוב ביצירתו ובהנחלתו של ידע אקדמי, לציבור החוקרים, לסטודנטים ולציבור הרחב.

[...]

בנוסף לקושי של השימוש ההוגן, ישנו קושי גדול הרבה יותר והוא שימוש מלא² ביצירות. ברירת המחדל של רישיון השימוש בזכויות יוצרים הינו "כל הזכויות שמורות" דבר המונע מכל אדם זולת בעל הזכויות להשתמש ביצירה. כדי להמנע מהפרת זכויות יוצרים, נדרש תיאום מראש עבור כל שימוש ביצירה מוגנת, ולעיתים אף קניית רישיון שימוש. במקרים שונים קיימים גופים שנועדו

2 כביכול שימוש "לא הוגן", אך המונח שימוש ביצירה המלאה יתאים יותר.

להסדיר את הנושא (אקו"ם בעולם המוזיקה), אך במקרים אחרים ישנו צורך ליצירת קשר והתעסקות עם גופים רבים, דבר המהווה תקורה משמעותית.

כתוצאה מאופי תחום זכויות היוצרים, ומברירת המחדל הקיימת בחוק, עיקר ההתעסקות בזכויות היוצרים נוגעת לאיסורים הקשורים לשימוש ביצירה או בפיצוי (לרוב כספי) במטרה לאפשר שימוש כזה או אחר. גם כאשר הצד המסחרי/כלכלי אינו רלוונטי, רישיון השימוש נשאר "כל הזכויות שמורות" והדבר מדלל באופן משמעותי את הפצת המידע או את היכולת להתבסס עליו לשם התקדמות אקדמית. מסיבות אלה, נושא הבעלות על זכויות היוצרים הינו מהותי בעבודה משותפת או מרובת מחברים³. כתוצאה מכך נפגעים שיתופי פעולה, כאשר תרומה לעבודה בה אין לתורם זכויות יוצרים עליה עלולה להוביל לחוסר גישה ליצירה בעתיד.

לזכויות יוצרים ישנם היבטים נוספים באקדמיה, בשלב זה אציין חלק מהם בקצרה:

- העברת זכויות היוצרים לכתבי עת בתמורה לפרסום, דבר שיוצר מגוון בעיות ובראשן הצורך של מוסדות אקדמיים לרכוש גישה למאמרים אלה.
- זכויות היוצרים של מרצים ביצירות שלהם, מאחר ובאופן עקרוני זכויות היוצרים שייכות למקום העבודה ולא לעובד (חוק זכויות יוצרים התשס"ח-2007, סעיף 34). ישנם מוסדות אקדמיים שמתייחסים לעובדים כבעלים של זכויות יוצרים, אך הדבר תלוי במדיניות המוסד.
- זכויות היוצרים שבין המרצים לבין הסטודנטים שמעבדים באופן קבוע את חומרי המרצה (סיכומים, הקלטות וכו') וממזגים מקורות נוספים לתוכן (היכן מדובר בהעתקה, ביצירה נגזרת ואף יצירה עצמאית לחלוטין).
- זכויות היוצרים של סטודנטים בעבודות שלהם (Chan, 2006) ביחס למוסד האקדמי בו הם לומדים (בהנחה שהסטודנט לא יוצר עבודה בחלל ריק) וביחס למרצים שלהם שעשויים לתרום לעבודות.
- זכויות היוצרים של סטודנטים בעבודות שלהם המשמשות כבסיס עבור מחקרים ומאמרים של המרצים.
- זכויות היוצרים של סטודנטים ביחס לעמיתיהם ללימודים ולמחקר ביחס לתרומה שלהם לעבודה של עמיתים.
- זכויות היוצרים של האוניברסיטה כמוסד הממומן ברובו על ידי המדינה באמצעות משלם המיסים ביחס לאזרחי המדינה.

בעיה נוספת וכללית יותר בנושא זכויות היוצרים הינה המעבר לעולם דיגיטלי. בעוד המגמה הכללית של המעבר לעולם דיגיטלי היא חיובית לרוב (לדוגמה גישה לספרים וכתבי עת מרחוק), מבחינת זכויות יוצרים מדובר בנסיגה משמעותית. בעידן הדיגיטלי, כל שימוש שהוא ביצירה מבוסס על היכולת להעתיק אותה (בין אם באופן פעיל על ידי המשתמש או באופן אוטומטי על ידי המערכת). זכות זאת היא הזכות הבסיסית שעליה מנסים זכויות היוצרים לשלוט ולהגביל, כך שהנושא נהפך לסבוך יותר.

רעיון התוכן החופשי לקוח מעולם התוכנה החופשית שהתחילה בשנות ה-80. רעיונות החופש הקשורים לתוכנת מחשב התפשטו לטקסטים אחרים ובהמשך גם לצורות תוכן אחרות כגון

3 ישנה הבחנה בין תורם לעבודה אשר אחראי על חלק מוגדר (פרק לדוגמה), לבין מחבר שותף אשר לא ניתן להפריד את תרומתו מזאת של שאר התורמים. לכל מחבר שותף יש זכויות יוצרים על העבודה באופן נפרד.

תמונות, קול ווידאו. לא מדובר באנרכיה או בויתור מוחלט על זכויות היוצרים, אלא במתן רישיון שימוש מתירני הדואג לחופש של המשתמש בתנאי שהקרדיט על היצירה נשמר. הזכויות הבסיסיות מוגדרות כזכות לשימוש ביצירה, הזכות ליצור עותקים ולהפיץ אותם, והזכות ליצור יצירות נגזרות (ולהפיץ אף אותן).

ישנם גם יישומים חלקיים יותר של הרעיון, כגון תוכן נגיש, אליהם אתייחס במהלך העבודה. תוכן חופשי, וגם מספר יישומים חלקיים שלו מוכרים על ידי הסיסמה "חלק מהזכויות שמורות" כמשחק מילים על הביטוי "כל הזכויות שמורות". במילותיו של לורנס לסיג (Lessig, 2004, p. 283):

What's needed is a way to say something in the middle – neither “all rights reserved” nor “no rights reserved” but “some rights reserved” – and thus a way to respect copyrights but enable creators to free content as they see fit. In other words, we need a way to restore a set of freedoms that we could just take for granted before.

השימוש בתוכן חופשי מאפשר להתגבר על חסמים שונים שנובעים מההגבלות של זכויות יוצרים ורישיון "כל הזכויות שמורות". מטרת המחקר המוצע הינה להבין את החסמים הקיימים כיום באקדמיה בישראל וכיצד שימוש בתוכן חופשי עשוי לפתור או לעקוף אותם. תרומת המחקר הצפויה הינה מצד אחד הבנה של בעיות זכויות יוצרים וכיצד הן משפיעות על פעילות אקדמית ומצד שני איזה כלים נמצאים בשימוש כדי לעקוף את הבעיות ומה הסיבות לשימוש באותם כלים.

רקע תיאורטי

במסגרת אקדמית בעולם, תוכן חופשי ותוכן נגיש⁴ מופיעים בשני תחומים עיקריים:

• כתבי עת, מאגרי הפקדה וארכיונים, רובם תחת הכותרת Open Access⁵.

• חומרי לימוד (Open Educational Resource) וקורסים מקוונים (OpenCourseWare).

שני התחומים קשורים באופן הדוק לעולם האינטרנט כאמצעי לחיפוש ומציאת מידע, כמו גם להתפתחות טכנולוגית. ברשימה המתעדת אירועים הקשורים ל-Open Access Movement ניתן לראות את הניצנים של תחום עוד לפני עולם האינטרנט כפי שהוא מוכר לנו כיום (Suber, 2009). בתחום זה מוגדרות כיום שתי דרכים עיקריות להשגה של גישה פתוחה: דרך הזהב (Gold Road) על ידי פרסום בכתב עת המקדם גישה פתוחה או הדרך הירוקה (Green Road)⁶ על ידי העברת המאמר לארכיון גישה פתוחה (Harnad et al., 2004). הגישה הפתוחה מגדילה באופן ודאי את מספר הקוראים הפוטנציאלי. מספר מחקרים הראו כי לגישה הפתוחה ישנה השפעה משמעותית גם על כמות ההפניות והציטוטים של המאמרים (Lawrence, 2001; Harnad & Brody, 2004; Antelman, 2004), אם כי יש החולקים על מאמרים אלה (Craig, Plume, McVeigh, Pringle & Amin⁷, 2007).

התחום השני התפתח מאמצע שנות ה-90 ואילך עם הפיכת הגישה לאינטרנט לנפוצה והתחלת השימוש בו לנושא הגישה ולימוד מרחוק. בשנת 2001 הכריזה MIT על פרוייקט ה-OpenCourseWare שמטרתו לשחרר במהלך של 10 שנים כמעט את כל ה-2,000 הקורסים שלה לגישה מהאינטרנט (Massachusetts Institute of Technology, 2001; Goldberg, 2001) ההכרזה היוותה מהפך בתחום וזרז משמעותי לגופים אקדמיים נוספים לפעול בכיוון של גישה לתכנים אקדמיים, חלקם ברישיון חופשי. היוזמה של MIT הוצגה בשנת 2002 בכנס של UNESCO תחת הכותרת "Forum on the Impact of Open Courseware for Higher Education in Developing Countries". בכנס זה גם הוחלט להשתמש במונח Open Educational Resource כמונח המייצג עבור השילוב בין גישה חופשית לעולם החינוך שהתעסק ב-Learning Objects ובלימוד מרחוק (UNESCO, 2002). בשנת 2005 נפתח ב-MIT ה-OpenCourseWare Consortium במטרה לאגד פעילות הקשורה ל-OCW. בקונסורציום חברים מעל ל-200 מוסדות השכלה גבוהה ביותר מ-40 מדינות (ביניהם גם האוניברסיטה הפתוחה).

תנועת התוכנה החופשית (Free Software) התחילה ב-1983 עם הקמת פרוייקט GNU על ידי ריצ'רד סטולמן (Stallman). כ-15 שנה לאחר מכן, בשנת 1998 נוצר המושג Open Content⁸ על ידי David Wiley, כאשר מדובר במימוש עקרונות החופש המוצגים של תוכנה חופשית עבור תכנים שאינם קוד תוכנה (Wiley, 2006). בשנת 1999 הוא הציג לראשונה רישיון שימוש עבור תוכן כזה (Open Publication License). רישיונות נוספים יצאו בשנת 2000 (GNU Free Documentation License) ו-2002 (Creative Commons). כיום, משפחת רישיונות ה-Creative Commons מהווים את

4 על ההבדל בין תוכן חופשי לתוכן נגיש ראו סעיף הגדרות בעמוד 10.

5 לרוב במשמעות של גישה ללא עלות (גישה חינוכית), ללא יתרונות של תוכן נגיש או חופשי. ראו גם MacCallum, 2007.

6 המושג שואב את שמו מכך שהמוציא לאור מאשר (נותן אור ירוק בסלנג) לשליחת המאמר לארכיון בגישה פתוחה.

7 חוקרים אלה עובדים עבור חברות הוצאה לאור המתפרנסות בין השאר ממכירת גישה למאגרי מאמרים.

8 בשנת 1998 נוצר המונח Open Source כתחליף למונח Free Software וזאת לאור מגבלה של השפה האנגלית בה המילה Free קיימים שני פירושים: חופשי (Libre) וחינם (Gratis). המונח Open Content מתבסס על חידוש זה. בהמשך המילה Open תובן במשמעות של נגיש/זמין ויבוצע מעבר למונח Free Content למרות הבלבול האפשרי באנגלית.

הכלי העיקרי עבור רישוי תוכן חופשי או נגיש, וזאת כתוצאה ממגוון האפשרויות בתוך משפחת הרישיונות ומהבהירות שלהם לגבי המותר והאסור בשפה אנושית בנוסף לשפה משפטית.

את החיבור בין תוכן חופשי לגישה חופשית ניתן למצוא בשתי הצהרות משנת 2003 והן הצהרת בטסדה לפרסום בגישה פתוחה (*Bethesda Statement on Open Access Publishing, 2003*) והצהרת ברלין לגישה פתוחה לידע במדעים ומדעי הרוח (*Berlin Declaration on Open Access to Knowledge*) (*in the Sciences and Humanities, 2003*), על האחרונה חתומים כ-250 מוסדות אקדמיים באירופה). בשתי ההצהרות ישנן הגדרות לגישה פתוחה הדומות לאלה של התוכן החופשי. כפי שמציין Fitzgerald (2007), רישיון תוכן חופשי הוא הבסיס עליו נוצרים מאגרי ה-OER והוא זה שמאפשר את השיתוף והשימוש החוזר של המשאבים במאגר. ללא הבהרה ברורה של נושא זכויות היוצרים במאגר, יתכנו הפרות עתידיות של זכויות יוצרים, וזאת בנוסף לבעיות של זכויות יוצרים במדיה דיגיטלית (Fisher & McGeveran, 2006).

יש לזכור כי למרות שהחוק מגדיר מחקר והוראה או בחינה על ידי מוסד חינוך כשימוש הוגן (חוק זכויות יוצרים התשס"ח-2007, סעיף 19(א)) הוא אינו מגדיר מהם הקווים הברורים לשימוש ההוגן, אלא רק מבחנים לבדיקת שימוש כזה. דרך הגדרה זאת, שדומה להגדרה בחוק האמריקאי, הופכת את השימוש ההוגן לאזור אפור בו לא ניתן לדעת בוודאות לגבי המותר או האסור. פרוייקט השכלה נגישה מנסה לפתור בעיה זאת (Dotan, Elkin-Koren, Fischman-Afori & Haramati-Alpern, 2010), אך באמצעות רישיונות תוכן חופשי אפשר לעקוף את הבעיה כמו גם לאפשר מגוון רחב יותר של שימושים.

הדגש הרב שניתן לזכויות היוצרים מעלה מספר שאלות כגון מי הבעלים של זכויות היוצרים ביצירות אקדמיות (חוק זכויות יוצרים התשס"ח-2007, סעיף 34; McCracken, 2003; McSherry, 2007), כמו גם היכולת של מוסדות אקדמיים לפרסם מדיניות מחייבת בנושא פרסום בגישה פתוחה (או כתוכן חופשי). במספר מוסדות בחו"ל (Harvard, Stanford ו-MIT לדוגמה) התקבלו החלטות על ידי המרצים או על ידי המוסדות לעודד (ואף לחייב) פרסום בגישה פתוחה.

בישראל ישנם כיום שני מיזמים מובילים בנושאים של OER ו-OCW והם פרוייקט פתיחת אוצרות רוח (פא"ר) של האוניברסיטה הפתוחה ומאגר אובייקטי למידה ברשת (מאו"ר). פא"ר מכיל בחלקו תוכן נגיש (רישיון CC-BY-NC-ND), אך נדרשת הרשמה (ללא עלות). הרישיון הנגיש מוגבל ל-14 ספרים (פא"ר, ח"ת; Tal-Elhasid, 2010) מתוך 129 ספרים שזמינים באתר הפרוייקט⁹. חומרים נוספים באתר זמינים עם "כל הזכויות שמורות" או ללא אזכור של רישיון מפורש. מאו"ר מכיל כ-3,200 אובייקטים בעברית, מתוכם כ-460 אובייקטים עם רישיון CC¹⁰ (ממספר בדיקות נראה שהרישיון המועדף הוא CC-BY-NC-ND). בנוסף, הוא משמש גם כשער ל-MERLOT (מאגר בינלאומי לאובייקטי למידה).

פרוייקטים שיוצרים תוכן חופשי, משמשים לעתים כחומר גלם למחקר אקדמי. בתחומים הקשורים לבלשנות וניתוח שפות מדובר בבניה וניתוח של קורפוסים. ההתבססות על טקסט בעל זכויות יוצרים מטילה מגבלות על החוקר ודורשות ממנו לתאם מראש פעולות הקשורות בטקסט שאינן בתחום השימוש ההוגן (גבאי, 2008, עמוד 6):

9 על פי המופיע בעמוד הסטטיסטיקות של פא"ר, נבדק ב-3 באפריל 2011.
10 נבדק ב-3 באפריל 2011 ע"י סימון כן באפשרות "קריאטיב קומונס" בעמוד החיפוש של המאגר.

הקורפוס העיקרי, שייקרא להן "קורפוס הבלוגים", הורכב מטקסטים שנלקחו מתוך בלוגים (יומני רשת) של תשעה כותבים... שנתנו את הסכמתם לשימוש שנעשה בטקסטים שלהם ולפרסומם (כדי שיתאפשר להעמיד את הקורפוס לשימוש חוקרים נוספים).

גבאי יצר קורפוס נוסף המתבסס על פרוייקט בן יהודה - פרוייקט שמטרתו לשמר יצירה עברית באופן אלקטרוני ודואג להקלדה של טקסטים בעברית שזכויות היוצרים בהם פגו ונכנסו לנחלת הכלל. הטקסטים שמקליד הפרוייקט מפורסמים גם הם לנחלת הכלל. כתוצאה מהיותם של הטקסטים זמינים בנחלת הכלל, ניתן לעשות בהם כל שימוש ללא אף מגבלה ולכן במקרה זה לא נדרש גבאי לאישורים בדומה לקורפוס הבלוגים שלו. פרוייקט בן יהודה מהווה בסיס למחקרים נוספים (Eliam, 2008) ואף מהווה חלק מקורסים בתחום הלשון העברית באוניברסיטת תל אביב, כאשר פרופ' תמר סוברן רושמת את ההכרות עם הפרוייקט כמטלה בשניים מהקורסים שהיא לימדה¹¹ (המגמה ללשון עברית, ח"ת). גם במחלקה ללשון עברית בבר אילן מתייחס לפרוייקט בן יהודה כמשאב טוב להפנות אליך את הסטודנטים והוא מופיע ברשימת הקורפוסים (המחלקה ללשון העברית וללשונות שמיות, 2011).

המחלקה ללשון עברית בבר אילן מפנה ברשימת הקורפוסים שלה גם לקורפוסים של מילה (מרכז ידע לתקשוב בשפה העברית בטכניון). מילה דואגת ליצירת קורפוסים בעברית ממקורות שונים, ומפורסמים ברישיון תוכן חופשי בשם ה-GFDL¹² (מילה, ח"ת). בעוד רוב הקורפוסים מגיעים ממקורות המחזיקים את זכויות היוצרים (כך שנדרש תיאום מראש או חשיפה לבעיות הקשורות לזכויות יוצרים בעתיד¹³), אחד מהם הוא קורפוס המבוסס על אתר ויקיפדיה העברית. במקרה של ויקיפדיה, יש לציין כי שימוש בתוכן האתר לצורך בניית הקורפוס מותר באופן מפורש כתוצאה מהרישיון החופשי של ויקיפדיה.

הקורפוסים של מילה נמצאים בשימוש לא רק בישראל, אלא גם בחוץ לארץ. פרוייקט hebrewCorpus של National Middle East Language Resource Center מכיל גישה לקורפוסים של מילה. ובנוסף מספק גישה לקורפוס המבוסס על פרוייקט בן יהודה ולקורפוסים המבוססים על פרוייקטי קרן ויקימדיה בעברית: ויקיספר, ויקיחדשות, ויקיפדיה¹⁴, ויקיצוטוט וויקיטקסט (hebrew Corpus, n.d.). קורפוס נוסף המבוסס על ויקיפדיה נוצר באוניברסיטת בן גוריון (Gabay, Ben-Eliahu & Elhadad, 2008), גם הוא זמין על פי רישיון ה-GFDL¹⁵ (בהתאם לרישיון של התוכן המקורי מויקיפדיה).

11 בקורסים לשונה של השירה העברית החדשה, לשונה של הסיפורת העברית החדשה.

12 רישיון תוכן חופשי זה נפוץ בעולם הקוד הפתוח, ומשם הגיע לעולם התוכן החופשי. ויקיפדיה השתמשה ברישיון זה כרישיון העיקרי שלה עד אמצע 2009.

13 מצד אחד מדובר בהעתקת תוכן רב, מצד שני מדובר במאגר לניתוח לשוני בו הטקסט המוגן הוא רק חומר גלם ואינו מופיע בצורתו המקורית (או כזאת שיכולה להחליף את התוכן המקורי).

14 נראה שמדובר בקורפוס שונה מזה של מילה.

15 <http://www.cs.bgu.ac.il/~nlpproj/wiki-seg-corpus/>

שאלות המחקר

מהם החסמים הקיימים כיום באקדמיה בישראל הקשורים לזכויות יוצרים?
כיצד משפיע השימוש בתוכן חופשי על מחקר ופעילות אקדמית?

שיטה

אוכלוסיה ומדגם

אוכלוסיית המחקר הינה חוקרים במוסדות אקדמיים, לרבות סטודנטים המבצעים עבודות חקר בשלבים שונים של לימודיהם (כולל עבודות סמינריוניות ותיזות), אשר יוצרים ומפרסמים מידע כחלק מהפעילות שלהם.
המידע שמפורסם אינו מוגבל רק למאמרים אלא גם לכלי מחקר שונים כגון נתונים, קורפוסים, ספרים דיגיטליים או משאבים המפותחים במשותף עם סטודנטים כחלק מתהליך הלימוד.
המדגם מהאוכלוסיה נבחר באמצעות עבודת מיפוי שאני מבצע כדי לגלות שימושים בתוכן חופשי באקדמיה. במהלך המיפוי נעזרתי ברשימת ערכים בוויקיפדיה שנכתבים במסגרת אקדמית (ויקיפדיה:כתיבה בוויקיפדיה כמטלה אקדמית, 2011), חיפוש עצמאי באינטרנט, בקשת הפניות מגופים שונים (פרוייקט בן יהודה לדוגמה) והפניות שקיבלתי מאנשים שנחשפו להצעת מחקר זאת (דומה לשיטת כדור השלג, אך ההפניות אינן מגיעות מהנדגמים עצמם).
המדגם צפוי להכיל כ-20 נבדקים על פי המתואר לעיל.

הגדרות

ההגדרות הבאות מתבססות על רישיונות Creative Commons לשם בהירות, ניתן למצוא הגדרות מקבילות באמצעות רישיונות אחרים. בהגדרות הבאות נושא הייחוס (מתן הקרדיט) אינו נתפס כמגבלה.

תוכן חופשי - תוכן שאינו מוגן בזכויות יוצרים (נחלת הכלל / Public Domain) או כזה המאפשר שימוש לכל מטרה באמצעות הרישיונות CC-BY ו-CC-BY-SA.

תוכן נגיש - תוכן שניתן לגישה ולשימוש בחינם ומאפשר שימושים לא מסחריים לרבות יצירות נגזרות באמצעות הרישיונות CC-BY-NC ו-CC-BY-NC-SA.

תוכן חינמי - תוכן שניתן לגישה ולשימוש ללא תשלום, אך מכיל מגבלות על יצירות נגזרות באמצעות הרישיונות CC-BY-ND ו-CC-BY-NC-ND.

כלי המחקר

המחקר יבוצע באמצעות ראיונות מובנים למחצה עם הנבדקים. השאלות כוללות שיחה כללית על זכויות יוצרים באקדמיה (דיעה אישית, קשיים שחווית בעבודתך), פירוט שימושים שנעשו בתוכן חופשי או נגיש במהלך העבודה האקדמית ומה היו השיקולים והיתרונות לשימושים אלה.
במהלך השיחה צפויים לעלות החסמים השונים הקיימים בנושא זכויות יוצרים, כמו גם פירוט כיצד הם נפתרו (או שעדיין מהווים בעיה ללא פתרון).

הליך

ראיון יבוצע עם חברים במדגם לאחר תיאום מראש, הבהרת מטרות המחקר והחתמת טפסי הסכמה מדעת. אורך הראיון לא יעלה על 90 דקות. הראיון יבוצע בעל פה והקלטה שלו (אם תאושר ע"י המרואיין) תתומלל לאחר הראיון. במקרה ולא ניתן לבצע ראיון בעל פה יבוצע ראיון באמצעות הצגת שאלות דרך מדיה דגיטלית (צ'אט, דוא"ל).

ניתוח התוכן של הראיונות יבוצע ע"י קידוד טקסט הראיונות שבעזרתו אפשר יהיה להוציא את התמות המרכזיות. התמות יאפשרו הבנה טובה יותר של הנושא המחקר (שקדי, 2003).

מגבלות המחקר

המחקר מתמקד בתכנים כחלק ממחקר ופרסום אקדמי, כמו גם לתכנים לצורך לימוד במוסדות אקדמיים. המחקר אינו מתעסק בנושאי זכויות יוצרים של עובדי אוניברסיטה שלא לצורך פעילות אקדמית, או בתחומים אחרים הקשורים לקניין רוחני כגון פנטנטים.

מאחר והאוכלוסיה אינה ידועה באופן מלא, המדגם אינו אקראי ואינו בהכרח מהווה מדגם מייצג של האוכלוסיה. עם זאת, בהתאם לגודל האוכלוסיה שתתגלה, המדגם עשוי לכלול את רובה.

רשימת מקורות

- אייזק ניוטון. (2011, פברואר 20). ויקיפדיה, האנציקלופדיה החופשית. אקדמיה. (2009, אוקטובר 3). ויקיפדיה, האנציקלופדיה החופשית. גבאי, ד' (2008). זיהוי מחבר על פי סגנון בטקסטים בעברית החדשה. חיבור לשם קבלת תואר "מגיסטר במדעי הטבע", אוניברסיטת בן גוריון בנגב. אוחר מהכתובת <http://www.cs.bgu.ac.il/~nlpproj/dudi/Authorship%20Attribution%20in%20Hebrew-1.pdf>
- המגמה ללשון עברית. (ח"ת). חומרי עזר לקורסים. אוחר מהכתובת <http://www.tau.ac.il/humanities/hebrew/useful-courses-s.heb.html>
- המחלקה ללשון העברית וללשונות שמיות. (2011, אפריל 4). משאבים. אוחר מהכתובת <http://www.biu.ac.il/js/hb/resources.html>
- השכלה נגישה. (ח"ת). אודות הפרוייקט. אוחר מהכתובת <http://openeducation.org.il>
- ויקיפדיה:כתיבה בוויקיפדיה כמטלה אקדמית. (2011, פברואר 18). ויקיפדיה, האנציקלופדיה החופשית.
- חוק זכויות יוצרים, התשס"ח-2007. (2007). בתוך רשומות - ספר החוקים 2119. אוחר מהכתובת <http://www.justice.gov.il/NR/rdonlyres/32331F27-82B1-4191-AE64-7F3A1024892B/9216/2119.pdf>
- יסודות (ספר). (2011, פברואר 18). ויקיפדיה, האנציקלופדיה החופשית. פא"ר. (ח"ת). זכויות יוצרים. אוחר 3 באפריל 2011 מהכתובת <http://ocw.openu.ac.il/files/copyright.html>
- מילה. (ח"ת). קורפוסים לעברית. אוחר מהכתובת <http://www.mila.cs.technion.ac.il/hebrew/resources/corpora>
- שקדי, א' (2003). מילים המנסות לגעת: תיאוריה ויישום. תל אביב: רמות אוניברסיטת תל אביב.
- Antelman, K. (2004). Do Open Access Articles Have a Greater Research Impact?. *College & Research Libraries News*, 65(5). pp 372-382. Retrieved from <http://crl.acrl.org/content/65/5/372>
- Bethesda Statement on Open Access Publishing. (2003, June 20). Retrieved from <http://www.earlham.edu/~peters/fof/bethesda.htm>
- Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. (2003, October 22). Retrived from <http://oa.mpg.de/lang/en-uk/berlin-prozess/berliner-erklarung/>
- Chan, Phil C. W. (2006). Copyright ownership in university students' academic works. *Journal of Intellectual Property Law & Practice*, 1(10). pp 664-674. doi: 10.1093/jiplp/jpl120
- Craig, I. D., Plume, A. M. McVeigh, M. E., Pringle, J. & Amin, M. (2007). Do Open Access Articles Have Greater Citation Impact?: A critical review of the literature. *Journal of Informetrics*, 1(3), 239-248. Retrived from <http://www.publishingresearch.net/documents/Citations-finalversion.pdf>

- Dotan, A., Elkin-Koren, N., Fischman-Afori, O., Haramati-Alpern, R. (2010). Fair Use Best Practices for Higher Education Institutions: The Israeli Experience . Journal of the Copyright Society of the U.S.A., 57(3). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1648408
- Eilam, A. (2008). Beyond Regularization: The Role of Native Learners in Shaping Modern Hebrew Relative Clauses. Paper presented at Generative Approaches to Language Acquisition in North America (GALANA) 3, UConn, Sept. 4-6, 2008.
- Fisher, W. & McGeeveran, W. (2006). "The Digital Learning Challenge: Obstacles to Educational Uses of Copyrighted Material in the Digital Age. Berkman Center Research Publication No. 2006-09. Retrieved from http://cyber.law.harvard.edu/publications/2006/The_Digital_Learning_Challenge
- Fitzgerald, B. (2007). Open Content Licensing (OCL) for Open Educational Resources . Organisation for Economic Co-operation and development. Centre for Educational Research and Innovation (CERI). Retrieved from <http://www.oecd.org/dataoecd/33/10/38645489.pdf>
- Gabay, D., Ben-Eliahu, Z. & Elhadad, M. (2008). Using Wikipedia Links to Construct Word Segmentation Corpora. Proceedings of the AAAI Workshop on Wikipedia and Artificial Intelligence: An Evolving Synergy (WikiAI). Association for the Advancement of Artificial Intelligence. California: AAAI Press. Retrieved from <http://www.cs.technion.ac.il/~gabr/publications/papers/Bunescu2008WikiAI.pdf>
- Goldberg, C. (2001, April 4). Auditing Classes at M.I.T., on the Web and Free. The New York Times. Retrieved from <http://www.nytimes.com/2001/04/04/us/auditing-classes-at-mit-on-the-web-and-free.html>
- Harnad, S. & Brody, T. (2004). Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals. *D-Lib Magazine*, 10 (6). ISSN 1082-9873. Retrieved from <http://eprints.ecs.soton.ac.uk/10207/1/06harnad.html>
- Harnad, S., Brody, T., Vallieres, F., Carr, L., Hitchcock, S., Gingras, Y, Oppenheim, C., Stamerjohanns, H., & Hilf, E. (2004). The Access/Impact Problem and the Green and Gold Roads to Open Access. *Serials Review*, 30(4). pp 310-314, ISSN 0098-7913, doi:10.1016/j.serrev.2004.09.013
- hebrewCorpus. (n.d.). Instructions. National Middle East Language Resource Center (NMELRC). <http://hebrewcorpus.nmelrc.org/search.php>
- Lawrence, S. (2001). Free online availability substantially increases a paper's impact. *Nature*, 411 (6837). p 521. doi:10.1038/35079151
- Lessig L. (2004). Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity . New York: The Penguin
- MacCallum, CJ. (2007). When is open access not open access? PLoS Biology 5(10): e285. doi:10.1371/journal.pbio.0050285
- Maury, Jean-Pierre. (1992). Newton: Understanding the cosmos. London: Thames and Hudson
- Massachusetts Institute of Technology. (2001, April 4). MIT to make nearly all course materials available free on the World Wide Web. Retrieved from: <http://web.mit.edu/newsoffice/2001/ocw.html>

- McCracken, R. (2006). Cultural responses to open licences and the accessibility and usability of open educational resources . Paper for Expert Meeting on Open Educational Resources. Retrieved from <http://www.oecd.org/dataoecd/48/38/36539322.pdf>
- McSherry, C. (2003). Who Owns Academic Work? Battling for Control of Intellectual Property. London: Harvard University Press. ISBN 0-674-01243-7
- Suber, P. (2009). Timeline of the Open Access Movement. Retrieved from <http://www.earlham.edu/~peters/fos/timeline.htm>
- Tal-Elhasid E. (2010). Free Access to E-Books and 'OER's at the Open University of Israel. Paper presented at OCWC Global 2010, Educational Policy and OpenCourseWare, Hanoi, Vietnam. Retrieved from http://www.openu.ac.il/Personal_sites/download/OUIsrael-ocwc2010.pdf
- UNESCO. (2002). Forum on the Impact of Open Courseware for Higher Education in Developing Countries: Final report. Retrieved from <http://unesdoc.unesco.org/images/0012/001285/128515e.pdf>
- Wiley, D. (2006, February 6). Paper for Expert Meeting on Open Educational Resources. Organisation for Economic Co-operation and development. Centre for Educational Research and Innovation. Retrieved from <http://www.oecd.org/dataoecd/19/26/36224377.pdf>