

## רלוונטיות, איכות ומרכזיות: המשמעות של אחזור על פי רלוונטיות ועל פי מדדים ביבליומטריים בכלי גילוי

אילת נבון, התכנית למדע טכנולוגיה וחברה, אוניברסיטת בר-אילן; ספריית מכללת דוד ילין  
[eilat.navon@gmail.com](mailto:eilat.navon@gmail.com)

ד"ר ישראל בלפר, התכנית למדע טכנולוגיה וחברה, אוניברסיטת בר-אילן [Israel.Belfer@biu.ac.il](mailto:Israel.Belfer@biu.ac.il)

לזכרה של יהודית בר-אילן ז"ל, מורה יוצאת דופן ששילבה כלים מתמטיים בעולם מחקרי-מילולי. למדנו ממנה שהגדרות נעשות על ידי אחרים, שלא חייבים להיות שגרתיים בגישה אל עולמות ידע. יש סיפורים שאפשר לקבל מידע ותבניות כמותיות ולא רק מאותיות ומילים, ושיש מגבלות גם ללקחים אלה.

### תקציר

חיפוש בכלי גילוי של הספרייה הפך בעידן המידע למיומנות נדרשת ואינטואטיבית מכל קורא, חוקר וסטודנט. אולם מקומו של המדד הביבליומטרי, המכמת את מספר הציטוטים הודחק מתוצאות חיפוש בכלי הגילוי הנפוצים ותוצאות החיפוש מתבססות על רלוונטיות של מילות מפתח וחווית משתמש. בה-בעת, חשיבותם של מדדים ביבליומטריים כאינדיקטורים לחשיבות ומרכזיות מאמרים וחוקרים, ככלי ניתוח של מטא-דאטה מדעי ומחקרי וככלים לניהול תקציבים מדעיים הולכת ומתחזקת.

להעדרו של מדד ביבליומטרי מתוצאות החיפוש, ישנה משמעות כפולה; מצד אחד חוסר יכולת להערכת מרכזיותו וחשיבותו של החומר שנדלה במסגרת הספרות המקצועית-אקדמית. מאידך, תוצאות החיפוש עשויות להקנות מרחב לגילוי ולידע חדש, לתקן הטיות של מצוינות (במובן המצומצם של ריבוי ציטוטים) ולאפשר גילוי רעיונות חדשים של מעגלי יוצרים וחוקרים המודרים ממרכזי המחקר עתירי מימון.

המאמר מציע בחינה של המשמעות של שתי שיטות לאחזור מידע המתפתחות זו לצד זו, האחת על בסיס רלוונטיות וחווית משתמש והשנייה מבוססת מדדי ציטוט ביבליומטריים.

**מילות מפתח:** תוצאות חיפוש, חיפוש וגילוי, מדדי ציטוט, אינדיקטור ביבליומטרי, קריטריון רלוונטיות, תפוקות מדעיות, חווית משתמש

## Relevancy and Centrality in Scientific Research: the significance of retrieval by relevancy vs. Bibliometric Indicators in Discovery Tools

### Abstract

The use of library Discovery tools has become a requisite intuitive skill for readers, scientists and students. However, as this article explores, the bibliometric indicator that quantifies citations, signaling centrality and importance of the scientific communication, is suppressed from search results in the major discovery tools. Instead, search results are retrieved based on keyword relevancy and user experience.

The absence of the citation indicator from search results has dual significance; it creates an inability to estimate centrality and importance ("impact") of scientific communication, yet at the same time, unhindered by citation count, research results can generate a new space for discovery of new knowledge, and correct biases of science produced in resource rich scientific institutions.

The article offers a review of the significance of the two types of information retrieval that are developing alongside each other, one based on citation indicators, the other on keyword relevancy and user experience.

**Keywords:** search results, search and discover, citation count, bibliometric indicator, relevance criteria, research output, user experience

## מבוא

בכל רגע נתון עשרות אלפי חוקרים ותלמידים מבצעים חיפוש בכלי הגילוי של הספרייה האקדמית במטרה ללמוד, לנתח, לחקור ולהוסיף לגוף הידע המדעי את תרומתם המחקרית. מאמר זה בוחן את ההשלכות – עבור תלמידים וחוקרים – של ההכלה או ההדחקה של הדירוג הביבליומטרי מתוצאות החיפוש בכלי הגילוי, בשעה שחשיבותו ככלי ניתוח של מטא-דאטה מדעי ומחקרי הולכת ומתחזקת.

עמוד התוצאות הראשי של כלי הגילוי שמשמשים מוסדות אקדמיים אינו כולל מדד ביבליומטרי. מדדי ציטוט גם אינם מובנים בתוך הפלטפורמה של כלי הגילוי, או במסנני החיפוש (Facets). הם יופיעו, בהנחה שהמוסד רכש אותם באופן פרטני, רק לאחר כניסה לפריט המבוקש. במקום זאת מציע כלי הגילוי נתיבי ציטוט (Citation Trail) ללא אינדיקטור מספרי.

המאגרים הביבליומטריים Web of Science ו-Scopus מתממשים עם כלי הגילוי האקדמיים המרכזיים Primo של Ex Libris<sup>1</sup> ו-EBSCO Discovery Service (EDS)<sup>2</sup>, אולם הגישה אליהם תלויה בבחירת המוסד האקדמי לרכוש אותם, שכן הם נמצאים מאחורי חומת תשלום גבוהה. בפועל, גם במוסדות שרכשו גישה למאגר ביבליומטרי ומשתמשים ב-EDS לא יראו את הדירוג הביבליומטרי כחלק מתוצאות החיפוש בעמוד הראשי של תוצאות החיפוש, ולעיתים גם לא בעת כניסה לפריט.

ברירת המחדל עבור אחזור מידע בכלי הגילוי הינו ממשק של תיבת חיפוש יחידה. האחזור מניב חוויית משתמש משופעת בתכנים, חלקם מכתבי עת שפיטים, חלקם מדעיים וחלקם פופולריים; לצד התוצאות מופיעים מסנני מיקום, שיפוט, תוכן, שפה, זמינות ואחרים. אולם בהעדר מדדים ביבליומטריים, אף שאלו אינם חפים מביקורת, נשללת מהמשתמש אינדיקציה לאימפקט (במובן המדיד של ספירת ציטוטים ונגזרותיה) על הספרות המקצועית ולמרכזיות של מאמרים בתחומם ואף נשללת היכולת לזהות חוקרים מובילים בתוצאות החיפוש. התוצאה היא מעין השטחה של הידע המדעי לרובד מאוחד.

אולם יחד עם זאת העדר המרכיב הביבליומטרי מתוצאות חיפוש עשוי גם להקנות לחוקרים מרחב לגילוי ולידע חדש, לתקן הטיות של מצוינות ולאפשר גילוי רעיונות חדשים של חוקרים שמודרים ממרכזי המחקר עתירי המימון שבהם חוקרים בעלי צבר ציטוטים משמעותי, וכתבי עת מולטידיסציפלינריים חדשים ללא ותק<sup>3</sup> ביבליומטרי ראוי אך בעלי חשיבות חדשנית.

<sup>1</sup> לשם הערכת סדר הגודל של חברת אקס-ליבריס ומוצריה, וחדירתן לשוק צרכנות המידע האקדמי, ראו "הפתרונות של החברה מוטמעים ב-43 מתוך 50 האוניברסיטאות המובילות בעולם וביותר מ-40 ספריות לאומיות." אורפז, ע (2015) עסקת ענק: פרקווסט האמריקאית רוכשת את אקס ליבריס הישראלית בכחצי מיליארד דולר. דה מרקר.

<sup>2</sup> <https://www.ebsco.com/public-libraries/products/ebsco-discovery-service> EBSCO על כלי החיפוש של EDS

09:51:34 10/12/2021

מטרת מאמר הסקירה הנוכחי היא בחינה ביקורתית של המשמעות של תוצאות אחזור מידע על בסיס חוויית משתמש לעומת על בסיס מדדים ביבליומטריים. המאמר אינו מבקש לערוך השוואה כמותית בין תוצאות האחזור או להצביע על תבנית מובהקת, אלא להאיר בפני הקורא את השוני בין התוצאות ולנתח את המשמעויות ששוני זה יוצר. לצורך כך נעשתה בחינת התנסות מצד חוויית משתמש של תוצאות אחזור – אחזור על פי רלוונטיות ללא אינדיקטור ביבליומטרי ועם אינדיקטור ביבליומטרי – על שלוש פלטפורמות: פלטפורמת Primo ללא מינוי על מדד ציטוט, פלטפורמת EDS של EBSCO ופלטפורמת Web of Science (WoS). ממחקר ראשוני זה עולה כי מרכזיותם של המדדים הביבליומטריים לצורכי מדידת המפעל המדעי הולכים ומשתכללים, וכי השימוש בהם כאינדיקטורים לחלוקת מימון מחקרי וקידום אישי מתרחב וכולל קבוצה חדשה של כלי מדידת מטא-דאטה של תקשורת מדעית מקוונת – כלים אלטמטריים (Altmetrics). אולם בד בבד מרכזיותם של המדדים הביבליומטריים כאינדיקטורים של איכות ומרכזיות של מחקר מודחקת לטובת חוויית משתמש (User Experience) עשירת תכנים המבוססת על אחזור לפי רלוונטיות של מילות מפתח. מכאן שניתן לראות כיצד מתפתחות זו לצד זו גישות אחזור וגיטוש אחר מידע מועיל ומשמעותי, האחת מבוססת על רלוונטיות סמנטית והשנייה על מיקום בשדה ההשפעה והכוח של הספרות המקצועית.

המאמר מחולק כדלהלן: החלק הראשון של המאמר דן במדדים ביבליומטריים תוך התייחסות לביקורת המופנית כלפיהם לצד השימוש הגובר בהם לצורכי מימון וניהול מדעי. חלקו השני עוסק ברלוונטיות של כלי הגילוי של PRIMO ו-EDS ובחוויית המשתמש שהם מעניקים. החלק השלישי מתייחס לגיוון במקורות ולסוגי הפריטים (התקשורת המחקרית) המאוחזרים בתוצאות החיפוש של כלי הגילוי. פרק הדיון בוחן את המשמעות המחקרית והמדעית של תוצאות החיפוש עם מדדים ביבליומטריים ובלעדיהם.

## שאלות המחקר

א. במה שונה אחזור על פי רלוונטיות וחוויית משתמש מאחזור על בסיס מדדים ביבליומטריים?

ב. מה המשמעות של מחקר ומדע המבוסס אחזור על פי רלוונטיות לעומת אחזור על בסיס מדדים ביבליומטריים?

## שני משלים ומועקת מידע

### משל ראשון על קטלוג ותוכן

בספר "האיש ללא תכונות" מתאר רוברט מוסיל את הספרן האולטימטיבי, שנוזף בגנרל הסבור שמציאת רעיון כלשהו בספרייה בת מיליוני הכרכים כרוכה בקריאת ספר כלשהו:

'if you want to know how I know about every book here, I can tell you! Because I never read any of them . . . The secret of a good librarian is that he never reads anything more of the literature in his charge than the titles and the table of contents. Anyone who lets himself go and starts reading a book is lost as a librarian,' he explained. 'He's bound to lose perspective.'

'So,' I said, trying to catch my breath, 'you never read a single book?'; 'Never. Only the catalogs.' (Musil, 1995, pp. 501-502)

לפי תמונה ספרותית זו של אמנות-מדע המידענות והספרנות, מציאת פריט רלוונטי איננה קשורה בהבנת התוכן של הפריט כלל וכלל. יתר על כן, ניסיון ההבנה על ידי קריאה איננו אלא הסחה מן המשמעות האמיתית – ההתמצאות במיון וקטלוג הכרכים. המומחיות המיוחדת של הספרן היא בדרגה גבוהה יותר של ההתמצאות בנבכי המיון עצמו. הספרייה איננה אלא גיאוגרפיה של מרחב רעיוני, שבו המיקום מתלכד עם הרעיון, והמיקום איננו אלא ניווט באמצעות כלי המיון.

כשהספרייה, או בגרסה הנוכחית – כלי הגילוי במסד הנתונים, שואפים לגודל אין-סופי, ניתן להגיע לרמת דיוק או אחזור הולכת ומתחדדת לפי המיקום הממוין. כל ביטוי ומהלך מחשבה הולכים ומקבלים כרכים המתפרטים ומתמיינים כך שהמיקום על המדף איננו אלא קוד לתכולת הספר, ופעולת המיון היא יוצרת של מידע חדש לא פחות משהיא הוספת כרך לאוסף. הקוראים המחפשים ידע צריכים את סיועו של מורה-דרך האמון על המיפוי של תחום למיקום.

במונחים של עידן המידע הנוכחי, מושג החיפוש הוא המעצב של פרופיל המידע שאותו הוא משיג, ולא רק מאפשר הגעה אל יעד מוגדר.

### משל שני על הספרייה של בבל

חורחה לואיס בורחס – המגלם את ארכיטיפ הספרן העיוור – מוכר בשל כתביו הפיקטיביים-פנטסטיים המתאגרים רעיונית ואנליטית. הוא נותן ביטוי לרעיונות וחלומות שונים המבליחים בעידינים, בתודעות ובתרבויות שונות. אחד מאותם מוטיבים הינו זה של הספרייה, ההופכת לזו של מוסיל בכך שאין ולא יכול להיות לה אמצעי קטלוג קובע וסמכותי.

ספרייה זו מורכבת ממבנה משושה החוזר על עצמו בחיבורם של מסדרונות קצרים. ארבעה מקירותיו של כל חדר מכוסים בחמישה מדפים, שעל כל אחד מהם שלוש וחמישה ספרים בעלי גודל קבוע – ארבע מאות ועשרה עמודים, שבכל אחד ארבעים שורות בנות כשמונים תווים. הספרייה העצומה מכילה את כל הווריאציות האפשריות בתבנית המאכלסת כל פרמוטציה אפשרית של צירוף אותיות. כמות הספרים הנוצרת היא עצומה, הרבה מעבר ליכולת האנושית לסקור באופן אישי. בתמורות האפשריות של ספרייה-עולם זה מרחף תמיד הצורך למצוא את הפריט הנכון, וזה בתורו מוביל לשלל תרחישים:<sup>4</sup>

“ישנם חמישה מדפים לכל אחד מקירות המשושה; כל מדף מכיל שלוש וחמישה ספרים בפורמט אחיד; כל ספר בן ארבע מאות ועשרה עמודים; כל דף בן ארבעים שורות, בכל שורה כשמונים אותיות בצבע שחור. יש גם אותיות על עמוד השדרה של כל ספר; אותיות אלה אינן מציינות או מגדירות את מה שהדפים יגידו... אכן היה להם כל שילוב אחד של אותיות זמין, כלומר כל סיפור אפשרי היה זמין. הוגה זה הבחין כי כל הספרים, מגוונים ככל שיהיו, מורכבים מאותם אלמנטים - הרווח, הנקודה, הפסיק, ועשרים ושתים אותיות האלף-בית. הוא טען גם עובדה אשר המטיילים אישרו: בספריה העצומה אין שני ספרים זהים. משתי הנחות היסוד הבלתי מעורערות הללו הסיק כי הספרייה היא כוללת וכי מדפיה מתווים את כל השילובים האפשריים של קצת יותר מעשרים סמלים אורתוגרפיים. ... ישנם מחפשים רשמיים, אינקוויזיטורים. ראיתי אותם בביצוע תפקידם: הם תמיד מגיעים עייפים מאוד ממסעותיהם” (בורחס, 1998, עמ' 74)

בספרייה של בורחס מבוטאת הקצנה מתמטית של משחק המחשבה הקומבניטורי של קופים ומקלדות שבסופו של דבר מממש את כל הסונאטות של שייקספיר, או בעצם את כל התבניות האפשריות של תמורות האותיות. האווירה העולה מתוך "הספרייה של בבל" משלבת מוטיבים של מסתורין מימי הביניים עם עתידנות של עידן המידע; עם מסדר של מלומדים-נוודים וספרנים מקומיים המבליים את חייהם במקום אחד או בחיפוש מתמיד אחר גרסת מקור חמקמקה. האתגר המוחשי ביותר המבוטא אצל בורחס הוא מציאת הפריט מקורי, הגרסה האותנטית, יצירה מסוימת בטהרתה. באפשרות של חזרה בכל פרמוטציה אפשרית טמונה גם החרדה המתמדת מפני מציאת הגרסה הכמעט-מושלמת-אך-שקרית של היצירה המבוקשת. הרובד העמוק יותר הוא הקושי שבמציאת *אותנטיות ומשמעות* כאשר כל חיבור אפשרי והגרסאות הסובבות אותו מתעתעים כמו ענן של אפשרויות פוטנציאליות.

<sup>4</sup> בורחס חזר למוטיב זה של הספרייה השואפת לאין-סוף שוב, לאחר הפרסום המקורי בשנת 1941 בניגוד לסדר המופתי המאפיין את המיון לפי מוסיל, האפשרויות השונות להופעה בספרייה של בורחס כוללות את 'הקטלוג המהימן של הספרייה, מילין על מילין של קטלוגים מזויפים, הוכחת זיופם של הקטלוגים הללו, הוכחת זיופו של הקטלוג האמיתי' (בורחס, 1998, עמ' 74). לא ניתן להסתמך על המיפוי המלא שמציע מוסיל כאמצעי מוסדר לידע ומשמעות.

בגרסתו של בורחס החיפוש המתמיד הוא חלק אינהרנטי מן המתח שהוא מראה בין האפשרי לבין הנגיש, בין המידע המשמעותי וזה שמגיע אלינו בפועל. פעולת השיטוט ברחבי העולם-ספרייה יש בה מן המזל ומן היצירתיות האפשריים, וכל ביקור הוא מיוחד.

עולם המידע בגלגולו הנוכחי שואף לקנה המידה של הספריות הבדיוניות; גישת המיון-שהוא-המשמעות אצל מוסיל מצינת ודאות של יחסי מפה-טריטוריה בין המיון והתוכן, יחסים המאפשרים תמרון במרחב הרעיוני אפילו מחוץ לספרים עצמם. זאת בעוד שאצל בורחס מתערערת האפשרות למיון מלא, לסידור של כל המסלולים האפשריים ברחבי עולם הידע ואף לביטחון במשמעות המדויקת של כל ספר כשלעצמו. בהופעת הספרייה אצל בורחס מתחדדת מועקת התרים אחר פיסות מידע משמעותיות – כאשר אלה עומדות בפני עצמן, ללא הקשר וללא יחס עם מידע אחר. אפילו הספרנים העומלים חיים שלמים בספרייה הזאת – בניגוד לספרן של מוסיל המספק מערכת התמצאות – אינם פטורים מהתחושה המתמדת של אי ודאות, ויחד עם זאת של הפוטנציאל המתמיד לגילוי.

### מדדים ביבליומטריים

מדדים ביבליומטריים הם מושג כולל לקבוצה של מדדים המשמשים לניתוח כמותי של היחסים והקשרים בין מאמרים מדעיים וכתבי העת שבהם הם מתפרסמים. בגילוי ואחזור מידע המדדים משמשים כאינדיקטורים למרכזיות וחשיבות של סוגיות מחקריות ומעידים על צמיחה ובשלות של תחום מדעי. המדדים מאפשרים זיהוי חוקרים מובילים, מחקרים מרכזיים וכתבי עת בולטים באיכותם. ביבליומטריקה משמשת לניתוח תוצרים מדעיים ולמיפוי רשתות ביבליומטריות ולמענה על שאלות כגון מי חקר, מתי, היכן ומה נחקר (Romanelli et al., 2021). כנתוני מקרו המדדים מאפשרים מיפוי של תמורות מדעיות ברמת הדיסציפלינה המחקרית, ברמת המוסד האקדמי והמדינה וככלי לזיהוי פריצות דרך מדעיות ברמה גלובלית.<sup>5</sup> ככלי להיסטוריוגרפיה של המדעים הנתונים מקנים כלי להסקה על קשרים בין תמות מדעיות והתפתחותן (McBurney & Novak, 2002), וכן כלי לכימות ולהארה של תהליכי התקשורת המדעית והתפתחות הדיסציפלינות באמצעות פרסומיהם וניתוח של מטא-דאטה של התקשורת המדעית.<sup>6</sup>

ככלי אדמיניסטרטיבי המדדים מקנים יכולת כימות והשוואה של תפוקה מדעית של חוקרים, מחלקות ומוסדות להשכלה גבוהה. קרנות מימון משתמשות בנתונים ביבליומטריים ואלטמטריים ככלי סינון והעדפה של חוקרים וגופי מחקר על פני עמיתיהם, ומדינות משתמשות בכלים אלו כבסיס לפיתוח מדיניות מחקר (Prichard, 1969).

<sup>5</sup> להרחבה ראו Shape of Science, כלי הוויזואליזציה של SCIMAGO: <https://www.scimagojr.com/shapeofscience>

<sup>6</sup> בשנת 1969 ההתייחסות לתקשורת מדעית כתובה היא לפורמט של מאמר

## היסטוריה

מדדים ביבליומטריים החלו להתפרסם בסוף המאה ה-19 ובתחילת המאה ה-20. המדד הראשון מיוחס ל-Shepards' Citation, שיצא לאור בשנת 1873 (Garfield, 1955). אולם קפיצת המדרגה המדעית התרחשה בשנים שלאחר מלחמת העולם השנייה בד בבד עם התפתחותן של תיאוריות סוציולוגיות שחקרו את המדע בתוך המסגרת החברתית שממנו הוא נוצר ואת האימפקט של מדענים על תגליות מדעיות.

גארפילד הציג את הרעיון של שימוש בניתוח ציטוטים מתוך מאמרים מדעיים בשנת 1955, ואז גם עמד על הצורך באינדקס לשם "איגוד רעיונות" (association of ideas index) – אינדקס של יחסים בין תוצרי מחקר, שלא ניתן לקבלם תחת חיפוש ישיר של מילות מפתח ובכך יאפשר לחוקר למפות התפתחות מחקרית של רעיונות במדעים. בשנת 1961 יצא לאור הכרך הראשון של ה-Science Citation Index (SCI), שזו הייתה מטרתו העיקרית. המדדים הביבליומטריים שהאינדקס מסתמך עליהם נכתבו עוד קודם לכן, אך לאיגוד המידע באינדקס אחד כפי שעשה גארפילד, הייתה השפעה מרחיקת לכת על עולם המדע.

המדדים הביבליומטריים ניתנים לתיאור כחלק מקבוצה של מדדים המבוססים על עיקרון פארטו (Pareto principle) (יאיר, 2018), מעין 'כלל אצבע' לפיו 80% מהפעילות בתופעות רבות מקורם ב-20% מהגורמים הפעילים. בביבליומטריקה מדדי פארטו משתמשים בכימות המטא-דאטה של תוצרים מדעיים באמצעות חישובים מתמטיים וסטטיסטיים. אלפרד לוטקה פרסם בשנת 1926 מדד התפלגות תדירות של הפריון המדעי (Lotka, 1926). המדד מסביר כיצד מספר קטן של מדענים יפרסמו מספר גדול של פרסומים. חוק הפיזור של ברדפורד (Bradford's Law of Scattering) משנת 1934 מסביר את הפיזור של מחקרים מדעיים בכתבי עת. חוק זיפף (Zipf, 1949) מתאר את התפלגות השכיחות של מילים בתוך טקסט וקובע כי מילים מעטות חוזרות פעמים רבות ומילים רבות חוזרות פעמים מועטות (Price, 1963).

מדדים המבוססים על עיקרון פארטו בביבליומטריקה נתונים במחלוקת ארוכת שנים לגבי איכות התוצרים. הטענה שמספר קטן של חוקרים ופרסומים אחראים למספר גדול של מאמרים וציטוטים, משמעותה היא - שקיימת עילית של חוקרים שמחקריהם מתפרסמים בכתבי עת מובילים ומצוטטים בתדירות ניכרת על פני עמיתיהם. (Belter, 2015; Garfield, 1980) למרות הרתיעה מהכללה שכזו, עיקרון זה משמש בסיס לתיאור היסטורי, התנהגותי וסוציולוגי של המדע.

בשנים הבאות גארפילד הרחיב וביסס את מנעד השימושים של כלי הציטוט למתודולוגיות בהיסטוריוגרפיה של המדעים, וככלי למדידת אימפקט מדעי ודפוסי פרסום של מדענים. כדי

לבסס את החשיבות של מדדי הציטוט ככלי היסטורי וככלי מתודולוגי חדש עבור ההיסטוריון המדעי, שמאפשר לחשוף את המקור של הגילוי והיצירתיות המדעית (Garfield et al., 1964, p. i), גארפילד הציג ניתוח השוואתי של ספרו של אסימוב "הקוד הגנטי" עם המחקרים כפי שהם מופיעים באינדקס הציטוטים של ה-Genetics Citation Index 1961, אינדקס ציטוטים בתחום הגנטיקה שמקורו ב-SCI של אותה שנה (Garfield et al., 1964). הספר מתאר את ההתפתחות המדעית שהובילה לגילוי הקוד של ה-DNA והמחקרים שהובילו לפריצת הדרך המחקרית, והבחירה בו לא הייתה מקרית – שכן הגילוי של קוד ה-DNA נעשה במקביל לפיתוח ה-SCI.

בהמשך, בדברים שנשא גארפילד בפני קונגרס ספרנים במקצועות הרפואה בשנת 1970, הוא הציג את עצמו כביבליוגרף שהפך להיסטוריוגרף וסוציומטריסט.<sup>7</sup> גארפילד ציין שהוועדה המייעצת של SCI הושפעה מחבריה הסוציולוג רוברט מרטון והפיזיקאי-היסטוריון דרק דה סולה פרייס. הוועדה, כך תיאר זאת גארפילד, ייצגה את הזן החדש של "מדעני המדע", "שם המתאר זן חדש של סוציומטריסטים בתחום ההיסטוריה, הסוציולוגיה, הכלכלה וחקר ההתנהגות של המדע והמדענים" (Garfield, 1971).

התפתחות כלי הציטוט המושתתים על מדדים ביבליומטריים הביאה ליכולות כימות והשוואה של תוצרים מחקריים וחקרים והפיקה אינדיקטורים לחשיבות ומרכזיות מדעית ומחקרית. אך דווקא במסגרת הטכנולוגית הנוכחית, שבה כיווני פיתוח אלה מתאפשרים באמצעות המימוש של טכנולוגיות התקשורת והמחשוב, נחשפות גם בעיות עם השימוש במטריקות אלו כאמצעי להערכת הפעילות המחקרית.

### **ביקורת על ביבליומטריקה: איכות, אימפקט ומשמעות**

Who dares to balance one paper of Einstein on relativity against even a hundred papers by John Doe, Ph.D., on the elastic constant of the various timbers (one to a paper) of the forest of lower Basutoland? (Price de Solla, 1963, p. 36)

כבר עם פרסום הכרך הראשון של ה-SCI החלה להישמע ביקורת על השימושיות שלו. בחלק זה של המאמר נעמוד על חלק לא ממצה של ביקורות על הכלים הביבליומטריים.

את הביקורת של אחד המבקרים הראשונים של ה-SCI, ס' ו' קלוורדון (Cleverdon, 1964), ניתן לחלק לשניים: חדשנות ועיתוי. החלק הראשון של הביקורת קשור בחדשנות של האינדקס,

---

<sup>7</sup><https://www.britannica.com/science/sociometry>

אשר יצא בחמישה כרכים ועלה סכום כסף משמעותי לתקופתו, כך שהוא לא היה נגיש מחוץ לאקדמיה. נוסף לכך, במהדורותיו הראשונות האינדקס לא פירט את שם המאמר אלא רק את שם המחבר וכתב העת – מה שהקשה על בחינת הרלוונטיות של הציטוט. בעיות יסוד אלה בחלקן כבר אינן רלוונטיות היום, ובמאה ה-21 אנו רואים התפתחות מהירה בכל האגפים הכמותיים (Bar-Ilan, 2008). עם זאת, גרסאות אחרות של אתגרים להסתמכות ונגישות של מידע מחליפות את מקומן. לדוגמה גם בעידן הנוכחי יש לזכור כי האחזור אינו מושלם מפאת חוסר האחידות במטא-דאטה של הספקים (Sonawane, 2017) וכן שחשיפת החוקרים איננה מתרחשת רק מעל דפי כתב עת מודפס – יש נוכחות מקוונת הולכת וגוברת באתרים אישיים, רשתות חברתיות ופורטלים אחרים (Bar-Ilan et al., 2012).

הביקורת השנייה רלוונטית גם כיום, וקשורה לפער בין מועד ההוצאה לאור של המאמר בכתב העת לכניסתו לאינדקס (Cleverdon, 1964). יש הטיות מובנות, כגון "אפקט מתיו" (The Matthew Effect) של מרטון – ככל שמאמר ומחבר מצוטט יותר, כך סביר שהוא ימשיך ויצוטט מספר רב של פעמים גם בעתיד (Merton, 1968).

גם בהנחה שמדובר במסד נתונים מלא בעל אמינות אחידה – הביקורת גדלה לצד השימוש ההולך וגובר במדדים הביבליומטריים.

### **שיקוף ראוי של איכות מדעית**

בורמן ודניאל (Bornmann & Daniel, 2008) מחלקים את המשתנים שמשפיעים על כמות הציטוטים לשבעה תחומים: תלויי זמן, דיסציפלינה ומאמר, כתב עת, מחבר, זמינות ובעיות טכניות.

השאלה הבסיסית המערערת על המדדים הביבליומטריים של מספרי ציטוטים ודרכי המעקב אחריהם נוגעת לאיכות, קרי כמות אינה כשלעצמה מדד לאיכות (McBurney & Novak, 2002), ויש לגשר בין ממדי האיכות שרלוונטיים לחוקרים עצמם לבין המדדים הביבליומטריים, יחד עם הנגשת אלה האחרונים באופן ברור יותר (Szomszor et al., 2021).

למקה ואחרים (Lemke et al., 2021) ציינו את הנטייה של חוקרים להעדיף מחקר ופרסומים שמקנים מדדים ביבליומטריים גבוהים, לצד רתיעה ממחקר בנושאים חדשניים ופרסום בכתבי עת חדשים, שהם אולי משפיעים אך בעלי אימפקט פקטור נמוך (Bornmann & Daniel, 2008; McBurney & Novak, 2002).

מכאן שהשגת מספר ציטוטים גבוה הופכת למטרה משמעותית יותר מהמחקר. מדדי ציטוט אף מקדמים ניוון מחקרי – החוקר יימנע מלקחת סיכון מחקרי אשר עשוי שלא להניב הערכת עמיתים גורפת.

ביקורת נוספת מתייחסת לכמות הפרסומים שמגיעה על חשבון איכות (McBurney & Novak, 2002; Romanelli et al., 2021). בהחלטה שהתקבלה על ידי קרן המחקר הגרמנית Deutsche Forschungsgemeinschaft (DFG) (Deutsche Forschungsgemeinschaft, 2010) בשנת 2010 תחת הכותרת "איכות ולא כמות" הוגבל מספר הפרסומים שיש להגיש במסגרת בקשה לתקצוב מחקר. החוקרים התבקשו להגיש מספר מוגבל של פרסומים משמעותיים במיוחד. בקורות החיים הוגבל מספר הפרסומים שיש לצרף לבקשת מחקר והחוקרים אף התבקשו להתמקד בפרסומים הרלוונטיים לבקשה בלבד. באטלר, שמצוטטת במאמר של דה ריק ואחרים (De Rijcke et al., 2016, p. 163), מציינת את העלייה בפריון הפרסומים האקדמיים באוסטרליה שחלה בד בבד עם ירידה באימפקט של הפרסומים. נתונים דומים נמצאו גם בספרד.

מעבר לביקורות מסוג זה, יש אלמנטים טכניים-מעשיים המשפיעים על כמות הציטוטים (Bornmann & Daniel, 2008):

- התכיפות של פרסום הגיליונות של כתב העת מעיד על חשיבות הנושא ומכאן על מספר הציטוטים.
- סדר הופעת המאמר בגיליון רומז על השפעתו העתידית; המאמר הראשון בגיליון מיצר יותר ציטוטים מאלו של הבאים אחריו.
- היוקרה, האיכות, נגישות נראות ובינלאומיות של כתב העת.
- מספר המחברים של מאמר ישפיע על מספר הציטוטים.
- אורך המאמר; ככל שהמאמר ארוך יותר הוא יכול יותר ציטוטים.
- זמינות מאמרים עם גישה פתוחה מביאה לחשיפה גבוהה יותר וכך ליותר ציטוטים.
- ככל שיש יותר מאמרים חדשים, מספר הציטוטים הכללי עולה הסתברותית (בגידול מעריכי).
- מאמרים חדשים יצוטטו יותר שכן כמותית יש יותר מהם.
- שפת פרסום המאמר והרשת המחקרית-חברתית של החוקר עצמו.

נוסף לכך, עולות שאלות לגבי הגדרת תחום ההפעלה של הכלים. המשמעות הראשית של ספירת הציטוטים היא השוואה פנים-דיסציפלינרית, שכן אי אפשר להשוות איכות של כתבי עת בין דיסציפלינות, אלא רק לבחון השפעה בתוך הדיסציפלינה (McBurney & Novak, 2002). גודל הדיסציפלינה קובע את כמות הציטוטים – דיסציפלינה קטנה תנפיק מספר קטן של ציטוטים ולהפך (Bornmann & Daniel, 2008).

כמו כן יש לציין שהגדרת הדיסציפלינה איננה ברורה לגמרי או מוסכמת מנקודות מבט שונות, ואם ההסתמכות היא על מסגרת הציטוטים על מנת לייצר את התיחום, אז יש כאן בעיה של הגדרה מעגלית. לפיכך בחירת מסגרות פרסום שונות באותה דיסציפלינה אך בברירת מעגלים מסוימים שלה, יכולה להטות את הכף באופן מלאכותי לטובת ציטוטים מרובים באגף אחד על פני השני, כפי שהראו פרז ובר-אילן לגבי כתבי עת משפטיים במסגרות של כתבי עת מקצועיים לעומת אלה השייכים ליחידות האקדמיות-משפטיות (Perez et al., 2019).

### **אדמיניסטרציה, קידום ומימון**

כיווני ביקורת נוספים עוסקים בהשלכה של ההסתמכות על המדדים הכמותיים במסגרות של ניהול, תקצוב ודירוג:

- חדשנות אינה מתוגמלת – פחות כדאי לפרסם בכתבי עת חדשים, אף שהם עשויים להיות משפיעים מאוד, מפני שהם בעלי אימפקט פקטור נמוך (Bornmann & Daniel, 2008; McBurney & Novak, 2002).
- אדמיניסטרציה – שימוש במדדים כמותיים לקבלת החלטות ללא התייחסות לאיכות השימוש במדדים לתיאור ולא להערכה (McBurney & Novak, 2002).
- הערכת העמיתים יכולה להיות מושפעת מהפן הכמותי.
- ההסתמכות המינהלית והכספית על המדדים הביבליומטריים יכולה להביא לאיבוד האוטונומיה המחקרית של החוקרים על ידי הפעלת לחצים חיצוניים שאינם נוגעים למחקר גופא (Jappe, 2018).

### **צמיחה ללא הסדרה**

ביקורות מסוג נוסף העולות בשנים האחרונות נוגעות בשימוש-יתר וביישום בלתי מוסדר של מדדים ביבליומטריים או דומיהם בטווח הולך וגדל של תחומים (González-Alcaide, 2021). השימוש בטכניקות הולכות ומתפרטות, הולכות ומשתנות באופן עצמאי של ניתוח הנתונים והגישה למספרים הרלוונטיים בפרסומים, מביאים לכך שאין אחידות או סטנדרטיזציה מוסדרת של אופן הכימות, זאת דווקא עם עליית ההסתמכות על טכניקות אלה (Mejia, 2021). אין "שיטת מיון לשיטות המיון" כמו זו המשמשת את הספרן של מוסיל. המצב המתקבל הוא מגוון של שיטות הערכה המבוססות על עיבודים שונים של המטא-דאטה של הפרסומים. שיטות אלה משמשות בתחומים שונים ולמטרות שונות, ללא עיקרון מאחד, תוך שחיקה של משמעות ההערכה הכמותית עצמה.

### **עלות Serial Crisis**

ראוי אף להתייחס לעלותם הגבוהה של מאגרי הציטוטים (Tabacaru, 2019). ב-2019 עלות המינוי שהושתה על מוסד אקדמי גדול<sup>8</sup> הייתה מעל ל-212,000 דולר עבור WoS וכמעט 140,000 דולר עבור Scopus. בחמש השנים האחרונות הייתה עלייה של 4% בכל שנה בעלות המינוי ל-WoS, בעוד של-Scopus הייתה עלייה של 5% בכל שנה. מצב עניינים כלכלי זה מעמיד את הספריות ואת המשתמשים – החוקרים והקוראים – בפני צורך בבירורות נוספות. אפשרויות נוספות אלה מביאות עימן גם חוויית משתמש שונה.

### רלוונטיות וחוויית משתמש

כלי הגילוי הנגיש למשתמש בכלי הספרייה הדיגיטלי כיום מורכב מקטלוג הספרייה, מאגרי מידע ואגרגטורים של מידע אלקטרוני הנרכשים על ידי הספרייה, ספרים דיגיטליים, ה"מרחב הקהילתי" של Primo (Community Zone), ובו חומרים פתוחים לקהילת "אקס ליבריס", וחומרים בגישה פתוחה. הספרייה יכולה לתעדף את אוספיה באמצעות תוצאות החיפוש (Blended Results). כלומר פעולות המחפשים והעדפותיהם משפיעות בהיזון חוזר על המאגרים ותתי-המאגרים הזמינים.

בתוך תכנים אלו קיימים במקביל עולמות מגוונים של ידע מדעי, שפיט ופופולרי, ואלה מצריכים אסטרטגיית סינון דו-שלבית ותלת-שלבית כדי להגיע לאחזור תוצאות אופטימלי. זאת בשל ריבוי סוגי הפריטים וגיוון מקורות המידע שנכללים בכלי הגילוי, הכוללים פריטים רבים שאינם מכוסים על ידי מאגרי ציטוטים – חומר לא-אקדמי, כגון כתבי עת מקצועיים (Trade Journals), פרסומים ממשלתיים ומאמרים מעיתונות יומית וחומרים משלימים-פריפריאליים לתכנים מדעיים שפיטים, כגון פרוטוקולים קליניים, הרצאות ודו"חות טכניים וסטנדרטים.

כלי הגילוי מבוססים על הקונספט של תיבת חיפוש אחת (Single Search Box) – מושג המושאל משיטת החיפוש במנוע החיפוש של חברת "גוגל". תיבת החיפוש הבודדת מפשטת את תהליך החיפוש על חשבון דיוקו, ומשקפת את המתח בין הרצון ליצור ממשק משתמש פשוט והמחויבות של הספרייה לאוריינות חיפוש ולדיוק האחזור (Haggerty & Scott, 2019). עם זאת, תיבת חיפוש בודדת מעבירה מסר של ביטחון למשתמשים שכלי החיפוש יכול להתמודד עם צרכי המחקר מנקודת כניסה אחת (Lown et al., 2013).

תיבת החיפוש באתר הספרייה היא נקודת הגישה ומקום האינטראקציה הבלעדי של מרבית המשתמשים בספרייה האקדמית. רוב השימוש בכלי הגילוי מתבצע באמצעות אתר האינטרנט ללא תיווך של ספרן. ממחקר שערכו האגרטי וסקוט (Haggerty & Scott, 2019) עולה כי חיפוש ללא מחשבה או אסטרטגיה יכול להוביל לתוצאות מבוקשות, וכי הצורך ההיסטורי בדיוק החיפוש באמצעות שדות חיפוש הוא תוצר של מורשת מקצועית של העדפת אסטרטגיית חיפוש

<sup>8</sup> הנתונים מתייחסים לאוניברסיטת טקסס שבה יש כ-73 אלף תלמידים. מקור: אתר אוניברסיטת טקסס <https://www.tamu.edu/about/at-a-glance.html> נדלה: 25.4.2022

על פני יעילות וקלות. מחקרים רבים שנעשו העלו שהשימושיות שבתביבת החיפוש הבודדת תואמת את צרכי המשתמשים, בעיקר כאשר האפשרות הקלאסית של חיפוש מורכב בשדות תוך שימוש באופרטורים בוליאניים ובחירת המאגרים מסורבלת בהרבה ופוגעת בחוויית המשתמש.

יש לציין כי מורכבויות גילוי מוכרות בספרנות גם ללא קשר לכלי הגילוי, כגון תכנים אשר משתמשים במנגנוני שיפוט עצמאיים, כמו במקרה של מכוני מחקר עצמאיים – לדוגמה במכון ראנד (Rand Corporation), שמקיים מנגנון שמחייב את כל הפרסומים לעבור ביקורת עמיתים על ידי איש צוות ומומחה חיצוני. מכונים אחרים בעלי אוריינטציה פוליטית<sup>9</sup> אינם דורשים בקרה שיפוטית לאובייקטיביות מדעית, מבלי לציין זאת במפורש בפרסומים.

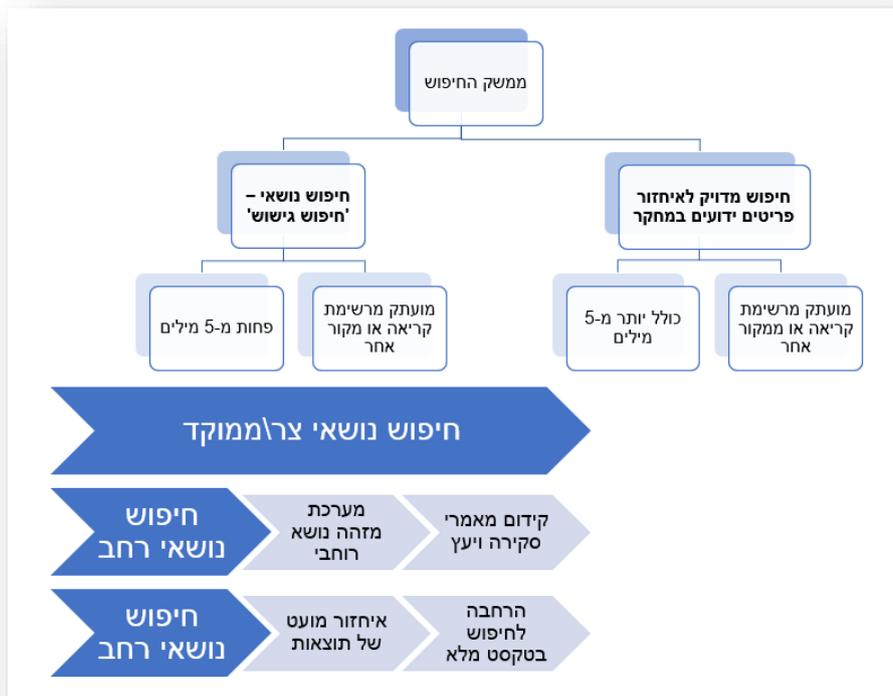
ישנה סוגיה נוספת של זיהוי של חוקרים מובילים שאינם אקדמאיים בעיקר בתחומי אומנויות וספרות. בדיסציפלינות שבהן הפן היוזאלי (תמונות, איורים, עיצובים, סרטונים וכו') מהוות את החלק הארי של המאמר והביבליוגרפיה והציטוט מצומצם, או כאשר מומחים בעלי שם הם מחזאים, אמנים או מומחים מולטי-דיסציפלינריים.

### תהליך הגילוי והאחזור ב-Primo וב-EDS

Primo הוא מגה-אגרגטור (mega-aggregator) של ספקי מקורות מידע (Exlibris, 2011; Sonawane, 2017). כאשר המשתמש מבצע חיפוש, Primo יחפש גם באוסף הספרייה וגם במקורות מידע נוספים.<sup>10</sup> המשקל של הפריטים שיאוחזרו נקבע על פי פרמטרים של רלוונטיות. לשם הפרסום או לשם המחבר יהיה משקל גדול יותר מאשר פרסומים שנכתבו על הפרסום או על המחבר, התאמה בין מילות החיפוש או המושג, מספר הפעמים שמופיעות מילות החיפוש או המושג. חשיבות אקדמית כוללת את כתב העת ומספר הציטוטים וסוג הפרסום. פרסומים חדשים יופיעו גבוה בתוצאות החיפוש.

כלי הגילוי משתמש באלגוריתם שמזהה את סוג החיפוש שנעשה ומחלק אותו לשני סוגים: חיפוש ממוקד וחיפוש נושאי (ראו תרשים 1). זיהוי של נושא רחבי יביא להתאמת תוצאות החיפוש על ידי הכנסת סקירות (reference works), מיעוט תוצאות יוביל להרחבת החיפוש לחיפוש בטקסט מלא. ממחקר שנעשה על ידי אקס ליבריס (Stohn, 2015) עולה כי תהליך החיפוש בכלי הגילוי הוא גם תהליך של למידת הנושא הנחקר.

<sup>9</sup> להרחבה על תהליך השיפוט במכון ראנד, ראו <https://tinyurl.com/2fz38z2k>  
להרחבה על מכוני מחקר ראו: The 50 Most Influential Think Tanks in the United States  
<https://thebestschools.org/magazine/most-influential-think-tanks>  
2020 Global Go To Think Tank Index Report  
[https://repository.upenn.edu/cgi/viewcontent.cgi?article=1019&context=think\\_tanks](https://repository.upenn.edu/cgi/viewcontent.cgi?article=1019&context=think_tanks)  
<sup>10</sup> הספרייה יכולה לנהל רכיב זה. תכנים מקומיים ואלו המסופקים על ידי פרימו נקראים "מעורבבים" (blended)



### תרשים 1: תהליך אחזור תוצאות חיפוש ב-Primo

(Stohn, 2015)

### תהליך הגילוי והאחזור ב-EDS

בדומה ל-Primo, כלי הגילוי EDS של EBSCO מובנה לאחזור רלוונטיות תוך בחינת התדירות של הופעת מילות המפתח במטא-דאטה ובטקסט מלא.

EBSCO (EBSCO, 2020) מאפיין את תהליך החיפוש כסביבת למידה ( Learning Environment) שמובילה את המשתמש לשיפור מילות או מושגי החיפוש. הכוונה זו המבוססת על עבודת מומחים לתחומי ידע (Subject Matter Experts) באמצעות מנגנון ששמו Knowledge Graph, מנגנון המגשר על פערים בין משתמשים חדשים למשתמשים ותיקים באחזור מידע רלוונטי ומוודא אחזור איכותי.



### תרשים 2: תהליך אחזור תוצאות חיפוש ב-EDS

(EBSCO, 2015)

ב-Primo וב-EBSCO יש רצון ליצור התאמה בין רמות שונות של מיומנויות חיפוש ומטרות חיפוש, כך שסטודנט לתואר ראשון וחוקר ותיק יוכלו להשתמש באותו הכלי. ב-Primo האלגוריתם משתמש במספר מילות החיפוש כדי לזהות אם החיפוש ממוקד (מאפיין חוקר ותיק) או שמדובר בחיפוש ראשוני, גישוש בתוך תחום ידע (מאפיין סטודנט חדש). לפי זיהוי ראשוני זה הוא מפנה את המחפש לסוגי חומר רלוונטיים. ב-EBSCO יאזנו את תוצאות ואיכות החיפוש באמצעות העשרה והכוונה נושאית (סמנטית). שילוב פונקציית ההשלמה האוטומטית

(Auto-Complete Function) של מושג החיפוש משמש גם הוא כאמצעי להכוונת המשתמש למושגים שהאלגוריתם מזהה כשכיחים.

### חויית המשתמש

כלי הגילוי התפתחו מתוך עולם המסחר והצרכנות, שבו חויית המשתמש משמעותית להצלחת מיזם ולרכישה חוזרת. אינטרס דומה קיים גם בספריות, שמעוניינות לתת לתלמיד ולחוקר (הלקוח) חויית שימוש עשירה, בעיקר בצל מגבלות תקציביות. ב-Primo חויית המשתמש באה לידי ביטוי באחזור מספר גדול של תוצאות, הרחבת אחזור הפריטים כאשר מספר התוצאות נמוך ושימוש במנגנון ההמלצות. לדברי סטון:

"Recommendations provide additional sources of information. Just like the Amazon feature "Customers Who Brought This Item Also Bought..." (Stohn, 2015, p. 10)

ב-Primo וגם ב-EBSCO קיימת האפשרות להתאמה אישית, שיתוף והמלצות. זו מוסברת (EBSCO, 2020) כך:

"Access. Search. Choose. Use. These are the stages of the library user's journey. However, users take similar journeys with Netflix, Amazon, Spotify or Google, and their experiences with the personalized dashboards, sharing options, and recommendation capabilities on these platforms have shaped their user interface expectations. Rather than reinvent the wheel when it comes to upgrading discovery tools, the new version of *EBSCO Discovery Service*<sup>TM</sup> (EDS) combines the popular features of these commercial websites with the functionalities necessary for libraries."

בדומה ל Primo, גם ב-EBSCO EDS המטרה היא לאחזר כמה שיותר תוצאות רלוונטיות בעמוד הראשון (EBSCO, 2015). על מנת לייצר חויית משתמש חיובית, האחזור מבוסס על רלוונטיות סמנטית למילות מפתח ומספר גדול של תוצאות חיפוש. הצורך במספר גדול של תוצאות אחזור מוסבר על ידי ריבוי המקורות והמסמכים שמקובצים בכלי הגילוי והצורך לספק חויית משתמש עשירה.

### מקורות וסוגי פריטים

כאמור, כלי גילוי יפעלו כדי לאחזר מידע הרלוונטי למילות החיפוש (בהתאמה סמנטית לצירוף של המילים עצמן ובדרגות שונות של היצמדות לניסוח המקורי). לצורך חוויית משתמש טובה יכללו תוצאות החיפוש פריטים רבים. בחלק זה של המאמר אנו מבקשים לבחון מהם המקורות והפריטים שמהם מורכבות תוצאות החיפוש. מהממצאים בלוח 1 עולה כי חלק מהמקורות הם אקדמיים: ספרים ופרקי ספרים, מאמרים מכתבי עת, דיסרטציות ופרסומים מכנסים. חלק אחר הוא פרפריאלי ולעיתים יסודי ומכונן ביחס למחקר: סקירות, דו"חות, מסמכים משפטיים, מפות, פטנטים, חומרי ארכיון ומערכי נתונים (data-sets). קבוצה שלישית כוללת חומרים בעלי מאפיינים פופולריים ולא מדעיים, כגון עלונים מקצועיים (trade publications), סרטים, משחקים, עיתונות, תמונות, דו"חות, סקרי שוק, סטנדרטים ועוד.

EBSCO (EBSCO Industries, 2020)		Exlibris (ExLibris, n.d.)	
<b>Academic journals</b>	Academic Journals and Peer-Reviewed Journals	Archival Manuscript	Material/
<b>Magazines</b>	Periodicals, Journal Articles, and Journals	Article	
<b>Trade publications</b>		Audio	
<b>News</b>	Newspapers and Newswires	Book	
<b>Books</b>	Books, Almanacs, Reference Books, Book Collections, and Essays	Book Chapter	
<b>Reviews</b>	Book Reviews, Film Reviews, and Product Reviews	Book Review	
<b>Reports</b>	Country Reports, Market Research Reports, Educational Reports, Analyses, Industry Reports, Industry Profiles, and Grey Literature	Computer File	

<b>Conference Materials</b>	Conference Conference Symposia	Papers, Proceedings,	Conference proceeding
<b>Dissertations</b>	Dissertations and Theses		Database
<b>Biographies</b>			Dataset
<b>Primary Source Documents</b>	Legal Documents and Primary Source Documents		Dissertation
	Government Documents		Government document
			Image
<b>Music Scores</b>	Music, Music Score, Multiple Score Format, Voice Score, and Printed Music		Journal
<b>Research Starters Content</b>	Research Starters Content		Legal Document
<b>Electronic Resoures</b>			Map
<b>Non-print Resources</b>	CD-ROMs, Games, Kits, Mixed Materials, Multimedia, Objects, Realia, Visual Materials, and Technical Drawings		Market Research
<b>Audio</b>	Audio, Audiocassettes, Music, and Sound Recordings	Audiobooks, Audiovisual,	Microform
<b>video</b>	DVDs, Motion Pictures, and Videos		Newsletter article
<b>eBooks</b>			Newspaper
<b>Maps</b>			Newspaper article

Unclassifiable non-textual sources. Includes art, artifacts, realia, special collections (Other)
Patent
Reference entry
Report
Score
Standard
Text Resource
Video
Web Resource

#### לוח 1: מקורות ופריטי מידע

בהשוואה בין EBSCO ל-Primo לא ניכר הבדל בין סוגי החומרים אלא באופן הצגתם ומיין הפריטים.

החומרים שנגישים באמצעות כלי הגילוי משקפים עושר תוכני עצום ומלהיב. אולם בה בעת נדרשת התייחסות לאופי תוצאות החיפוש, מכיוון שמילות המפתח ידלו מגוון עשיר ביותר של תוצאות רלוונטיות, אך גם כאלה שאינן בהכרח מדעיות. כאמור, לטענתנו במאמר זה אחזור עשיר כל כך ללא מדדי ציטוט לא יאפשר למשתמש לברור בין מקורות איכותיים יותר או פחות.

#### תהליך המחקר

כאמור מטרת המאמר היא בחינה ראשונית של השוני בין אחזור על פי רלוונטיות וחווית משתמש לבין אחזור התוצאות כאשר יש אינדיקטורים ביבליומטריים – במקרה זה מדדי ציטוט. לצורך בחינת תוצאות האחזור בחרנו במאמרים בשפה האנגלית וערכנו חיפוש בשני כלי גילוי במוסדות אקדמיים: ב-Primo ללא מדדי ציטוט ובמוסדות עם EDS של EBSCO.

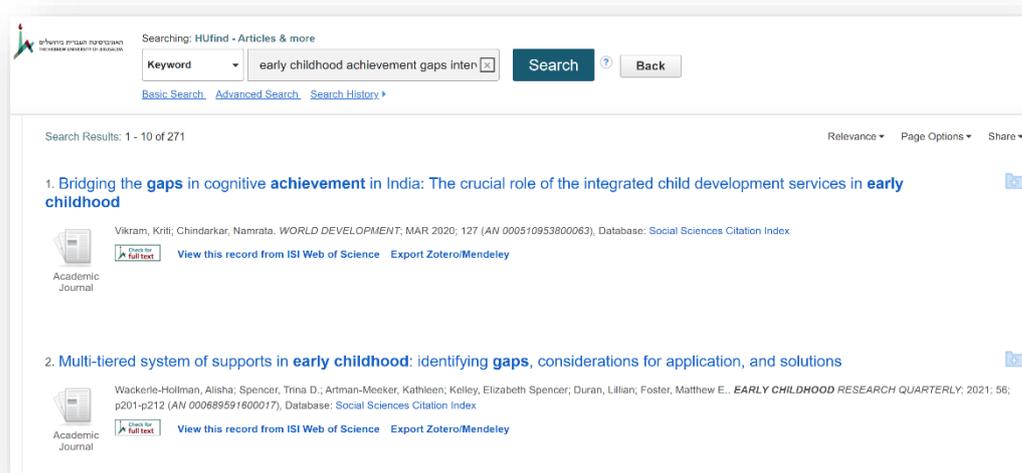
החיפוש הוגבל לשנים 2016–2022 וסומנו המסננים "כתבי עת שפיטים". כדי להימנע מהטייה שנובעת מרכש הספרייה סומן המסנן "הרחב מעבר לאוספי הספרייה" והוסר המסנן "טקסט מלא". כך וידאנו שכל המאמרים שיעלו בתוצאות יהיו שפיטים, לא מוגבלים לאוספי הספרייה

ולא בהכרח בטקסט מלא. לצורך המחקר נלקחו עשרת המאמרים הראשונים שאוחדו בכל פלטפורמה.

בשל תיחום השנים, ולאור העובדה שמדובר בשנים האחרונות וכן במאמרים שנכתבו עד לאחרונה, מספר הציטוטים לכל מאמר אינו גבוה. עם זאת, למספר הציטוטים אין משמעות שעומדת בפני עצמה במחקר זה, אלא כאינדיקטור בלבד.

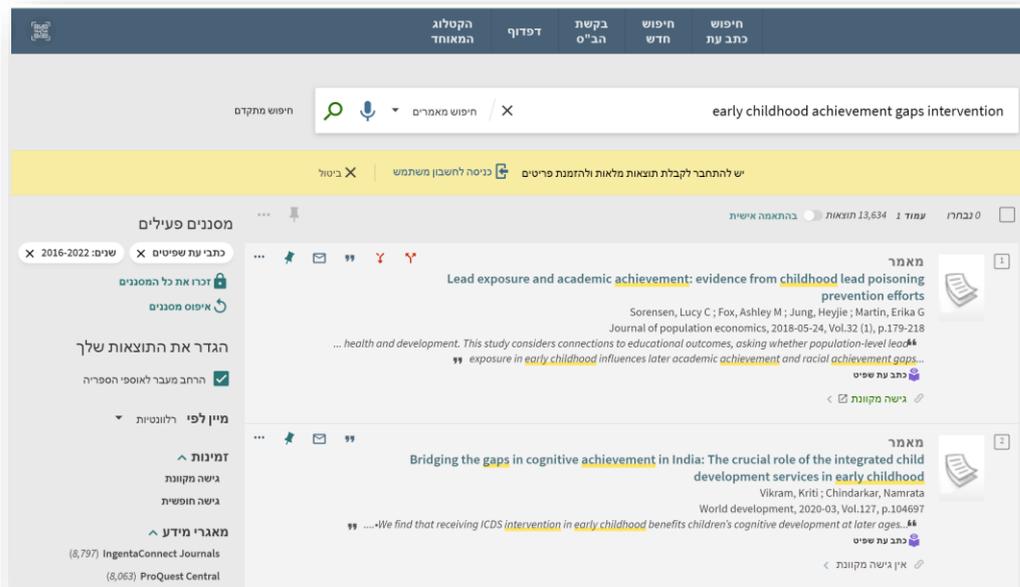
לצורך בחינת השוני בין אחזור תוצאות חיפוש על פי רלוונטיות בלבד ואחזור הכולל מדדי ציטוט, נערך חיפוש בשלוש פלטפורמות גילוי על מושג מורכב זהה. המושג נבחר באופן שרירותי והוא מורכב משלושה מונחים: "התערבות", "הגיל הרך", פערים בהישגים" (intervention, early childhood, achievement gaps).

בשלב השני, כדי לקבל אינדיקציה למספר הציטוטים, נבדקו המאמרים שנדלו ב-Google Scholar וב-WoS לצורך שיוך מספר הציטוטים למאמרים.

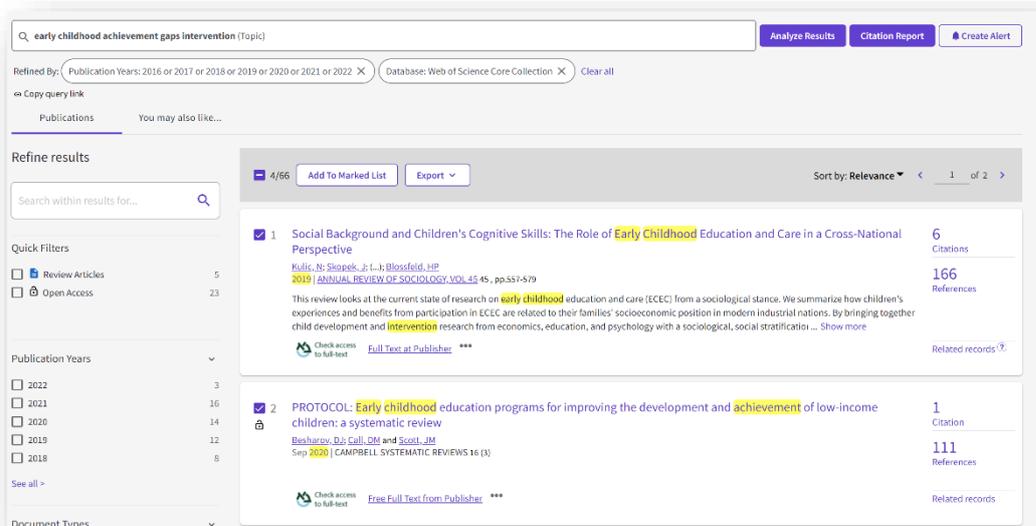


### תרשים 3: כלי הגילוי EDS של EBSCO בספרייה אוניברסיטאית

בתוצאות החיפוש אין מדדי ציטוט. כניסה לפריט נותנת קישור לפריט ב-Web of Science – בהתאם לגישה של המוסד האקדמי שהספרייה שייכת אליו. לא כל המוסדות רוכשים את אותן חבילות מידע מן הספקים, כגון WoS, ונדיר למצוא מוסד שרוכש את כל חבילות המידע.



#### תרשים 4: הצגת תוצאות ב-Primo במוסד שאינו מקבל מדדים ביבליומטריים



#### תרשים 5: תוצאות מתוך WoS הכוללות מדדי ציטוט

כתיבוד של חוויית משתמש, וכמובן לא כתחליף לניתוח עומק שיצריך נתונים רבים הרבה יותר, ניתן לדווח על השוני בתוצאות החיפוש, שוני שבחלקו משקף את הביקורות שנסקרו לעיל על המדדים הביבליומטריים: המקורות הכוללים מדדי ציטוט (תרשים 5) מסודרים בסדר מוטה לטובת רשומות מרובות ציטוטים וכאלה שנכתבו על ידי חוקרים מרכזיים, והפריטים החדשים יותר זוכים למקום מרכזי פחות. יש להטיה זו משמעות עבור סוגי המחקר המקבלים חשיפה – מקורות מענפי מדעי החברה המקבלים תיעוד ועיבוד ציטוטים נרחב יותר, כגון כלכלה ומחקר

סוציו-אקונומי, מקבלים יותר תוצאות וחשיפה גבוהה יותר. כלומר – המלל לחיפוש והמונחים הכלולים בו ("התערבות, גיל הרך, פערים בהישגים") הניב תוצאות מוטות לטובת תחומים מסוימים כאשר נבדק על בסיס ביבליומטרי, לעומת התוצאות בכלי EDS ו-Primo (תרשימים 4,5), הכוללים שילוב של תחומים מרובים כפי שמתבטא בפן הסמנטי של השאלתה.

## דיון

### האם רלוונטיות תפיק מדע אחר או שיח ציבורי אחר מתוצאות חיפוש עם אינדיקטור ביבליוגרפי של איכות?

העוצמה של המדדים הביבליומטריים – החדשים והוותיקים – לצורכי מעקב אחרי התקדמות מחקרית ותיעוד שרשרת המחקר קשורה בטבורה לגיבוש מדיניות מחקרית ומימון מחקר במוסדות האקדמיים. בשנים האחרונות התפתחו מדדי ביבליומטריקה חדשים – אלטמטריים, שמודדים אימפקט מחקרי מחוץ לעולם המחקר ברשת האינטרנט, ומדדים חדשים כגון Dimensions.ai, שאינם מוסתרים בחומת תשלום (Haustein et al., 2013). ההתפתחות של מדדים אלה משקפת את עוצמתם, והם מתפתחים גם לכיוון של איסוף והצלבה לצורך דיוק של מטא-דאטה שמייצר החוקר באמצעות כלי איסוף נתונים כגון מערכות CRIS (Current Research Information Systems). בכלי האיסוף האלו נצברים אובייקטים מסוגים מגוונים עוד יותר (כגון פעילויות, השתתפות בכנסים, קורסים, מכשירים, ואחרים) באמצעות הסטנדרט האירופי למערכות תפוקות מחקר – CERIF (Common European Research Information Format).

אם כך, מדדים ביבליומטריים משמעותיים ביותר להערכת מוקדי הידע המחקרי – ומאז עליית הפן הזה של חקר הידע, במרחב המחקרי התבססה כעובדה המוסכמה שללא תיעוד שרשרת המחקר לא ניתן לערוב לנכונותו. מדדים אלה מרכזיים ככלי להערכת קידום ומימון, אך עם זאת אינם בלעדיים לניתוח איכות הידע.

כלי הגילוי שנשען על רלוונטיות של מילות מפתח מאפשר לחוקרים אשר לא פרסמו בכתבי העת המובילים אפשרות להיכנס ולהשתתף בשיח המדעי. בכך כלי כזה ממלא את החסר עבור מחקר אינטרדיסציפלינרי וצעיר או כזה שנמצא מחוץ לטווח הכימות הביבליומטרי (בשפות מרובות, בדיסציפלינות ממדעי הרוח וכדומה) – היינו מחקרים ופרסומים שאינם מיוצגים במדויק ובמהירות על ידי מדדים ביבליומטריים, שכן המדדים מוטים לטובת דיסציפלינות "בוגרות" מחקרית. כלי גילוי שמאחזר על בסיס רלוונטיות, נותן גישה להקשר מחקרית אלטרנטיבי, כגון דוחות ממשלתיים ומאמרים מקצועיים שאינם שפיטים, אך מכילים ידע משמעותי וראוי לציטוט.

מקורות ידע מדעי שאינו אקדמי מתפתחים בשנים האחרונות גם מתוך מסגרת של מדע אזרחי Civic Science שבו הציבור הכללי לוקח חלק בפעילות המדעית. מקור ידע זה צפוי להעשיר ולהרחיב את מסגרות המדע האקדמי (Gura, 2013).<sup>11</sup>

### השתתפות הוגנת ביצור ידע מדעי

לדברי סטון (Stohn, 2015, p. 10), מטרת כלי הגילוי הוא לאפשר לספרייה לקשר בין המשתמשים, מטרות וצרכי המחקר שלהם, ובין המידע הזמין, כך שהמגוון ועושר הנתונים הופך למשמעותי. מטרות אלו יוצאות מנקודת פתיחה שאיכות אחזור המבוססת רלוונטיות היא הכלי המרכזי, וכן שאין הכרח בשימוש במקורות שפיטים (שפיט/לא שפיט כמסן חיפוש אפשרי). במקרים כאלה חומר חדש יקבל עדיפות על פני פרסומים שנתפסים כמושנים. מאמרים שפיטים יופיעו במקום גבוה בתוצאות אם הם רלוונטיים, אך ללא רכישת מאגר ציטוט מספר הציטוטים לא יופיע בתוצאות החיפוש ולמשתמש לא יהיה אינדיקטור של איכות הפריט שאוחזר.

האם יש חשיבות לאיכות, אימפקט ומרכזיות החומר המדעי כפי שהוא בא לביטוי בציטוטים, באחזור מידע? במאמר זה הצגנו את התהליך שבו אובד הפן של מציאת החומרים לפי ההיררכיה הקלאסית-מחקרית, בין השאר בשל הנטל הכלכלי על ספקי המידע ועומס המידע הניתן לסיווג ומיון. החלופות הצומחות בספריות מאפשרות לקוראים ולחוקרים לקבל תוצאות משמעותיות; דווקא באפשרות שהיא פותחת לגישוש ברובדי המידע השונים באופן החופשי ממוקדי המשקל המחקריים, הספרייה מנווטת עבור הקוראים בין החשיפה הישירה לעומס האינפורמציה, תעוקת המידע, לבין הצורך בידע.

כפי שהדגימה מגפת הקורונה, האנושות זקוקה למדע באיכות גבוהה, לפריצות דרך יוצאות דופן ולחוקרים שיובילו אותן. בה בעת הציבור נזקק לכלי יידוע בני-סמכא ובעלי רלוונטיות גבוהה, במיוחד לאור האופי הבעייתי של הדיון והצפת כלי התקשורת ברמות שונות של מידע. מכל האגפים עולה צורך דוחק במידע משמעותי, אמין ונגיש. במאמר סקרנו כיצד מדדים ביבליומטריים נדחקים מלשמש כאינדיקטור לאיכות ומרכזיות מדעית במסגרת הספרייתית, לטובת רלוונטיות של מילות חיפוש, שמיועדות בראש ובראשונה לספק חוויית משתמש טובה. הפן השלילי של עידן המידע הנוכחי מוכר – דלדול האמון בבעלי המקצוע והתגדרות צרכני המידע במקורות מידע מבוצרים. מגדל השן של האוניברסיטה יחד עם מעבדות המחקר של החברות הגדולות מן העבר האחד, והציבור על גווניו מן העבר השני. הצורך בגישור של רלוונטיות המידע גדול מתמיד, ואף מעבר לכך, נחוצה פתיחת הגשר בין מחפשי המידע ליוצריו.

<sup>11</sup> ראו לדוגמה <https://scistarter.org/citizen-science>

ניתן להמשיל את הפן של כלי האחזור על בסיס רלוונטיות כביטוי לדמוקרטיזציה של המדע ואת הביבליומטריקה כמייצגת המסורת של הקידמה המדעית הקלאסית. האם חוויית המשתמש שבאה לביטוי בתוצאות רלוונטיות, והמעורבות של אלגוריתמים באחזור, יעודדו כיוונים חדשים במדע, יצירתיות ומצוינות?

לחלופין יש לתת את הדעת על השיח הציבורי – האם הישענות על הסמכותיות המסורתית, המתוחזקת על ידי הביבליומטריקה, היא המוצא היחיד מן המצב של דיוני סרק בין קבוצות בעלות עמדות בלתי ניתנות לגישור? או שמא דווקא האוטונומיה של המחפשים ב"מגרש משוטח" של תוצאות, ללא העדפה ביבליומטרית על פני רלוונטיות סמנטית, ואף השפעתם על גופי הידע הזמינים, היא שתייצר יותר אמון בידע המצטבר?

### סיכום ומחשבה להמשך

ב"ספרייה של בבל" (בורחס, 1998) בורחס מצביע על האתגר העומד בפנינו בעידן המידע הנוכחי, עם הבדל מרכזי בין הסיפור לבין האפשרות העכשווית למימוש, שבסופו של חשבון מעצים את הלקחים מן הספרייה של בבל. ההבדל הוא יכולת אחזור המידע, החיפוש. האפשרות למצוא כל רצף אותיות (או שילוב אחר באמצעים הולכים ומשתכללים) במהירות גבוהה פותרת לכאורה את הבעיות ואת האופי המאתגר של הספרייה בסיפור. במקום ספרנים מקומיים ונוודים, תסכול מתמשך ובעיית המקוריות – יש צורך "רק" במנגנון אחזור המידע. יש המבטאים גרסה זו בתיאור של הספרייה כולה כספר בודד הנמצא בסופרפוזיציה קוונטית, או במימוש דיגיטלי של המספר העצום של פרמוטציות "בענן".<sup>12</sup> דווקא ניסיונות אלה ממחישים את המסר העולה מן הסיפור של בורחס ובעידן הנוכחי – תהליך הכילוי והיצירה לבודים זה בזה.

מסדי הנתונים שפועלים בלי המטריקות של ציטוטים ויחסים מרושתים בין-טקסטואליים, מחזירים אותנו לכאורה צעד אחורה – אך לעיתים יש צורך לקחת צעד אחורה כדי לראות ולהעריך את הנוף.

מהסקירה הראשונית שהובאה במאמר זה עולה כי בעוד שהמרכזיות של הכלים הביבליומטריים – לצורכי יעילות אחזור של פריטי מחקר מרכזיים, למיפוי דיסציפלינה, לצורכי קידום ומימון – הולכת ומתעצמת, הם נדחקים מתוך תוצאות החיפוש בכלי הגילוי המרכזיים לטובת רלוונטיות וחוויית משתמש.

במאמר זה אין בכוונתנו להשוות בין שני מנגנוני גילוי, אך ברצוננו לפתוח דיון ולעודד מחקר עתידי בנוגע למשמעות של גילוי מבוסס רלוונטיות ללא שימוש באינדיקטורים ביבליומטריים.

<sup>12</sup> בשנת 2015 הקים הסטודנט לספרות ג'ונתן בסיל (Jonathan Basile) את האתר "הספרייה של בבל", שבו ניתן לגלוש בחוויה וירטואלית של הספרייה של בורחס. מובן שאין ברשותו של בסיל חוות שרתים שמסוגלת לאכלס את כמות המידע האמיתית של הספרייה מן הסיפור, ומדובר בעצם במנוע פרמוטציות שמייצר לבקשת הגולש דגימות מן הספרייה. ראו <https://libraryofbabel.info>

האם שיטת אחזור זו תביא לפריצת חסמי ידע ותאפשר גילוי וקידום חוקרים שאינם במרכז העשייה המדעית? האם הדמוקרטיזציה של הידע וגיוון סוגיו ("תקשורת מדעית") תסלול אפיקי ידע חדשים?

## מקורות

בורחס, ח"ל. (1998). הספרייה של בבל. בתוך גן השבילים המתפצלים (י' ברונובסקי, מתרגם). הוצאת הקיבוץ המאוחד. עמ' 74-77.

יאיר, ג', גואטה, נ', ודודוביץ, נ' (2018). חוק המצוינות הנדירה: הפיריון המדעי של הזוכים בפרס ישראל במדעי החיים ובמדעים המדויקים. סוגיות חברתיות בישראל, 25, 6-27.

Bar-Ilan, J. (2008). Informetrics at the beginning of the 21st century—A review. *Journal of Informetrics*, 2(1), 1–52.

<https://doi.org/10.1016/j.joi.2007.11.001>

Belter, C. W. (2015). Bibliometric indicators: opportunities and limits. *Journal of the Medical Library Association : JMLA*, 103(4), 219.

<https://doi.org/10.3163/1536-5050.103.4.014>

Bornmann, L., & Daniel, H. D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), 45–80. <https://doi.org/10.1108/00220410810844150>

Cleverdon, C. W. (1964). Citation Indexing. *Nature*, 203(4944), 446–446.

<https://doi.org/10.1038/203446b0>

De Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2016). Evaluation practices and effects of indicator use—a literature review. In *Research Evaluation* (Vol. 25, Issue 2, pp. 161–169).

<https://doi.org/10.1093/reseval/rvv038>

Deutsche Forschungsgemeinschaft. (2010). "Quality not Quantity" – DFG Adopts Rules to Counter the Flood of Publications in Research (Issue 7, pp. 9–10).

EBSCO. (2015). *How is relevance ranking determined in EBSCO Discovery Service (EDS)?* EBSCO. [https://connect.ebsco.com/s/article/How-is-relevance-ranking-determined-in-EBSCO-Discovery-Service-EDS?language=en\\_US](https://connect.ebsco.com/s/article/How-is-relevance-ranking-determined-in-EBSCO-Discovery-Service-EDS?language=en_US)

- EBSCO. (2020). *Learning to Search/Searching to Learn: The Evolution of EBSCO Discovery Service™*.  
<https://www.ebsco.com/blogs/ebscopost/learning-searchsearching-learn-evolution-ebsco-discovery-servicetm>
- EBSCO Industries. (2020). *What content is included in my EBSCO Discovery Service (EDS) profile?* [https://connect.ebsco.com/s/article/What-content-is-included-in-the-different-EBSCO-Discovery-Service-Source-Types?language=en\\_US](https://connect.ebsco.com/s/article/What-content-is-included-in-the-different-EBSCO-Discovery-Service-Source-Types?language=en_US)
- ExLibris. (2011). *Ex Libris Primo Reaches a Major Milestone as Its Customer Base Exceeds 750 Institutions - Ex Libris*. <http://exlibris.eoi.co.il/press-release/ex-libris-primo-reaches-a-major-milestone-as-its-customer-base-exceeds-750-institutions/>
- ExLibris. (n.d.). *Resource Types in CDI*.
- Garfield, E. (1955). Citation indexes for science. *Science*, 122(3159), 108–111. <https://doi.org/10.1126/science.122.3159.108>
- Garfield, E. (1971). Citation indexing, historio-bibliography, and the sociology of science. *Currents Contents*, 1(5273), 158–172.
- Garfield, E. (1980). Bradford's Law and Related Statistical Patterns. *Current Contents*, 19, 5–12.
- Garfield, E., Sher, I., & Torpie, R. (1964). *The Use Of Citation Data In Writing The History Of Science* (Issue May).
- González-Alcaide, G. (2021). Bibliometric studies outside the information science and library science field: Uncontainable or uncontrollable? *Scientometrics*, 126(8), 6837–6870. <https://doi.org/10.1007/s11192-021-04061-3>
- Gura, T. Citizen science: Amateur experts. *Nature* 496, 259–261 (2013).  
<https://doi.org/10.1038/nj7444-259a>
- Haggerty, K. C., & Scott, R. E. (2019). Do, or Do Not, Make Them Think?: A Usability Study of an Academic Library Search Box. *Journal of Web Librarianship*, 13(4), 296–310.  
<https://doi.org/10.1080/19322909.2019.1684223>

- Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2013). Coverage and adoption of altmetrics sources in the bibliometric community. *Proceedings of ISSI 2013 - 14th International Society of Scientometrics and Informetrics Conference*, 1(2), 468–483. <https://doi.org/10.1007/S11192-013-1221-3/FIGURES/8>
- Jappe, A., Pithan, D., & Heinze, T. (2018). Does bibliometric research confer legitimacy to research assessment practice? A sociological study of reputational control, 1972-2016. *PloS One*, 13(6), e0199031. <https://doi.org/10.1371/journal.pone.0199031>
- Lemke, S., Mazarakis, A., & Peters, I. (2021). Conjoint analysis of researchers' hidden preferences for bibliometrics, altmetrics, and usage metrics. *Journal of the Association for Information Science and Technology*, 72(6), 777–792. <https://doi.org/10.1002/asi.24445>
- Lown, C., Sierra, T., & Boyer, J. (2013). How users search the library from a single search box. In *College and Research Libraries* (Vol. 74, Issue 3, pp. 227–241). Association of College and Research Libraries. <https://doi.org/10.5860/crl-321>
- Merton, R. K. (1968). The Matthew Effect in Science. *Science*, 159(3810), 56–63. doi: 10.1126/science.159.3810.56
- Mejia, C., Wu, M., Zhang, Y., & Kajikawa, Y. (2021). Exploring Topics in Bibliometric Research Through Citation Networks and Semantic Analysis. *Frontiers in Research Metrics and Analytics*, 6, 62. <https://doi.org/10.3389/frma.2021.742311>
- McBurney, M. K., & Novak, P. L. (2002). What is bibliometrics and why should you care? *IEEE International Professional Communication Conference*, 108–114. <https://doi.org/10.1109/ipcc.2002.1049094>
- Perez, O., Bar-Ilan, J., Cohen, R., & Schreiber, N. (2019). The Network of Law Reviews: Citation Cartels, Scientific Communities, and Journal Rankings. *The Modern Law Review*, 82(2), 240–268. <https://doi.org/10.1111/1468-2230.12405>

- Price de Solla, D. (1963). *Little science, big science... and beyond*. Columbia University Press. <https://doi.org/10.1007/BF02018109>
- Romanelli, J. P., Gonçalves, M. C. P., de Abreu Pestana, L. F., Soares, J. A. H., Boschi, R. S., & Andrade, D. F. (2021). Four challenges when conducting bibliometric reviews and how to deal with them. In *Environmental Science and Pollution Research* (Vol. 28, Issue 43, pp. 60448–60458). <https://doi.org/10.1007/s11356-021-16420-x>
- Sonawane, C. S. (2017). Library discovery system: An integrated approach to resource discovery. *Informatics Studies*, 4(3), 27–38. <http://eprints.rclis.org/32499/1/54-244-1-PB.pdf>
- Stohn, C. (2015). *How Do Users Search and Discover? Findings from Ex Libris User Research* (Issue May).
- Szomszor, M., Adams, J., Fry, R., Gebert, C., Pendlebury, D. A., Potter, R. W. K., & Rogers, G. (2021). Interpreting Bibliometric Data. *Frontiers in Research Metrics and Analytics*, 5, 30. <https://doi.org/10.3389/frma.2020.628703>
- Tabacaru, S. (2019). *Web of Science versus Scopus: Journal Coverage Overlap Analysis Texas A&M University Libraries*. 1–7.
- Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Addison-Wesley Press, Inc. <https://doi.org/10.1525/aa.1950.52.2.02a00290>