

# Semantic Network Analysis in Social Sciences.

Elad Segev (Ed.). Routledge, 2022. 248 pages.

שאול שנהב\*

בעשורים האחרונים אנו עדים לשינויים מרחיקי לכת בכל הנוגע לשימוש בטקסטים במדעי החברה. פיתוח מואץ של כלים אוטומטיים מבוססי מחשב לניתוח טקסטים, לצד נגישות למאגרי מידע הכוללים טקסטים ממקורות שונים, בהיקפים אדירים ובפורמטים דיגיטליים, למעשה מגדירים מחדש את אחת הפרקטיקות המסורתיות בחקר החברה - קריאה ופרשנות של טקסטים. תוך זמן קצר יחסית צמחה זירה עשירה, חיונית ומתפתחת של מחקר מבוסס טקסט. בחלק מהמקרים זירה זו מנותקת ממסורות ותיקות של ניתוח טקסטים ומתפתחת מתוך עצמה, ובאחרים יש לה מערכות יחסים עשירות ובעלות ערך רב עם גישות מסורתיות, שצמחו ברובן ממחקר איכותני. ההשלכות של השינויים הללו בולטות לא רק בתחומים שבהם מחקרים אמפיריים מבוססי טקסט היו יחסית נדירים בעבר ועתה מתמקמים בחזית, אלא גם בתחומים כמו חקר התקשורת, שבהם העניין בניתוח טקסטים היה רב מאז ומתמיד. לצד העניין שמתעורר נוכח שימוש במחשבים לניתוח טקסטים והפוטנציאל הגדול הטמון בכך, ישנה גם חשדנות טבעית כלפי מחקרים אלה. זו ניזונה ממקורות שונים, ובהם החשש מכך שמקצת הכלים שבהם נעשה שימוש אינם שקופים דיים עבור קהל הקוראים, וכן עמדה ספקנית עקרונית יותר ביחס למידת המורכבות שמאפשר ניתוח טקסט המתבסס על מחשב ולאובדן המגע ה"אנושי" בניתוחים אלה. מבלי להיכנס לעובי הקורה בוויכוחים חשובים אלה, נראה שעל דבר אחד קשה להתווכח: השילוב בין היצע של מאגרי מידע, הכוללים טקסטים בהיקפים חסרי תקדים, ובין כלים חדשים שמאפשרים לנתח אותם, מספק לחוקרות ולחוקרים הזדמנויות לחלץ תובנות ומסקנות על התנהגות אנושית בהיקפים ובאמצעים שלא היו לנו בעבר. השאלה שעדיין יש לשאול, ושעליה אין, ואולי לא צריך להיות, מענה נחרץ אחד, היא: כיצד עלינו לעשות זאת? כלומר, אילו כלים וגישות יוכלו לסייע לנו לנצל את ההזדמנות וללמוד מהיקפי הטקסט החדשים הזמינים לנו?

\* פרופ' שאול שנהב, המחלקה למדע המדינה, האוניברסיטה העברית בירושלים

shaul.shenhav@mail.huji.ac.il

זהו, בעיניי, ההקשר המתאים לקריאת הספר *Semantic Network Analysis in Social Sciences*, שראה אור לאחרונה, בעריכת פרופ' אלעד שגב מהמחלקה לתקשורת באוניברסיטת תל אביב. שגב, שמתמחה בניתוח רשתות סמנטיות, הוא העורך של הספר, אולם טביעת האצבע המחקרית שלו בולטת לכל אורכו. ההנחה הבסיסית בתחום זה, כפי שמסביר שגב בהקדמה לספר, היא שבחירת מילים וצירופן בטקסט יוצרת את המשמעות הייחודיות של הטקסט. אפשר לפשט את הרעיון של ניתוח רשתות סמנטיות באופן הבא: הטקסט מתורגם לרשת המציגה את המילים הבולטות ואת עוצמת הקשרים ביניהן, קשרים שאותם נהוג להעריך סטטיסטית מתוך קורפוסים גדולים לפי ההיקרות המשותפות (co-occurrences) של המילים ביחידות טקסט ברזולוציות שונות (משפט, מסמך וכו').

הספר נפתח בהקדמה, שבה מציג שגב את המונח ניתוח רשתות סמנטיות ומציב אותו בהקשר היסטורי ותיאורטי. הפרק הראשון, "How to conduct semantic network analysis", הוא בעיניי הפרק המרכזי בספר, לפחות עבור אלה שמבקשים ללמוד ממנו כיצד להשתמש בגישה זו הלכה למעשה. בפרק זה מציג שגב את התחום בשפה בהירה, מזמינה וקולחת, גם מההיבט הטכני וגם מבחינת ההיגיון התיאורטי שבבסיס הגישה. פרק זה, ההקדמה ובמידת מה גם הסיכום, הם עיקר התרומה של הספר עבור מי שישתמש בו כמדריך תיאורטי ומתודולוגי למחקר המתבסס על גישה זו. שגב מציע כאן קווים מנחים לניתוח רשתות וכן כלים עדכניים וזמינים לביצועו. הפרק, כמו הספר כולו, מאפשר לחוקרות ולחוקרים ללא רקע בנושא להבין את העיקרון של הגישה וליישם אותה במחקרים שלהם. שאר הפרקים מציגים שורה של מחקרים בתחום התקשורת שבהם נעשה שימוש בניתוח רשתות סמנטיות. החלק הראשון עוסק בניתוח "מלמעלה למטה", ובמאמרים הנכללים בו מוצגים ניתוחים של אמצעי תקשורת ונאומים פוליטיים. החלק השני מתמקד בניתוח "מלמטה למעלה", והמאמרים המופיעים בו מציעים ניתוח של שיח ברשתות החברתיות. החלק השלישי כולל מאמרים שבהם רשתות סמנטיות משמשות כלי לניתוח מאמרים אקדמיים וראיונות שנערכו במסגרת מחקרים אקדמיים. הפרקים המדגימים את השימוש בגישה פורשים בפני הקוראים את האפשרויות המגוונות הטמונות בה, ומראים שניתוח רשתות סמנטיות לא צריך בהכרח לעמוד במוקד המחקר, אלא יכול לשמש כלי עזר לניתוחים אחרים. למעשה, כפי שמראה הספר, השימוש בגישה זו יכול להיות חלק מתהליך של מחקר-מגשש ולא דווקא שיטת המחקר העיקרית.

כל אחד מהפרקים מסתיים בטיפים ובלקחים שהחוקרים למדו מהשימוש בניתוח הרשתות. יש אומנם מידה מסוימת של חזרתיות במסקנות של המאמרים השונים, ואף על פי כן מדובר בתוספת חשובה ומעוררת עניין, שבזכותה הפרקים הם לא רק הדגמה

של שימוש בניתוח רשתות סמנטיות, אלא חלק מתהליך למידה משמעותי שנכלל בספר. פרק הסיכום מציג את עיקרי הנושאים והפרויקטים שהובאו בספר, ומוסיף כמה תובנות כלליות, בהתבסס על המחקרים שנכללו בו.

הספר מסביר כיצד לבצע ניתוח רשתות סמנטיות בצורה פשוטה וזמינה, שמתאימה גם לחוקרות ולחוקרים ללא ניסיון מוקדם בשימוש בכלים ממוחשבים לחקר טקסט. הרשת הסמנטית משמשת למעשה גם לצורכי ויזואליזציה וגם לצרכים אנליטיים, ככלי אינטואיטיבי שמאפשר להציף ולזהות משמעויות ולנתח אותן. אפשר להציב את הגישה המוצגת בספר ביחס לגישות אחרות, כאשר חלקן פשוטות (כמו ניתוח מבוסס מילונים) וחלקן מתקדמות והוליסטיות יותר ביחסן לטקסט. גישות מתקדמות אלה אינן נשענות באותה מידה על מילים מרכזיות, אלא על מודלים מורכבים ורגישים לשפה, שמאפשרים להבחין בין ניואנסים בטקסט ובמקרים מסוימים אף לחקות היבטים של קריאה אנושית. שגב מתייחס בהקדמה לכמה גישות כאלה, אולם מנקודת מבטי היה אפשר להדגיש אף יותר את התרומה הסגולית של ניתוח רשתות סמנטיות בהשוואה אליהן, תוך התייחסות לנקודות התורפה שלה - וזאת מתוך הבנה שריבוי גישות וכלים יאפשר לנו להתאים את הניתוחים לשאלות המחקר שלנו. בכל זאת, הספר מאפשר לזהות את היתרונות הסגוליים של הגישה ובה בעת להצביע על כמה מחולשותיה. כך, אחד מיתרונותיה הוא שהיא מאפשרת לחוקרות ולחוקרים להיות מעורבים בפרשנות בקלות יחסית בשתי סוגיות מהותיות - הבחירה במילים המרכזיות שיוצרות את הרשתות, וההבנה של הרשתות עצמן. מצד שני, בדיוק בנקודה זו יש סיכון שהכרעות, למשל בסינון או בבחירה של מילים מרכזיות, שנתפסות כ"הגייוניות" או "מתבקשות" ישקפו הטיות שהחוקרים והחוקרות לא בהכרח מודעים להן. כאשר מילים מסוימות מוגדרות כמרכזיות בשדה המחקר (לדוגמה #MeToo בטוויטר, במאמר "Sexual assaults blindsided by politics on Twitter"), או כאשר המילים לקוחות משדה המחקר (לדוגמה, שימוש במילות מפתח במאמרים אקדמיים, שאותן בחרו מחברי המאמרים עצמם, בפרק "Identifying patterns in communication"), קל להצדיק את בחירתן. התמונה הופכת מורכבת יותר כאשר החוקרות והחוקרים מעורבים בקביעת המילים שייכללו ברשתות, אם באמצעות קביעת סף כמותי ואם באמצעות סינון לפי משמעות המילים. מעבר לכך, ניתוח המתבסס על רשתות מילים יהיה מראש רגיש לשאלה עד כמה עולמות התוכן המעניינים אותנו משתקפים במילים חוזרות, ועד כמה גישה כזו מתמודדת עם הבעיה המוכרת של עולמות תוכן חשובים המסתתרים "מתחת לרדאר". עיבוד מקדים (preprocessing) שיוצר האחדה בין מילים דומות או זהות מבחינה סמנטית הוא מוצא חלקי, אך החשש מהתלות של הגישה במילים בולטות עדיין קיים. הצעדים שמציע שגב בפרק הראשון,

המתמקד בשיטות המחקר של הגישה, אומנם לא יכולים לפתור קושי זה במלואו, אך יש בהם כדי לסייע לחוקרים להתייחס להכרעות שקיבלו בשלבי העבודה ולהצדיקן, ובמובן הזה הם בעלי חשיבות.

ככלל, הספר מביא השקפה מחקרית אחראית, שאינה מבקשת "לקדם" או "למכור" את הגישה, אלא להציע אותה, וזאת תוך התייחסות ליתרונותיה אך גם לנקודות התורפה שלה. זהו מהלך נכון, והוא רלוונטי לגישה זו כמו לגישות אחרות בתחום. למעשה, אל לנו לצפות שגישה אחת או קריאה אחת תיתן מענה מוחלט לאתגרים הרבים הכרוכים בנייתו טקסט. אחרי הכול, אפשר לדמות את השפה ואת השימוש בה למארג שבו משובצים יחד הצפוי והבלתי צפוי, הסדר והאי־סדר, הנשלט והבלתי נשלט. ובדיוק בשל כך אנחנו זקוקים לגישות שונות, שיאפשרו לנו לזקק את מה שמעניין אותנו כחוקרים, וככל הנראה להשאיר גם סימני שאלה למחקר הבא. אולי זו הסיבה שניתוח השפה והשימוש בה הוא שדה מחקר מעניין כל כך, ומעודד אותנו, החוקרות והחוקרים, לפתח ולאמץ כלים שונים שיעזרו לנו ללמוד על האופנים שבהם השפה משקפת את החברה ומעצבת אותה.

הספר החדש בעריכת אלעד שגב הוא בעיניי תרומה יפה וחשובה לניסיון הזה. זהו ללא ספק פרויקט שיכול לפתוח לקוראות ולקוראים את הדלת לעולם המרתק של שימוש בכלים אוטומטיים וממוחשבים לחקר טקסט. הוא מסביר בצורה קולחת ומזמינה גישה רבת ערך לניתוח מבוסס מחשב של טקסטים במדעי החברה, מאגד מחקרים מעניינים שעושים בה שימושים שונים, ויכול אף לחשוף את קוראיו לעולם מחקרי מרתק, שמבקש להבין לעומק את תפקידן של מילים כאמצעי לייצוג ולעיצוב המציאות שבה אנחנו חיים.