

# ניתוח דפוס ציטוטים במערכת המשפט הישראלית

מגישים :

גלעד X

ספיר X

נתנאל X

שי X

## תוכן עניינים

3	הקדמה
3	מבוא
4	סקירת ספרות
7	שאלות מחקר והשערות
7	מתודולוגיה
9	חלק ראשון – רשת שופטים וציטוטים
9	עיבוד הדאטה וייצוג הרשת
10	ניתוח הרשת
10	דרגות
11	בדיקת השופטים המדורגים גבוה
12	צפיפות
12	Betweenness centrality
12	Closeness centrality
13	קהילות ברשת
15	חלק שני – רשת פסיקות
15	עיבוד הדאטה וייצוג הרשת
15	ניתוח הרשת
15	האם הרשת מציגה מאפיינים של רשת עולם קטן
16	מדדים ומידת התאמתם למטרתנו
16	מדדים ומשמעותם
16	סקירת מדגמית של התיקים המובילים
17	סיכום ומסקנות
18	הפניות

## **הקדמה**

בעבודה זו, מטבעו הרב תחומי של הקורס וחברתם של סטודנטים ממגוון תארים שונים לצוות אחד, עלה במחשבתנו הרעיון לנתח בכלים ממוחשבים וגרפיים – רשתיים, תחומים ספציפיים במערכת המשפטית בישראל. מעבר לשילוב הרב תחומי הייחודי של משפט ורשתות בה עוסקת עבודתנו, שאבנו השראה מהאקטואליות הרבה והעיסוק הבלתי נדלה (לטוב ולרע) במערכת המשפט בכלל, ובפרט מרוח התקופה האחרונה, בה קולמוסים רבים נשברו תוך עיסוק עקבי ומתמשך במערכת המשפט. כוונתנו היא לראות לאלו מסקנות ניתן להגיע באמצעות ניתוח רשתי, ולבחון האם באמצעות הרשת ניתן לערער על תפיסות ואמונות רווחות לגבי מערכת המשפט.

בשעה בה כותרות העיתונים אינן חדלות מלזעוק ולהתריע על מערכת משפטית שקורסת תחת עומס של אינספור תיקים ונתונה (בעודנו כותבים עבודה זו) תחת לחצים פוליטיים כבירים, נידונה בתקשורת כמוטה ע"י דעות פוליטיות או לפחות ככזו המועדת בעתיד להטיה פוליטית זו (לדוגמה, "הרפורמה המשפטית" בדגש על התיקון המוצע לחוק יסוד השפיטה וחוק בתי המשפט הצפויים לשנות באופן דרמטי את אופן בחירת השופטים בוועדה לבחירת שופטים). לאור האמור, בעידן בו נראה לכאורה כי מקצועיות השופט מפנה את מקומה לטובת אינטרסים פוליטיים, רצינו לבחון סוגייה אקטואלית זו ולראות האם וכיצד ניתן למדוד או לכל הפחות להוסיף כלי מדידה אל ארגז הכלים הנתון בידי הגורמים הבוחרים את השופטים כך שיוכלו לכמת את ביצועיו, קרי את מקצועיותו, איכותו מידת ההערכה של חבריו, באופן ממוחשב ואובייקטיבי הנקי מאינטרסים פוליטיים. תקוותנו היא שעבודתנו תתרום ותסייע לאומדן אובייקטיבי של השופט כמוזכר לעיל כאשר ייתכן ואף תסייע בהצבעה על הליכי חקיקה נדרשים ותקל על העומס בבתי המשפט.

בנוסף להיותה של העבודה אקטואלית ולטעמנו גם מאוד חשובה כפי שפירטנו לעיל, ואף על פי שעבור רובנו נושא העבודה, קרי מערכת המשפט, רחוקה מאוד מתחום עיסוקנו ומהבנתנו, היא חשובה ורלוונטית לכולנו, שכן במידה מסוימת כולנו "משתמשים" במערכת המשפט, על כן רצינו לבחון מספר השערות שמעניינות אותנו לעניין האיזונים במערכת, צורת העבודה שלה, דרכים לאומדן מקצועיותם של שופטים וחשיבותם של תיקים משפטיים שונים.

לטעמנו הצורך במחקר זה, הוא צורך של המערכת ואזרחי ישראל. המוטיבציה למחקר היא הרצון לעזור ולסייע בכלים שברשותנו גם לענפי ידע אחרים שאולי לא היינו נתקלים בהם מחוץ לאקדמיה. כאשר התמריץ, הוא הידיעה שכולנו יחד, סטודנטים מתארים שונים, תורמים כל אחד כמיטב יכולתנו את הידע והכישורים שלמדנו עד כה לטובת עבודה זו.

העבודה תובא בסדר הבא: ראשית, בשל נקודת המוצא כי איננו משפטנים, נעמוד תחילה על מספר מושגים החשובים לעבודה זו, תוך סקירת ספרות משפטית-מחקרית בינלאומית עם מחקרים הדומים למחקר נשוא עבודתנו, נמשיך בהצגת מגוון שאלות מחקר והשערות ולאחר מכן נסביר בקצרה על יצירת מאגר הנתונים ועיבודו. שנית, בחלק הראשון של עבודתנו נעמוד על רשת שופטים וציטוטים תוך הסבר מילולי וגרפי על עיבוד הנתונים וייצוג הרשת. שלישית, בחלק השני נציג רשת פסיקות, עיבוד הנתונים וייצוג הרשת. רביעית, נסכם את העבודה ונציג מספר מסקנות.

## **מבוא**

כפי שצינו לעיל, נתחיל בהצגה קצרה של מספר מושגים משפטיים החשובים לעבודה זו:

- **ערכאות:** בישראל ישנן שלוש ערכאות, השלום, המחוזי והעליון.
- **תקדים מחייב:** פסק דין שמחייב בתי משפט לפסוק לפיו. בישראל פסיקה של בית המשפט העליון מחייבת את המחוזי והשלום, ופסיקה של המחוזי מנחה את בית המשפט השלום.

- **"ציטוט":** אמירה של שופט שנאמרת בפסק דין. מעבר לתקדים, בפסקי דין רבים יש אמירות של שופטים, ייתכן ומדובר באוביטר (הערת אגב), וייתכן ומדובר בציטוט דברי שופט שקיבלו גושפנקא של פסק דין. בליטיגציה (מקצוע הטיעון) לדוגמה יש חשיבות במציאת ציטוטים של שופטים שניתן "להלביש" על מקרה נתון בהתאם לאינטרס המשפטי של הלקוח.
- **שיטת המשפט בישראל:** נשענת לפי סדר על: חקיקה, פירוש החקיקה ע"י בתי המשפט ופסיקתו, תקדים מחייב מבוסס ערכאות (כפי שצוין לעיל תחת "תקדים מחייב") ופסיקת שופט ללא תפקיד אקטיבי במציאת העובדות אלא קבלתן מהצדדים במשפט. כמו כן ישנה התחשבות גם במשפט העברי ומספר חוקים שקדמו לחוק הישראלי.

מכיוון שבחרנו ליצור מאגר נתונים משלנו למחקר, היה עלינו להבין על מה נוכל להתבסס עבור יצירת רשת בעלת משמעות שתוכל לענות על השאלות שלנו.

לאחר קריאת מגוון מאמרים בנושא, בחרנו להסתמך על ציטוטים ותקדימים משפטיים כמדד למקצועיות ומוערכות דברי שופטים ולמידת הקונצנזוס של תקדים נתון.

כשחשבנו על נושא המחקר, אחת הסיבות בגינה סברנו שיש להתבסס על ציטוטים ותקדימים כמדד ל"טיב" השופט, היא שלדעתנו עצם ההתייחסות של שופט - לשופט אחר או לתוצאה של שיפוט, היא כמעין הבעת דעה על אותה פסיקה, ועל כן תייצג במידת מה את היכולת של השופט הראשון לחשוב ולבחון את הדקויות וההבדלים שבין המקרים, מה שלדעתנו מציג יכולת שיפוט גבוהה יותר. לשמחתנו כבר במאמר [4] נתקלנו במשפט מתוך מאמר [9] אשר טוען (באנגלית) כי: "כל הפניה לתקדים המוכלת בדעה, היא במהותה שיפוט כמוס/מוסתר/מאוחר על התקדים המצוטט". בהמשך המאמר אף נכתב: "בעצם, העובדה כי שופט בחר להקדיש את הזמן לציטוט תקדים מהווה שיפוט על איכות התיק המצוטט" [9].

לדעתנו, כאשר שופט רואה מקרה דומה שנפסק (שאינו בעל "תקדים מחייב"), הוא יכול לבחור להסיק ממנו או לבחור שלא, כאשר אם יבחר השופט להסיק ממנו ואף לאזכר חלקים רלוונטיים מפסק הדין, **משמע שהשופט המאזכר התרשם מנכונות דברי השופט המאזכר ואף טרח לאזכרו בפסק דינו.**

סבורנו כי התרשמות זו כאשר תגיע במספרים משמעותיים ועקביים על שופט נתון, תהווה אינדיקציה למקצועיותו של שופט ולעובדה שהוא מוערך יותר על ידי חבריו.

לעניין דירוג השופטים, בעקבות מספר מקרים בשנים האחרונות, נראה כי חלה פוליטיזציה של הועדה לבחירת שופטים והקריטריונים על בסיסם נבחרים שופטים הופכים ליותר פוליטיים, נשיאת העליון אף שיגרה מכתב בנושא, ובפרשת מין תמורת מינוי נחשף כי יו"ר לשכת עורכי הדין יפעל לקידום שופטת באמצעים פסולים. נראה נכון לומר כי אף על פי שיש מגוון קריטריונים לבחירת שופטים, אין מדד אובייקטיבי לבחירת שופט על בסיס מקצועיותו גרידא, כאשר לטעמנו בהינתן מדד שכזה יהיה קשה יותר לגורמים שונים להעדיף שופט שפחות ראוי על פי המדד האובייקטיבי, או לכל הפחות יאלצו אותם גורמים לנמק מדוע על אף שהמדד האובייקטיבי צידד במועמד X הם ביכרו את מועמד Y.

## **סקירת ספרות**

בחיפושנו אחר מחקרים דומים, נתקלנו במספר מאמרים אשר עושים שימוש בכלים דומים במדינות ומערכות משפט אחרות אשר סיפקו לנו נקודות למחשבה וכיוונים נוספים לבדיקה, כמו כן מסקנות שנוכל להתבסס עליהן בבחירת המדדים לניתוח והסקת המסקנות שלנו.

[1] Analysis of a German Legal Citation Network

במאמר זה מציגים הכותבים את התפתחותה של רשת הציטוטים הגרמנית, ומגיעים למסקנה המתבקשת כי אכן קיים קושי רב לאסוף מידע ולקשר בין ציטוטים שונים בשל העובדה שחוקים

ואמירות בפסקי דין לא בנקל ברות זיהוי. (פרק 3 במקור) אך על אף הקושי, ניתן להגיע מרשת הציטוטים למסקנות משמעותיות בדבר התנהגותה וטבעה של המערכת המשפטית. בפרק 5 במאמר מראים הכותבים כי רשת ציטוטים של בית המשפט העליון בארה"ב היא "Scale-free" ונראה כי אותה התנהגות מתקיימת גם בפסיקות בגרמניה, כאשר יש "Cluster" מאוד מצומצם של החלטות משמעותיות, כשרוב ההחלטות אינן ברות משקל רב. יפים דבריו של כותב המאמר בהמשך פרק 5 כי ציטוטים הם מדד למצוינות במגוון פרקטיות (Using citation counts as a measure for excellence is a common practice in many areas, including scientific research.)

[2] Emergence of network effects and predictability in the judicial system  
במאמר זה סבורים המחברים כי רשת של ציטוטים בין פסיקות בית משפט תעזור בהבנת המנגנון שמעצב את המערכת המשפטית, כאשר במאמר זה עושים שימוש בציטוטי תיקים ע"י שופטים ומראים כי לאורך זמן ניתן לצפות במגמת דיוק עולה את התגברות הציטוטים בין תיקים שונים ואף לחזות ציטוטים קיימים בתיקים עתידיים.

[3] Citation Analysis and Beyond: in Search of Indicators Measuring Case Law Importance

במאמר זה טוענים כי ההתפתחות הטכנולוגית והנגשת פסיקות בתי המשפט אל ציבור העוסקים במשפט אומנם מבורכת, אך נוצרה בעיית עומס של כמות המידע כאשר כלי חיפוש אינם מצליחים לעמוד בקצב גידול המידע המהיר ולאחר את המידע הרלוונטי. כותבי המאמר סבורים כי הפתרון לבעיה נעוץ בבניית רשת ציטוטים על בסיסה נקבל אינדיקציה לחשיבותן של הפסיקות השונות.

[4] Examining The Evolution Of Legal Precedent Through Citation Network Analysis  
במאמר זה נאמר כי ציטוט פסיקה ע"י שופט היא אינדיקציה טובה לכך שפסיקה זו מהווה תקדים עבור אותו שופט (One way that precedent can be examined is through citations in written judicial opinions. This method of analysis operates on the assumption that "[e]ach judicial citation contained in an opinion is essentially a latent judgment about the case cited"[7]. In other words, a citation is a good indication that a cited case is precedent for the case at hand [8]. Based on this assumption, scholars have begun using empirical methods based on the network of legal citations to rank the value of cases [9]).

בנוסף במאמר זה נאמר כי העובדה ששופט טורח לצטט פסיקה של שופט אחר בבואו לתת החלטה, הוא אינדיקציה טובה לכך לאיכות החלטת השופט המצוטט (the fact that a judge has taken the time to include a citation to a particular case is a judgment on the quality of that case). במאמר זה גם ניתן לראות כי אורך פסק הדין הוא אינדיקציה ל Out degree קרי ככל ששופט כותב יותר הוא מצטט יותר ולכן הוא עלול להיות מצוטט יותר, כיוון שייאלצו להתייחס לדבריו בנוגע לפסיקות האחרות. בנוסף בע"מ 254 למאמר מצוין כי יש נטייה לציטוט של שופט להיות מצוטט יותר ויותר ככל שצוטטו אותו יותר שופטים לפניו, שכן שופטים יעדיפו לצטט אמירה שכבר צוטטה ע"י שופטים אחרים (The rich get richer).

[5] CITATIONS IN THE U.S. SUPREME COURT: AN EMPIRICAL STUDY OF THEIR USE AND SIGNIFICANCE

במאמר מדובר על כך שציטוטי שופטים הם לעיתים רחוקות מושא למחקר אמפירי, ולאחר מחקר תיאורטי מגיעים למסקנה כי ניתן להסיק מסקנות ראשוניות בדבר הגורמים המשפיעים על פרקטיקת הציטוטים בבתי המשפט, כאשר נראה ששיעור הציטוטים הלך והתגבר עם השנים והבחירה האם לצטט והרצון להיות מצוטט משפיעים דרמטית על לגיטימיות השופט, אך הציטוטים כשלעצמם לא מספרים את הסיפור המלא. במאמר דובר על כך שציטוטים מהווים את ה"מטבע" של המערכת המשפטית (Citations function something like the currency of the legal system). בע"מ 552 למאמר דובר על כך שרשת ציטוטים משמעותית יותר ככל שיותר שופטים/מאמרים ציטטו את הציטוט הנתון על כן ניתוח רשת ציטוטים מהווה כלי יקר ערך לניתוח וצפיית תקדימים עתידיים. בע"מ 530 למאמר ניתן לראות רשימת פסקי הדין המצוטטים ביותר, כאשר ישנה הפתעה שרשימת פסקי הדין אומנם כוללת פסקי דין משפיעים אך לא בהכרח את אלו שייחשבו ע"י המשפטנים כפסקי הדין המשפיעים ביותר, את ההסבר לדבר מוצאים כותבי המאמר בכך שהרשימה הורכבה דווקא מפסקי דין שדחו תקדימים קיימים ומשכך צוטטו הרבה כדי להעניק להם לגיטימציה.

TABLE 1  
SUPREME COURT CASES CITING THE MOST PRECEDENT

Case	Total Cited Cases	Case Topic
<i>Miller Bros. Co. v. Maryland</i> (1954)	193	Due Process
<i>Minnesota Rate Cases</i> (1913)	160	Commerce Clause
<i>Comm'r v. Estate of Church</i> (1949)	150	Taxation
<i>Nebbia v. New York</i> (1934)	139	Due Process
<i>Glidden Co. v. Zdanok</i> (1962)	130	Separation of Powers
<i>Home Building &amp; Loan Ass'n v. Blaisdell</i> (1934)	113	Contract Clause
<i>Louis K. Liggett Co. v. Lee</i> (1933)	112	Equal Protection
<i>Ashwander v. Tennessee Valley Authority</i> (1936)	109	Legislative Power
<i>Joint Anti-Fascist Refugee Comm. v. McGrath</i> (1951)	108	Presidential Power
<i>Landgraf v. USI Film Products</i> (1994)	101	Retroactivity

Bias in Judicial Citations: A Window into the Behavior of Judges? [6]  
במאמר ניתוח ציטוטי שופטים והגיעו למסקנה כי שופטים שמונו ע"י מפלגות שונות נוטים לצטט את השופטים שמונו ע"י הצד הפוליטי שלהם ולא את השופטים שמונו ע"י הצד הפוליטי האחר.

Distinguishing Judges: An Empirical Ranking of Judicial Quality in the United States [7]  
Courts of Appeals

במאמר זה טוענים כי ניתן לעשות ואף עושים שימוש בציטוטי שופטים כמדד לאיכות השופטים לפני קידומם לערכאה עליונה יותר, אך בניגוד לאחרים, כותבי המאמר מפרידים בין שופטים שצוטטו בצורה שלילית לשופטים שצוטטו בצורה חיובית, תוצאות המחקר שונות לחלוטין ממחקריים שלא מפרידים בין ציטוטים חיוביים לשליליים. בע"מ 331 מגיעים מחברי המאמר למסקנה כי ציטוטי שופטים משקפים השפעה, פוריות, איכות ויצירתיות של השופט המצוטט. המאמר מגיע לאחר ניתוח הציטוטים למסקנה כי ישנם מספר שופטים בעלי ביצועים יוצאים מן הכלל ומספר שופטים בעלי ביצועים נמוכים

[8] יורם שחר, מירון גרוס, חנן גולדשמיט. מאה התקדימים המובילים בפסיקת בית המשפט העליון – ניתוח כמותי. מאמר המסכם את מאה התקדימים המצוטטים ביותר עד שנת 2004.

## שאלות מחקר והשערות

### שאלות המחקר:

1. האם כמות הציטוטים של שופט היא מדד למקצועיותו? נרצה לבחון האם הציטוטים הם מוטים - האם יש לשופטים הטיה לציטוט שופטים בעלי "קרבה" גבוהה יותר (מגדרית, מוסד לימודים, בית משפט, תחום)?
2. האם תיקים מצוטטים יותר אכן מוכרים יותר מאחרים לעוסקים במשפט?

### השערות:

1. אנו מניחים ששופטי בית המשפט העליון יהיו מהמצוטטים ביותר, בהתאם לעקרון התקדים המחייב עליו עמדנו לעיל. יחד עם זאת, אנו מניחים שדעות השופטים הושפעו מהמרצים מהם למדו (בעיקר בתואר הראשון) ולכן מצפים שתהיה הטיה בדפוסי הציטוט של שופטים על בסיס מוסד אקדמאי.
2. ברשת התיקים (פסיקות המאזכרות פסיקות, ללא קשר לשופטים), התיקים בעלי הדרגה הנכנסת הגבוהה יהיו תקדימים חשובים ומוכרים.

## מתודולוגיה

### א. שופטים וקו"ח:

את שמותיהם וקורות החיים של השופטים הורדנו מהאתר של הרשות השופטת (<https://judgescv.court.gov.il>), בה קיים עמוד לכל שופט מחולק לפי שנים ותפקידים. עבור כל שופט, פיצלנו את הטקסט שהתקבל לפי תאריכים שנתונים באתר, ובעזרת סינון על ידי שימוש במילים "חשובות" (שמות ערים, מוסדות לימוד וכו') התאמנו לכל שופט את הנתונים הבאים: מוסד לימודי התואר הראשון, ערכאה, תואר אקדמי (ראשון/שני/דוקטורט), מחוז בית המשפט בו הם מכהנים, מגדר וגיל. כ-500 מהשופטים הושמטו בעקבות חוסר נתונים קריטי (כגון: ארכעה לא ידועה) או דפים לא גמורים באתר. המאגר הסופי שהתקבל מכיל 1196 שופטים: 764 שופטים מכהנים, ו-432 שופטים בדימוס.

### ב. פסקי דין:

לצורך בניית מאגר פסקי דין בהתאם לדרישותינו, היה עלינו למצוא פסקי דין של כמה שיותר שופטים המאזכרים פסיקות אחרות. ביצענו חיפושים חוזרים באתרים הבאים: 1. [ruling.co.il](http://ruling.co.il), 2. תקדין, 3. נבו. השאלות התבססו על שם שופט בשילוב עם ביטויים המעידים על ציטוט או אזכור פסיקה או שופט אחר. בתחילה בנינו מאגר של 10 פסיקות לכל שופט, ולאחר מכן הגדלנו את המאגר ב-10 נוספות במידה וקיימות כאלו, וחוזר חלילה. בסופו של התהליך התקבלו 20,000 פסיקות ייחודיות.

### **ג. עיבוד פסקי דין לגרף שופטים:**

הקבצים שהורדו היו בפורמט WEB לא עדכני, ולכן נאלצנו להמירם לקבצים קריאים עבור פייתון, עבור הניתוח.

בנינו מתודה בפייתון אשר מקבלת את הפסיקות כאוגדן של 100 פסיקות בקובץ טקסט, מאתרת את הכותרות וגוף הפסיקה ע"פ מספר פרמטרים לדוגמת חיפוש המילים "כבוד ה-", "איתור" "פסק דין" ובעצם הישענות על דפוסי הכתיבה של הקלדנית בבית המשפט.

לאחר החלוקה, מתודה נוספת מצליבה שמות מתוך מאגר שמות להתאמה מתוך כלל השמות של השופטים. כמו כן, ישנה בדיקה לגבי גילו של השופט שמתקבל בהצלבה, ובמידה והיה צעיר מדי באותה פסיקה, ההתאמה נמחקת.

מכיוון שקיימים מספר שופטים עם שמות דומים שבמקרים שבהם השם יקוצר לאות ראשונה בשם הפרטי ולשם המשפחה (לדוגמה י. לוי כאשר ישנם שלושה כאלה במערכת המשפט) יגרמו לטעות או הכפלה בנתונים, הסרנו חלק מהשמות הכפולים שלא נצפו לעתים קרובות בפסיקות ושינינו את המצליב לעבוד בתצורת יש ספק-אין ספק, שמעלימה התאמות מוזרות או כפולות ולכן קשתות אלו נפסלו.

לאחר שלב ההצלבה הראשוני, מתקיימת הצלבה נוספת לפי השמות בראש ובגוף, ובסופה מתקבלת רשימת קשתות כקובץ מופרד פסיקים, לפי מס' זיהוי השופט המצוטט ומספר זיהוי המצוטט.

הנתונים שהתקבלו בשלב הזה לא היו מספקים, שכן במוצע קיבלנו התאמה וחצי לכל פסיקה, כ-30,000 ציטוטים/קשתות במשקל יחידה.

הוחלט לשנות אסטרטגיה ולבסס את ההתאמה על חיפוש חוזר של פסקי דין שצוטטו, ואיסוף שם השופט מהתיק שנמצא.

בשיטה החדשה, התקבלו לנו 85,000 קשתות במשקל יחידה, עם חזרות, נתון שהסתדר עם ההנחה, שכן ממוצע הציטוטים בכל פסק דין היה 8 ציטוטים. במעבר לגרף, אוחדו משקלי הקשתות והתקבלו 25,660 קשתות במשקלים שונים.

### **ד. עיבוד פסקי הדין לגרף פסקי דין:**

בעקבות הכמות העצומה של פסקי דין שכבר חיפשנו, וכן קבצי "יומן" שנוצרו כחלק מהבקרה של האוטומציות הקודמות, התקבל אצלנו מאגר שכלל כ-100,000 מספרי פסקי דין וכן ההפניות ביניהם.

בשלב זה בחרנו ליצור רשימה של כל התיקים (ע"פ מספר והליך) וכן כל התיקים אשר כל תיק מזכיר, הצלבנו את הרשימות, הסרנו את הלא מאזכרים וקיבלנו כ-150,000 קשתות (עם חזרות) בין כ-70,000 תיקים. במעבר לגרף, איחוד הקשתות בין אותם זוגות תיקים נעשה בסכימה לקשת במשקל הכולל ביניהם.

### **ה. בעיות בנתונים:**

1. מאילוצי זמן ויעילות, מאגר הפסיקות וכתוצאה מכך 2 הגרפים, התקבלו כתוצאה מחיפושים אשר הינם מוטים ע"פ אלגוריתם החיפוש שבאתר, על כן קיבלנו את הפסיקות המאוחרות יותר אשר ענו על השאלתה ולכן הנתונים הם ברובם מהשנים האחרונות פרט לשופטים שעבורם אותן פסיקות אחרונות היו עשורים אחורה.

2. שופטים מסוימים היו חסרים, בעלי שמות בעייתיים או לא נמצאו להם התאמות בחיפושים ולכן הרשתות הן רק מייצגות (כ-66%) ולא שלמות.



3. בגלל גודל הנתונים, לא ניתן היה לבנות "בקרת איכות" הרמטית, ולכן יש צורך להניח כ-10% שגיאה, אך בזכות כמות הנתונים, שגיאה שכזו היא זניחה.

4. ברשת הפסיקות, כל הפסיקות מעבר ל-20,000 המקוריות, הן תוצאות רנדומליות של חיפושים וקבצי מעקב שרשמו מספרי תיקים והפניות מתוכם, על כן יש להתייחס אליהם כדגימה של פסיקות רנדומליות.

#### ו. תוכנות וחבילות:

1. GEPHI.

2. פייתון

3. תוסף Selenium לאוטומציה ודליית מידע מהרשת.

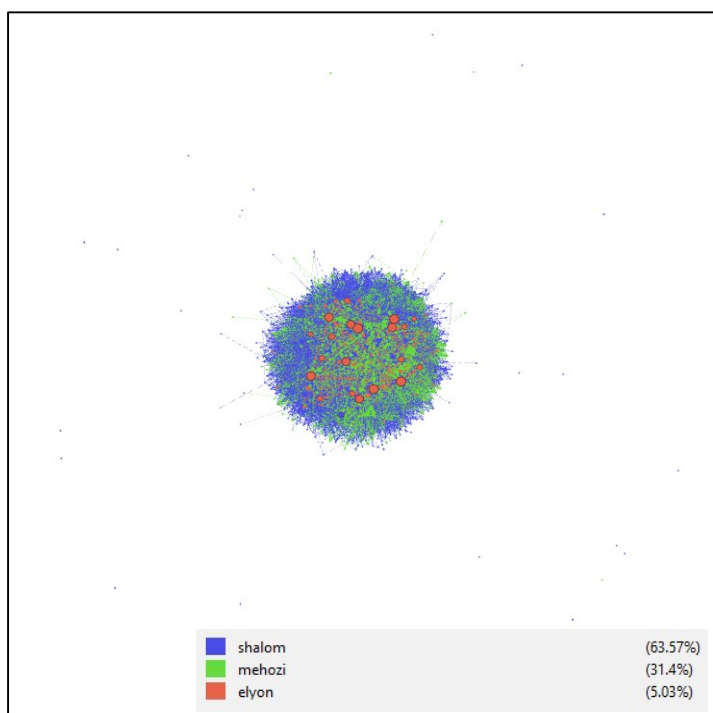
### חלק ראשון – רשת שופטים וציטוטים

#### עיבוד הדאטה וייצוג הרשת

עבור רשת זו השתמשנו בכל השופטים, שופטים מכהנים ושופטים בדימוס. הקודקודים ברשת מייצגים את השופטים, כאשר כל שופט מיוצג באמצעות קודקוד. הקשתות הן מכוונות וממשקולות, אם שופט א' ציטט את שופט ב' X פעמים, יש קשת משופט א' לשופט ב' במשקל X. לאחר עיבוד הנתונים, הרשת מכילה 1,172 קודקודים ו-25,660 קשתות. המטא דאטה עבור כל קודקוד מכיל מוסד לימודים, גיל (70 עבור שופטים בדימוס), מגדר, בית משפט נוכחי-ערכאה ומחוז, ותואר בחלוקה ל – ראשון, שני, דוקטורט.

ייצוג הרשת - לפריסת הרשת השתמשנו באלגוריתם הפריסה Force Atlas 2, עם מניעה של חפיפות. צבעי הקודקודים מסמנים את הערכאה אליה השופט שייך.

גודל הקודקודים פרופורציונלי לדרגה הנכנסת הממושקלת (מספר הפעמים שהשופט צוטט), כך שגודל הקודקודים גדל ככל שדרגת הכניסה הממושקלת שלו יותר גדולה (על מנת שיהיה ניתן לראות בבירור את ההבדלים בגדלי הקודקודים, טווח הגדלים נע בין 5 ל 25). 64 קודקודים עם דרגה 0, קודקודים אלו שייכים לבתי המשפט השלום או המחוזי, כל שופטי בית המשפט העליון בעלי דרגה גדולה מ 0.



ייצוג ויזואלי של הגרף

ניתן לראות מייצוג הרשת, שהשופטים בבית המשפט העליון כולם ציטטו או צוטטו והם הכי מצוטטים, ובפריסת הגרף אין חלוקה לפי ערכאות.

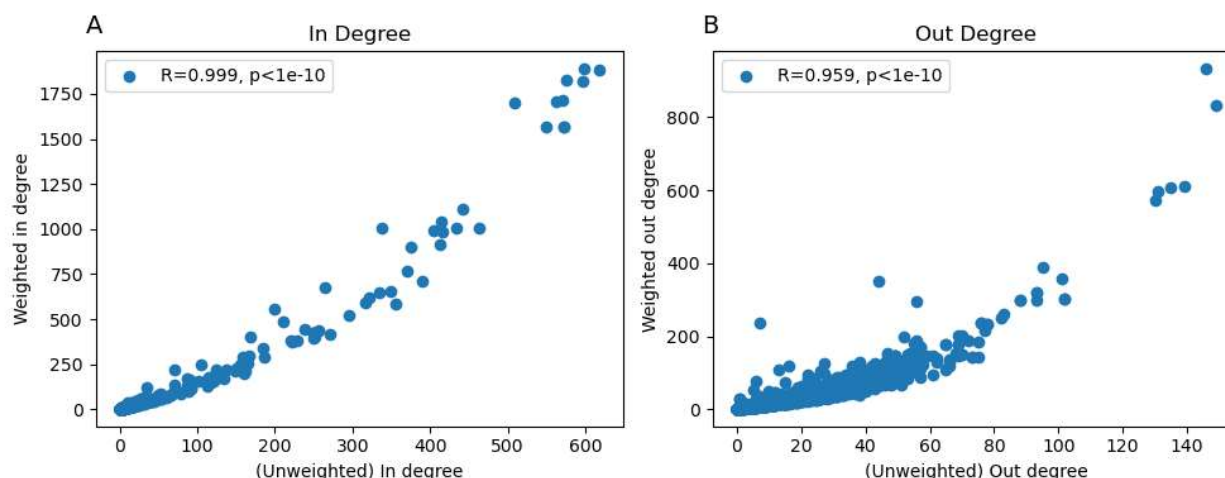
## ניתוח הרשת

**דרגות** - היות והרשת היא מכוונת וממושקלת נתייחס לארבעת סוגי הדרגות: דרגה נכנסת, דרגה יוצאת, דרגה נכנסת ממושקלת ודרגה יוצאת ממושקלת. הדרגה הנכנסת מהווה מדד לכמות השופטים שציטטו את השופט, בעוד שהדרגה הנכנסת הממושקלת מהווה מדד לכמות הציטוטים הכוללת של שופט זה. נרצה לבחון האם שופט הוא מאוד מצוטט כי הרבה שופטים ציטטו אותו, או שמא מעט שופטים מצטטים אותו אבל כל אחד מצוטט אותו מספר רב של פעמים. בדומה, עבור הדרגה היוצאת, דרגה יוצאת מייצגת את כמות השופטים השונים ששופט מסוים ציטט, בעוד שהדרגה היוצאת הממושקלת מייצגת את מספר הפעמים שהשופט ציטט.

ממוצע הדרגות הוא:

דרגה לא ממושקלת	דרגה ממושקלת	
21.9	44.5	ממוצע דרגה נכנסת/יוצאת

נבדוק את הקורלציה בין הדרגה הממושקלת לדרגה הלא ממושקלת. מקדם הקורלציה (R) מופיע ב legend.



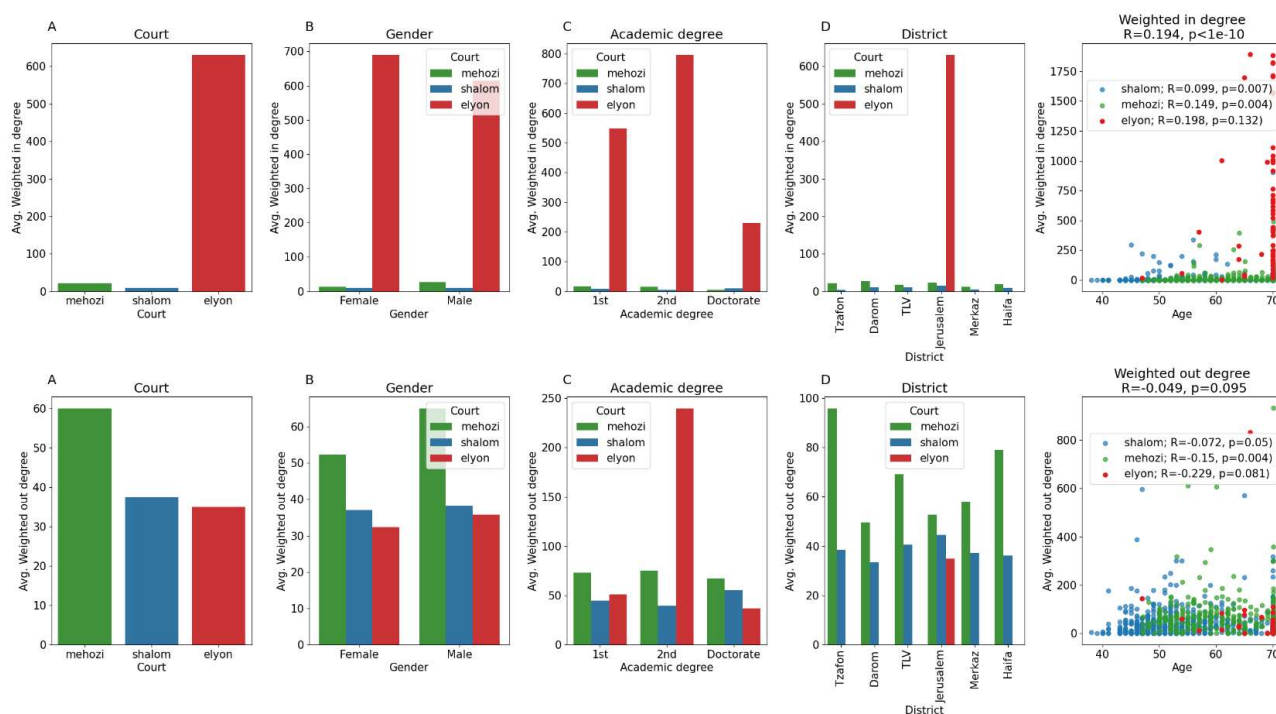
בבדיקה הקורלציה בין הדרגה הממושקלת לדרגה הלא ממושקלת עבור שני סוגי הדרגות, נכנסות ויוצאות, אנו רואים כי אכן יש קורלציה חיובית גבוהה. המסקנה היא ששופט שמצוטט הרבה פעמים, הוא מצוטט גם על ידי מספר רב של שופטים, ושהשופטים לא "מקובעים" על השופט אותו הם מצוטטים (לפי הקורלציה בדרגה היוצאת), ולכן ממצא זה תומך בכך שהערכת שופט לפי הדרגה הנכנסת הממושקלת שלו, היא מדד הוגן להערכת השופטים.

כעת נבחן את הדרגה הממושקלת, היות והיא מסמנת את כמות הציטוטים הכולל, ובעצם בה אנו הכי מתעניינים. כדי לראות מה הגורמים המשפיעים על הדרגה הממושקלת, השונו בין ממוצעי הדרגה הממושקלת הנכנסת/יוצאת. תחילה בין הערכאות השונות, ולאחר מכן בין מאפיינים שונים: מגדר, תואר אקדמי, מחוז וגיל. לבדיקת הקשר בין הגיל לדרגה הממוצעת הממושקלת, בדקנו את הקורלציה בין השניים (מקדם הקורלציה מופיע בכותרת של ה subplot), ובנוסף בדקנו עבור כל אחת מהערכאות (מקדם הקורלציה מופיע ב legend).

דרגה ממושקלת נכנסת - הדרגה הממושקלת הנכנסת הממוצעת של בית המשפט העליון גבוהה באופן מאוד משמעותי מבית משפט המחוזי והשלום. בית המשפט המחוזי מעט יותר מצוטט מבית משפט השלום. בהשוואה של מאפייני השופטים (מגדר, תואר אקדמי ומחוז) אנו רואים שהמאפיינים השונים לא משפיעים על כמות הציטוטים, היות וכל תת קטגוריה מסתדרת עם כך שבית המשפט העליון הוא המצוטט ביותר. אנו גם רואים שאין קורלציה בין הגיל לבין כמות הציטוטים, ניתן להסביר זאת בשל כמות גבוהה של שופטים בדימוס

מעל לגיל פרישה שכבר אינם בפועל בבית המשפט העליון וכן טענה מעניינת ממאמר [4] בעניין חיי המדף של פסיקות, כך שבפועל פסיקות מצוטטות שאינן תקדימים מאוד מוכרים נוטות לדעוך עם הזמן ולהיות מוחלפות בפסיקות שציטטו אותן לקביעת תקדים ממוקד יותר.

דרגה ממושקלת יוצאת – הדרגה הממוצעת הממושקלת היוצאת הגבוהה ביותר היא של בית המשפט המחוזי, לאחר מכן של בית משפט השלום ומעט אחריה בית משפט העליון. סדר זה נשמר גם עבור המאפיינים מגדר ומחוז. בחלוקה לתארים אקדמיים, נראה ששופטים בעלי תואר שני בבית המשפט העליון מצטיינים יותר מאשר שופטים בעלי תואר שני מערכאות נמוכות יותר. גם כאן לא נמצאה קורלציה בין גיל לכמות ציטוטים.



השוואה בין מספר הפעמים ששופט מצוטט(שורה עליונה)/מספר הפעמים ששופט מצטט(שורה תחתונה), בחלוקה לפי ערכאות, ולפי מאפיינים שונים.

**בדיקת השופטים המדורגים גבוה (השופטים המפורסמים מודגשים) - בחלק זה בדקנו מי הם השופטים בעלי הדרגה הנכנסת הגבוה ביותר, מבין שופטי בית המשפט המחוזי שגילם 65 ומטה.**

**רון סוקול:** שמו הועלה כמועמד לעליון  
אילן איטח: לא אוזכר כמועמד לעליון.  
נועה גרוסמן: מוכרת אך לא אוזכרה כמועמדת לעליון.  
**רבקה פרידמן פלדמן:** ראש הרכב השופטים במשפט נתניהו ואולמרט.  
**סיגל דוידוב מוטולה:** בית דין לעבודה, מספר תקדימים בולטים.  
תמר אברהמי: לא נמצאה סיבה מיוחדת.  
רועי פוליאק: בית דין ארצי לעבודה.  
מיכל ברנט: מוכרת אך לא אוזכרה כמועמדת לעליון.  
**משה בר-עם:** לא אוזכר כמועמד לעליון (דן בתיקי נתניהו).  
אביגיל כהן: לא אוזכרה כמועמדת לעליון.

name	Weighted In-Degree	district
RON SOKOL	393.0	tlv
ILAN ITAH	289.0	jerusalem
NOA GROSSMAN	257.0	tlv
RIVKA FRIEDMAN FELDMAN	157.0	jerusalem
SIGAL DAVIDOW MOTOLA	141.0	jerusalem
TAMAR ABRAHAMI	115.0	tlv
ROY POLIAK	76.0	jerusalem
MICHAEL BRANT	65.0	merkaz
MOSHE BAR AM	61.0	jerusalem
AVIGAIL COHEN	60.0	tlv

רשימה של עשרת השופטים מבית המשפט המחוזי, בגילאי 65 ומטה, עם הדרגה הנכנסת הממושקלת הגבוהה ביותר

נציין כי במקום ה 15 (עם דרגת כניסה ממושקלת 36) נמצא השופט גלעד נויטל ששמו הועלה כמועמד לעליון. ניכר כי בראש רשימת הציטוטים ניצבים מספר שופטי מחוזי מוערכים שאף היו מועמדים לעליון, והשופטים שדנו בתיקי נתניהו. אף יחד איתם ברשימה נמצאים מספר שופטים שלא נמצאה סיבה נראית לעין להיותם כה מצוטטים, אחד ההסברים האפשריים לכך – הציטוטים בעניינם לא היו בסנטימנט חיובי או אולי אפילו פסלו את החלטתם.

**צפיפות -** הצפיפות היא מספר הקשרים הקיימים מתוך כמות הקשרים המקסימלית האפשרית בגרף. נבחן את הצפיפות בגרף שלנו, באמצעות מספר הקשרים הקיימים ללא התחשבות במשקל הקשר (משקל הקשת).

כפי שצינו, בגרף יש 1172 קודקודים ( $|V|$ ) ו 25660 קשתות ( $|E|$ ).

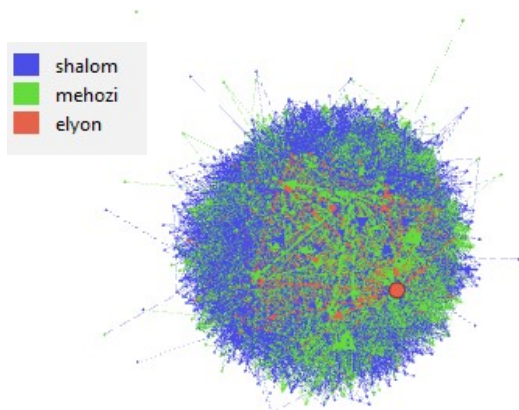
$$Density = \frac{|E|}{|V|(|V|-1)} = \frac{25660}{1172 \cdot 1171} = 0.019$$

צפיפות הגרף מאוד נמוכה, מה שבמקרה שלנו מצביע על כך שבדרך כלל שופט מצטט מעט מאוד שופטים שונים אחרים, דבר המסתדר עם העובדה שבמרבית התיקים אין לשופט סיבה לצטט שופטים מתחומים אחרים (לדוגמה תעבורה ומשפחה לא יצטטו זה את זה רוב הפעמים). נזכיר שמדובר בדאטה חלקי, לכן בהינתן כל הציטוטים לאורך ההיסטוריה, נצפה שהצפיפות תעלה, אם כי עדיין תישאר נמוכה (במצגת שהצגנו בכיתה, היו לנו פחות ציטוטים – 4659 קשתות והצפיפות הייתה 0.002).

**Betweenness centrality (B.C) -** מדד אשר מדרג את מרכזיותה של נקודה על הגרף לפי מספר הדרכים הקצרות ביותר אשר עוברות דרכו. ערכי מדד זה ברשת שלנו נעים בין 0 ל 135,662.62. השופט עם הערך הגבוה ביותר הוא אלכס שטיין, שופט בבית המשפט העליון. אחריו נמצא השופט שי שמואל מבית המשפט השלום בבאר שבע, עם ערך *betweenness centrality* של 35,456.4. שופטים עם ערך גבוה הם המרכזיים ביותר ברשת, ומתווכים בין קבוצות שונות של שופטים מתחומים שונים. סביר להניח שהם צוטטו על ידי שופטים מענפי משפט שונים או בעלי תחומי התמחות מגוונים. מרכזיות זו מעידה על כך ששופטים אלו ממלאים תפקיד משמעותי בחיבור בין חלקים שונים בגוף המשפטי.

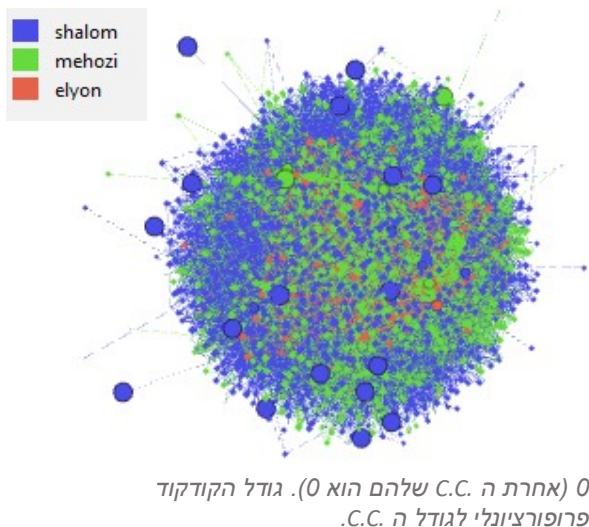
עשרת השופטים עם ה *betweenness centrality* הגבוה ביותר:

Betweenness Centrality	name	court	district	city
135662.620811	ALEX STEIN	elyon	jerusalem	jerusalem
35456.402813	SHAI SHMUEL	shalom	darom	bash
34401.597414	DR AVRAHAM AVRA...	mehoz	tzafon	nazrat
17754.667465	AVICHA DORON	shalom	tlv	tlv
13370.35434	SALIM JOUBRAN	elyon	jerusalem	jerusalem
11651.052018	Yehuda Cohen	elyon	jerusalem	jerusalem
10263.299352	RON SOKOL	mehoz	tlv	rag
9152.344654	Yitzhak Cohen	elyon	jerusalem	jerusalem
8797.23741	SHLOMO LEVIN	elyon	jerusalem	jerusalem
8771.210708	KEREN AZULAI	elyon	jerusalem	jerusalem



בגרף מוצגים רק הקודקודים בעלי דרגה גדולה מ 0 (אחרת ה B.C שלהם הוא 0). גודל הקודקוד פרופורציונלי לגודל ה B.C.

**Closeness centrality (C.C) -** מדד מרכזיות נוסף של קודקוד, לפי מידת קרבתו הממוצעת לכל הקודקודים ברשת. המדד מחושב לפי ממוצע הדרכים הקצרות ביותר מקודקוד לכל קודקוד אחר ברשת, כך שכל שהממוצע קטן יותר, הנקודה יותר מרכזית ברשת. מדד זה מאפשר לראות מי השופטים המשמעותיים ביותר ברשת שאינם בהכרח המרכזיים ביותר. לשופטים בעלי C.C גבוה עשויה להיות השפעה מהירה וישירה יותר על הנוף המשפטי הכולל. דעותיהם ופסיקותיהם עשויות להתפשט במהירות ברשת, ולהשפיע על החלטותיהם ופרשנויות משפטיות של שופטים אחרים מוקדם יותר. ערכי ה C.C ברשת שלנו נעים בין 1 ל 0 (אלו הערכים האפשריים לפי הגדרת המדד). בניגוד ל B.C, כל השופטים עם הערך הגבוה ביותר הם מבית משפט השלום (19 שופטים עם ערך C.C = 1). הערך הכי גדול לאחר 1, הוא 0.56. ערך ה C.C של השופט אלכס שטיין (בעל ערך ה B.C הגבוה ביותר) הוא 0.55, והוא מדורג במקום ה 22 לפי מדד זה.



השופטים עם מדד C.C. = 1

Closeness Centrality	name	court	district	city
1.0	ROY KERET	shalom	darom	bash
1.0	AYELET GOLAN TAVORI	shalom	merkaz	pat
1.0	ARNON KEMELMAN	shalom	tzafon	kiryatshmona
1.0	GIDEON BARAK	shalom	darom	bash
1.0	OPHIR KATAV RIVLIN	shalom	merkaz	natania
1.0	YOSSEF TORRES	shalom	tlv	rag
1.0	AMITSUR EITAM	shalom	jerusalem	jerusalem
1.0	ITTA KATZIR	shalom	heifa	heifa
1.0	TAL SHACHAR	shalom	tlv	tlv
1.0	ADI EISDORFER	shalom	darom	ashdod
1.0	MOSHE PERETZ	shalom	darom	bash
1.0	HILA GERSTL	mehoz	merkaz	lod
1.0	ALIA FOGEL	shalom	tlv	tlv
1.0	RIVA SHARON	shalom	merkaz	pat
1.0	YOCHVED GREENWACD RAND	shalom	merkaz	pat
1.0	EDNA GABRIELI	shalom	tlv	tlv
1.0	SHAUL AVINOR	shalom	tlv	tlv
1.0	EITAN ISRAEL ORENSTEIN	mehoz	darom	bash
1.0	AMIR SHAKED	shalom	jerusalem	jerusalem

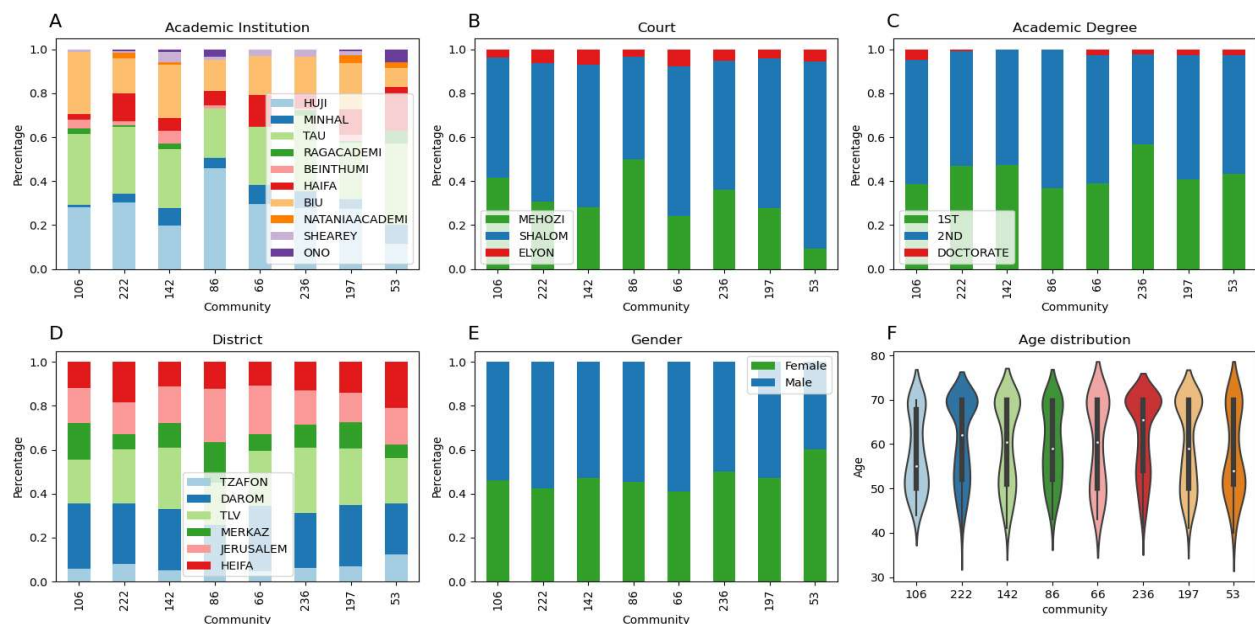
**קהילות ברשת** - השתמשנו באלגוריתם הדיפולטיבי של Gephi עם הרזולוציה הדיפולטיבית 1 לחלוקת הגרף לקהילות. האלגוריתם מחלק את הקודקודים לקהילות לפי ההסתברות של קודקוד להשתייך לקבוצה מסוימת בהתאם לקשתות בגרף.

הרשת התחלקה ל-72 קהילות. היות ו-64 קודקודים הם בודדים, כל אחד מהם שויך לקהילה נפרדת. לכן, הסרנו את הקהילות הללו מהאנליזה, והתייחסנו רק לקהילות שמכילות לפחות שני קודקודים (8 קהילות). גדלי הקהילות הם בין 53 ל-236 קודקודים (ניתן לראות את הגדלים בצירי ה x בתמונה מתחת). בדקנו האם השופטים מתקהלים לפי מאפיינים דומים. עבור גיל בדקנו את התפלגות הגילאים בקהילה בהשוואה לקהילות האחרות. בתמונה המצורפת למטה מימין בשורה התחתונה, ייצגנו התפלגות גילאים של כל קהילה באמצעות Violin plot. ניתן לראות שהתפלגות הגילאים נראית דומה בכל הקהילות. עבור המאפיינים הנוספים, בדקנו איזה אחוז מהווה כל תת קטגוריה מכל קהילה. גם כאן, אנו רואים שכל תת קבוצה מהווה אחוז יחסי מהקהילה באופן יחסי שווה.

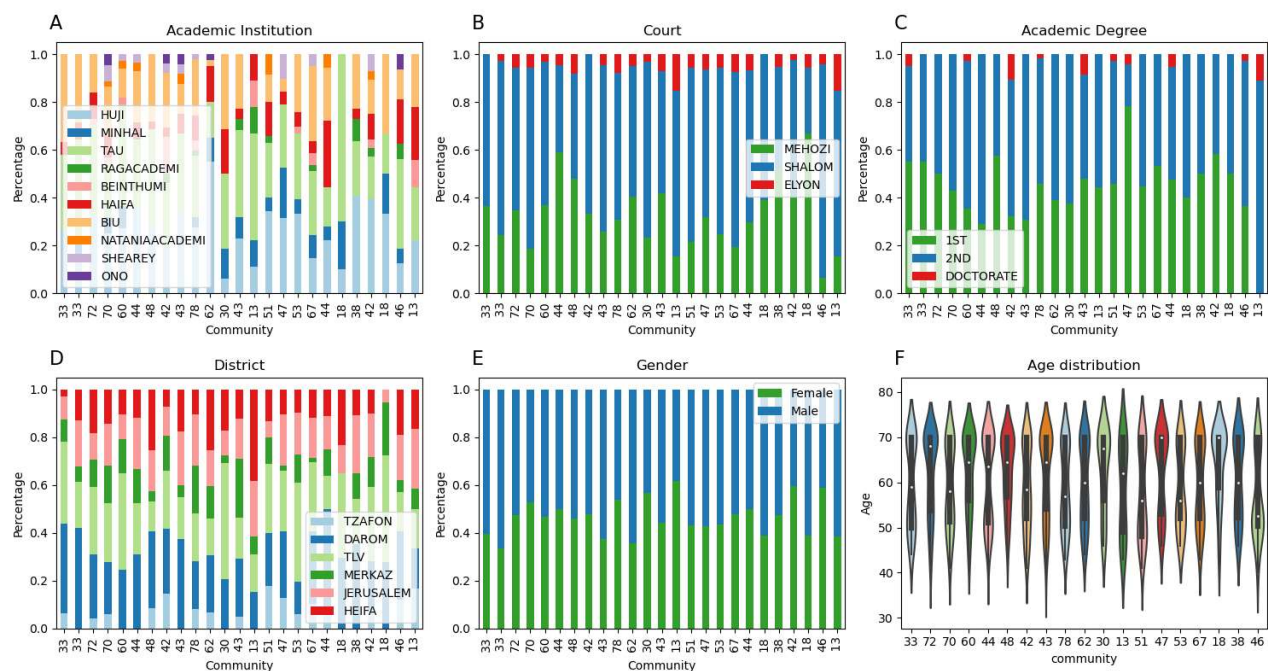
הפעלנו את האלגוריתם למציאת הקהילות עם רזולוציה נמוכה יותר של 0.5, כדי לחלק את הגרף ליותר קהילות. גם פה התקבלו תוצאות דומות.

בניגוד להשערותנו, כי שופטים שלמדו תואר ראשון באותו מוסד אקדמי יצטטו יותר אחד את השני, אנו רואים שאין קשר בין השניים.





חלוקה לקהילות עם רזולוציה 1. הערך בציר ה X הוא גודל הקהילה (אשר משמש גם כ ID של כל קהילה). ב E-A, ציר ה Y מציין את אחוז השופטים בקהילה. F – התפלגות הגילאים בכל קהילה.

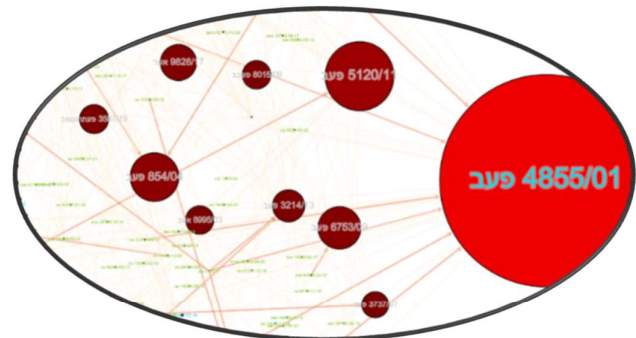
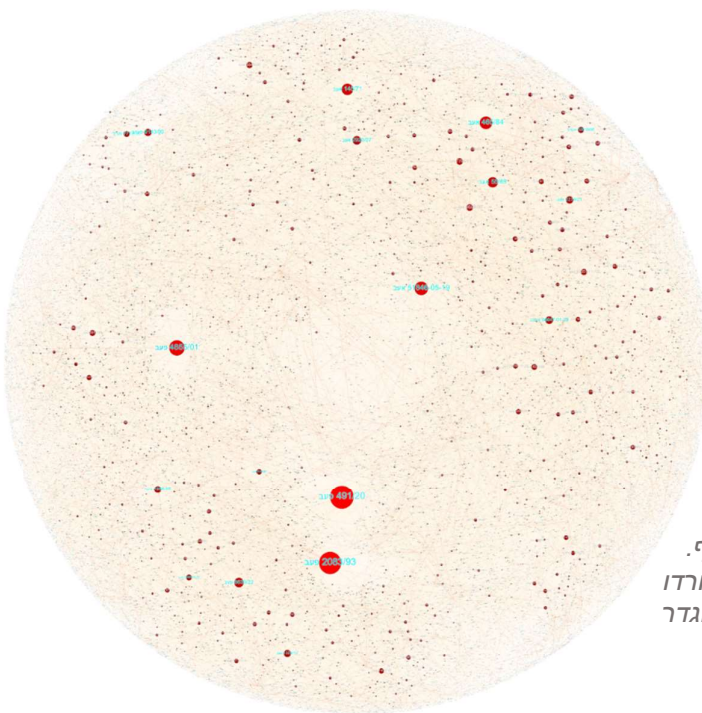


חלוקה לקהילות עם רזולוציה 0.5. הערך בציר ה X הוא גודל הקהילה (אשר משמש גם כ ID של כל קהילה). ב E-A, ציר ה Y מציין את אחוז השופטים בקהילה. F – התפלגות הגילאים בכל קהילה.

## חלק שני – רשת פסיקות

### עיבוד הדאטה וייצוג הרשת

עבור רשת זו השתמשנו בכל הפסיקות מהמאגר הפנימי שלנו אשר מפנות לפסיקה אחרת. הקודקודים ברשת מייצגים את התיקים השונים. הקשתות הן מכוונות וממשקולות ומצביעות לפי כיוון הציטוט. הסרנו תיקים שנראו חשודים, ע"פ מספר ציטוטים לא הגיוני בתוכם (מעל 25 ציטוטים). לאחר הסינון, הרשת מכילה 69,079 קודקודים, ו- 121,672 קשתות ממושקלות ביניהם. ייצוג הרשת- לפריסת הרשת השתמשנו באלגוריתם הפריסה Force Atlas 2 עם כבידה של 100, קנה מידה 100, וההגדרות שמונעות חפיפה. צבע ה label עבור קודקודים שהינם HUB הוא תכלת, אחרת לבן.



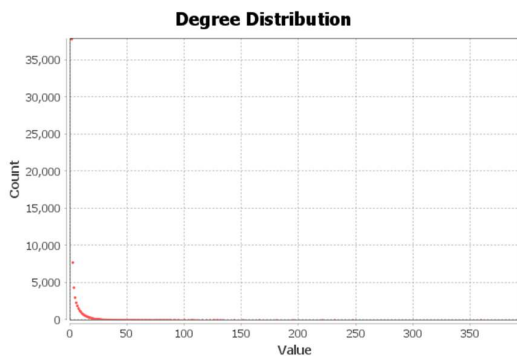
משמאל ייצוג של גרף הפסיקות. מימין Zoom In על חלק קטן מהגרף. ייצוג משקל הקשתות הנכנסות לכל תיק בסרגל צבעים מ לבן-בז' לבורדו ו- 5% העליונים הם מוטים לבורדו בהיר (אדום), וכן גודל הקודקוד מוגדר בהתאמה.

### ניתוח הרשת

#### א. האם הרשת מציגה מאפיינים של רשת עולם קטן?

בדיקת התפלגות הדרגות, כמו גם השינוי בהתפלגות (בהצלבת נתונים חלקיים) מעיד על כך שהרשת היא "סקייל-פרי", בהתאם למה שחשבנו וכן למסקנות ממאמר [1].

לצערנו, מקדם ההתקבצות הממוצע ברשת שלנו עומד על 0.003 שהינו נמוך ביותר. על כן ניאלץ להסיק שהרשת מציגה מצד אחד מאפיין אחד של רשת עולם קטן אך מצד שני ממקדם ההתקבצות היא מציגה מאפיין של רשת רנדומלית, הסיבה לרנדומליות הינה השימוש באוסף אקראי ברובו של נתונים. לצורך בניית רשת טובה יותר יש לבצע איסוף מידע של התיקים ע"י הרחבת מעגלים (שכנים מקרבה מסימת) וחפוש התיקים מהמעגל הראשון וכן הלאה כך שכל פסיקה תיצור סביבה תת רשת ובכך מקדם ההתקבצות יעלה.



## ב. מדדים ומידת התאמתם למטרתנו -

על בסיס המסקנות ממאמר [4] לגבי המדדים הטובים ביותר לבדיקת חשיבות תיק מסוים, השוונו את התיקים המדורגים בצמרת לפי כל אחד מהמדדים. בפועל ובהתאם לתוצאה במאמר (בעניין חוסר היעילות של PAGERANK), כל המדדים סיווגו את אותם תיקים בצמרת, אך PR עיוות מעט את הסדר כך שאינו בהתאם לדרגות נכנסות או נכנסות ממושקלות.

Id	casenum	In-Degree	Degree	Weighted In-De...	Eigenvector Ce...	PageRank
19847	בעא 465/84	220	220	186.0	0.5468	0.000242
17520	בעא 145/71	195	195	149.0	0.490835	0.000208
18645	בעא 2929/07	151	151	119.0	0.377625	0.000149
10215	בעא 9598/05	115	115	101.0	0.285701	0.000139
19921	בעא 475/77	92	92	69.0	0.22978	0.000111
21499	בעא 7649/18	66	66	57.0	0.165445	0.000107
20524	בעא 5777-09-17	53	53	51.0	0.132049	0.000106
30754	ברעא 4249/08	58	58	52.0	0.143166	0.000104
24950	בעפ 2910/90	99	99	68.0	0.248808	0.000102
17434	בעא 1359/20	69	69	59.0	0.172156	0.0001
18638	בעא 29159-11-22	82	82	67.0	0.203632	0.000099
18314	בעא 246/60	87	87	80.0	0.217274	0.000099
21498	בעא 7647-03-20	84	84	68.0	0.209257	0.000096
43976	ועא 64/87	80	80	69.0	0.200084	0.000094
29635	ברעא 10895/08	45	45	43.0	0.112302	0.00009
10637	עא 7426/14	88	88	71.0	0.22105	0.000084
24897	בעפ 28/04	33	33	31.0	0.085132	0.000083
22257	בעא 9653/06	56	56	56.0	0.140679	0.000082
17208	בעא 11039/07	70	70	64.0	0.174012	0.000081
13087	ברעא 3812/91	88	88	66.0	0.21913	0.00008
24074	בעפ 12606-04-19	100	100	59.0	0.255639	0.000079
29901	ברעא 1868/16	55	55	44.0	0.137598	0.000077
17032	בעא 10062/04	67	67	49.0	0.167913	0.000075
29967	ברעא 2063/10	49	49	41.0	0.122176	0.000071
22079	בעא 8995/03	35	35	34.0	0.086393	0.000071
32991	ברעפ 3676/15	33	33	30.0	0.081457	0.000069
20644	בעא 6018/03	33	33	35.0	0.082682	0.000067

Id	casenum	In-Degree	Degree	Out-Degree	Weighted In-De...	Eigenvector Ce...
19847	בעא 465/84	220	220	0	186.0	0.5468
17520	בעא 145/71	195	195	0	149.0	0.490835
18645	בעא 2929/07	151	151	0	119.0	0.377625
10215	בעא 9598/05	115	115	0	101.0	0.285701
24074	בעפ 12606-04-19	100	100	0	59.0	0.255639
24950	בעפ 2910/90	99	99	0	68.0	0.248808
19921	בעא 475/77	92	92	0	69.0	0.22978
10637	עא 7426/14	88	88	0	71.0	0.22105
13087	ברעא 3812/91	88	88	0	66.0	0.21913
18314	בעא 246/60	87	87	0	80.0	0.217274
21498	בעא 7647-03-20	84	84	0	68.0	0.209257
18638	בעא 29159-11-22	82	82	0	67.0	0.203632
43976	ועא 64/87	80	80	0	69.0	0.200084
24952	בעפ 2911/11	74	74	0	50.0	0.187023
17208	בעא 11039/07	70	70	0	64.0	0.174012
17434	בעא 1359/20	69	69	0	59.0	0.172156
17032	בעא 10062/04	67	67	0	49.0	0.167913
21499	בעא 7649/18	66	66	0	57.0	0.165445
29992	ברעא 2127/97	61	61	0	47.0	0.156234
32974	ברעפ 3588/16	58	58	0	41.0	0.145004
30754	ברעא 4249/08	58	58	0	52.0	0.143166
22257	בעא 9653/06	56	56	0	56.0	0.140679
29901	ברעא 1868/16	55	55	0	44.0	0.137598
29673	ברעא 11556/05	53	53	0	33.0	0.135875
20524	בעא 5777-09-17	53	53	0	51.0	0.132049
23859	בעפ 103/05	52	52	0	40.0	0.130806
29967	ברעא 2063/10	49	49	0	41.0	0.122176
20986	בעא 663/87	49	49	0	38.0	0.121563

דירוג התיקים לפי PageRank

דירוג התיקים לפי IN DEGREE

## ג. מדדים ומשמעותם -

מכיוון שהגרף נבנה על בסיס כמות ציטוטים בלבד, והינו גרף מכון וממושקל, בעל צפיפות 0~ מפתא גודלו, הצפיפות מעידה על כך שאכן, מרבית הקשרים האפשריים לא קיימים, הרי רוב פסקי הדין לא מצטטים את שאר פסקי הדין. מצד שני נצפו תיקים רבים, שנוטים להתקבץ בקבוצות קטנות של 3-4, בדרך כלל אלו יהיו תיקים של ערעורים אשר מזכירים ומצטטים את ההחלטות שהתקבלו בגלגולים הקודמים של התיק והקבוצות האלו בד"כ מזכירות פסיקה חיצונית (לא בקבוצה) כלשהי בעניין הסיבה לערעור, הזכות לערעור, נסיבות הערעור, כשלים ראייטיים וכדומה, ולכן עדיין חלק מהרשת. כמו שהסקנו בסעיף הקודם, אם הרשת תיבנה מתוך אחד המעגלים וכלפי חוץ, כמות פסקי הדין שיתקבצו יחד תגדל, ונקבל נטייה גבוהה יותר להתקבצות.

## ד. סקירת מדגמית של התיקים המובילים -

Id	casenum	PageRank
24494	בעפ 2083/93	0.000459
25949	בעפ 491/20	0.000413
25915	בעפ 4855/01	0.000341
21713	בעא 8104-11-14	0.000305
10162	בעא 51846-05-19	0.000283
19847	בעא 465/84	0.000257
33884	ברעפ 9140/99	0.000233
10423	לרעפ 7773/16	0.000221
17520	בעא 145/71	0.000215

Id	casenum	Weighted In-De...
25949	בעפ 491/20	403.0
24494	בעפ 2083/93	379.0
25915	בעפ 4855/01	268.0
10162	בעא 51846-05-19	237.0
19847	בעא 465/84	222.0
17520	בעא 145/71	201.0
20365	בעא 55/69	182.0
27547	בעפ 8639/22	167.0
18645	בעא 2929/07	153.0
12640	בבעא 66647-01-20	135.0



ע"פ 491/20 מיכאל מכלוף מיקי אילוז נ' י.א. מרכז מזון כפר תבור בע"מ: מאסר/סנקציה בשל ביזיון בית משפט יוטל רק כאשר ניכר כי המאסר יוביל להשגת המטרה בגינה הוטל המאסר/סנקציה. ע"פ 2083/96 תמר כתב נ' מדינת ישראל: עוסק בשאלה מהם הקריטריונים המאפשרים לעשות שימוש במנגנון החריג של אי הרשעה - "הלכת כתב".

ע"פ 4855/02 מדינת ישראל נ' ד"ר איתמר בורוביץ: ביטול כ"א בשל הגנה מן הצדק במקרים של הפלייה לרעה מול אחרים.

ע"א 465/84 גאזי טרביה נ' אליהו, חברה לביטוח בע"מ: עוסק בחובת הזהירות של המבוטח שהינו מעביד כלפי עובדיו בפוליסה (מילות הפוליסה - 'המבוטח ינקוט באמצעי זהירות מתקבלים על הדעת'), ומגיע למסקנה כי מקרה הנ"ל המבוטח (מעביד) לא חרג מתנאי הפוליסה על אף חוסר הזהירות שבמעשיו.

ע"א 145/71 מרים ורמברנד נ' מנדל אגולסקי: עוסק בצירוף נתבעים בתביעת סעד המופנית נגד שני נתבעים לחלופין. כאשר לא מן הנמנע ששתי תביעותיו יהיו סותרות זו את זו.

ע"א 55/69 קצין התגמולים נ' דורית שפירא: פסק דין העוסק בקשר הנסיבתי שבין הנכות לשירות הצבאי במקרים של תאונות המעלות טענות בדבר פיצוי מצה"ל.

ע"א 2929/07 פסק דין מוכר מהוועדה לאנרגיה אטומית נ' עדנה גיא-ליפל: שם נבחנת הידיעה בפועל והידיעה בכוח של התובע על פי מבחן אובייקטיבי של האדם הסביר והאמצעים הסבירים, תוך התחשבות בנסיבות העניין, להבדיל מתכונותיו ונסיבותיו האישיות של התובע; לעניין טיב הידיעה הנדרשת כדי שמרוץ ההתיישנות יחל, כאשר המבחן הוא קצה חוט או סברה ואם אדם סביר היה מגיש תביעה.

על אף שאין מדובר בתיקים הנחשבים ל"תיקים הכי משפיעים" של בית המשפט, ניכר כי מדובר בתיקים חשובים מאוד שאכן קבעו מסמרות במגוון סוגיות חשובות ופופולריות כגון רגע הידיעה במרוץ ההתיישנות בתביעות נזיקיות, ביטול כתב אישום בשל הגנה מן הצדק, חובת הזהירות של המבוטח וכדומה. כלל סוגיות אלו הן סוגיות רלוונטיות שעולות במגוון תביעות ודיונים משפטיים.

## סיכום ומסקנות

### גרף שופטים:

לשמחתנו, לא מצאנו התפלגות של ציטוטים המעידה על הטיה (bias) של השופטים, לדוגמה מגדר, לימודים משותפים וכדומה. כמו כן, ראינו ששופט שמצוטט הרבה פעמים, הוא מצוטט גם ע"י מספר רב של שופטים וגם מדד זה לא מוטה. אנו מסיקים מכך כי הערכת שופטים לפי כמות הציטוטים שלהם הוא מדד הוגן.

הגרף שנבנה על אף היותו מבוסס על מספר פסיקות חלקי בלבד ולכן לא מייצג את המצב בכללותו, בכל זאת הצליח לדרג את שופטי העליון גבוהה מעל שאר השופטים במדדי מרכזיות שונים שנבדקו, מכאן אנו מסיקים ששימוש בפסיקות כמטריקה לדירוג השופטים היא אכן בכיוון הנכון. לצורך המשך בכיוון זה נידרש לשימוש בכלים חזקים יותר לדוגמה הוספת זיהוי-רגש (SENTIMENT ANALYSIS) ועל בסיסו קשתות ממושקלות חיובית ושלילית, כך שאם שופט הוזכר רבות, אך כל הפעמים שהוזכר בהן היו כי טעה בשיפוט או שחלקו על דעותיו – ידורג נמוך משופטים שצוטטו בחיוב.

### גרף פסיקות:

על אף כל הבעיות שהוזכרו בבניית הגרף, עדיין ניתן היה למצוא תיקים בעלי חשיבות ומשמעות, מצאנו בהתאמה לטבלה במאמר [5] כי הסוגיות המשפטיות במדינתנו, בבסיסן דומות ומתייחסות לאותן אי הבנות ואי דיוקים המקבילים למערכת המשפט האמריקאית, וכן מדינות נוספות. פסקי הדין שנמצאו הכי מצוטטים עסקו בענייני: התיישנות פיצויים, חברות ביטוח, תהליכים משפטיים (גילוי ראיות, אובייקטיביות של צד בהליך, אי הרשעה), חוזים וכדומה.

אמנם הגרף הציג מדדים חלקיים של רשת רנדומלית אך במבט מקרוב ניתן להבחין בקהילות קטנות שאם היה ניתן ולהמשיך לדלות מידע, היו כנראה הופכות לקלאסטרים גדולים לפי נושאים מרכזיים בפסיקה.

בנוסף לכך, רבות מהפסיקות שהוזכרו במאמר מאה הפסיקות המצוטטות [8] נמצאו בגרף שלנו, אמנם עם דרגות נמוכות אך זה מראה על ההתאמה למסקנות מהמאמר בעניין מרכזיותן במערכת המשפט.

כמו כן למדנו רבות על האתגרים שביצירת רשתות מורכבות מהעולם האמיתי, על הקושי בניתוחן ועל התרומה החשובה של ההצגה הגרפית הויזואלית ליצירת שפה משותפת בין אנשים שונים מתחומים שונים.

## **הפניות**

- [1] Milz, T., Granitzer, M., & Mitrovic, J. (2021). Analysis of a German Legal Citation Network. In *KDIR* (pp. 147-154).
- [2] Mones, E., Sapieżyński, P., Thordal, S., Olsen, H. P., & Lehmann, S. (2021). Emergence of network effects and predictability in the judicial system. *Scientific reports*, 11(1), 2740.
- [3] van Opijnen, M. (2012, December). Citation analysis and beyond: in search of indicators measuring case law importance. In *JURIX* (Vol. 250, pp. 95-104).
- [4] Carmichael, I., Wudel, J., Kim, M., & Jushchuk, J. (2017). Examining the evolution of legal precedent through citation network analysis. *NCL Rev.*, 96, 227.
- [5] Cross, F. B., & Spriggs, J. F. (2010). Citations in the US Supreme Court: An empirical study of their use and significance. *U. Ill. L. Rev.*, 489.
- [6] Choi, S. J., & Gulati, G. M. (2008). Bias in Judicial Citations: A Window into the Behavior of Judges?. *The Journal of Legal Studies*, 37(1), 87-129.
- [7] Anderson IV, R. (2011). Distinguishing judges: an empirical ranking of judicial quality in the United States Court of Appeals. *Mo. L. Rev.*, 76, 315.
- [8] יורם שחר, מירון גרוס, חנן גולדשמיט. מאה התקדימים המובילים בפסיקת בית המשפט העליון – ניתוח כמותי.
- [9] Fowler & Jeon, *supra* note 7, at 17.