

עדות מהמכונה: מודל חישובי של הקשבה לעדויות שואה

רננה קידר,
עמית פינצ'בסקי,
מקסים איפרגן
ועמרי אבנד*

בעשורים האחרונים הוקדשו מאמצים עצומים לאיסוף עדויות אחרוני ניצולי השואה שעודם בין החיים. האתגר כיום הוא כיצד להקשיב לכל אותם אלפי סיפורים אנושיים, אשר חרף היותם שמורים בבטחה, עלולים בכל זאת להיעלם בתהום הנשייה. המפנה החומרי בחקר עדויות, אשר בא לידי ביטוי במפגש שבין עדות היחיד לריבוי העדויות המצויות בארכיונים, מעורר אתגרים אתיים וטכנולוגיים: כיצד ניתן לשמר את מעמדם של הסיפור והחוויה הפרטיקולריים בתוך מכלול עצום של עדויות פרטיקולריות אחרות?

במאמר אנו מצייעים מסגרת תאורטית ומודל חישובי שיאפשרו גישה חדשה לעדויות כמכלול. המאמר מציג מהלך חישובי בן ארבעה שלבים, המשלב בין ניתוח כמותני של מכלול העדויות לבין התמקדות בנרטיב ספציפי. המהלך המשולב מתבסס על מתודולוגיה של מידול נושאי (topic modeling) המאפשר לעקוב אחר הרצף הנרטיבי של קורפוס עדויות שואה ולחשוף באמצעותו את המבנה הנרטיבי הטיפוסי של הקורפוס, ועל בסיסו לאתר עדויות החורגות מן המבנה הטיפוסי. פעולת המודל תודגם על בסיס מסד נתונים של 1,000 עדויות מארכיון ה־Shoah Foundation. המטרה אינה זיהוי צורה טיפוסית קבועה, בבחינת צורה מייצגת, אלא זיהוי הטיפוסי כאמצעי להפניית תשומת הלב ליוצא הדופן – אל עדויות שייכתן שלא היו זוכות לתשומת לב בצורה אחרת. כך, הפעלה חזרתית של המודל על פי פרמטרים שונים תיצר בכל פעם תוצאות שונות. לצד תוכנות מחקריות, הגישה המוצעת עשויה לשרת גם שימוש הנצחתי, חינוכי ואוצרותי.

* רננה קידר היא מרצה בכירה בפקולטה למשפטים ובפקולטה למדעי הרוח וראשת המרכז למדעי הרוח הדיגיטליים באוניברסיטה העברית בירושלים.
עמית פינצ'בסקי הוא פרופ' מן המניין וראש המחלקה לתקשורת ועיתונאות באוניברסיטה העברית בירושלים.
מקסים איפרגן הוא תלמיד מוסמך בבית הספר להנדסה ולמדעי המחשב באוניברסיטה העברית בירושלים.
עמרי אבנד הוא פרופ' חבר בבית הספר להנדסה ולמדעי המחשב והחוג למדעי הקוגניציה והמוח באוניברסיטה העברית בירושלים.
המחברים מבקשים להודות למשתתפי הכנס Holocaust Memory in a Digital Age ולמשתתפי הסדנה History, Media and the Digital Age וכן לפרופ' ליאורה בילסקי ולד"ר ענת רוזנברג, לקוראים האנונימיים ולחברי המערכת על הערותיהם המועילות.

א. מבוא

בעשורים האחרונים הוקדשו מאמצים עצומים לאיסוף עדויות אחרוני ניצולי השואה שעודם בין החיים. פטירתם של אחרוני העדים ותחילתו של עידן העדות שלאחר העדים מתרחשים בד בבד עם התפתחויות טכנולוגיות, וכיום בעיקר עם המעבר ממדיה אנלוגיים, דוגמת קולנוע, וידאו וטלוויזיה, אל מדיה דיגיטליים המבוססים על מאגרים עתירי-נתונים (ביג דאטה), גישה מקוונת ורשתות חברתיות. המפגש שבין נתוני עתק ועדויות שואה מעורר אתגרים טכנולוגיים ומוסריים כאחד. מאגרי עדויות מחזיקים מידע, סיפורים וחוויות העשויים להזין מחקרים היסטוריים במגוון נושאים הנוגעים לשואה, ובמקביל לשרת מאמצי זיכרון והנצחה לקראת עידן ללא עדים חיים. אולם ריבוי העדויות המצויות בארכיונים מעלה את השאלה כיצד להתמודד עם הריבוי עצמו: כיצד ניתן לשמר את מעמדם של הסיפור והחוויה הפרטיקולריים בתוך מכלול עצום של עדויות פרטיקולריות אחרות? ובה בעת, האם וכיצד ניתן לתת ביטוי למכלול כולו – לעדות הקולקטיבית – באופן שיכבד את העדות הפרטיקולרית?

מטרת המאמר היא להציג מסגרת תאורטית ומודל חישובי שיאפשרו גישה חדשה לעדויות כמכלול. כיום קיימים כלים חישוביים רבים שבעזרתם ניתן לנתח, לעבד ולסווג טקסטים. בעבר נשמעו ביקורות נגד הניסיון לעשות שימוש בכלים כמותניים לניתוח עדויות שואה, כאשר הטענה העיקרית הייתה שיש בכך רדוקציה של העדים ושל זיכרון השואה לנתונים יבשים, צמצום של סיפורים למספרים. לפי גישה זו, לא מדובר בשאלה מתודולוגית גרידא אלא בכשל מוסרי – ראוי להתייחס לעדות כאל נרטיב אנושי ייחודי שיש לשמר ככזה. אין עוררין שעדויות עשויות להכיל מידע, פרטים ונתונים רבי-ערך על התרחשויות עבר, אך משמעותן וחשיבותן חורג מעבר לכך. עדות היא בראש ובראשונה מסמך מוסרי התובע הקשבה והכרה בסבל ובאובדן של הניצולים והקורבנות. הגישה שאנו מציעים מאמצת את ההתייחסות המוסרית לעדויות, אך כופרת בטענה כי שימוש בכלים כמותניים חישוביים הוא בהכרח רדוקטיבי. למעשה, כלים חישוביים אף עשויים לשרת את ההתייחסות המוסרית לעדות, בתנאי שכלים אלו פותחו במיוחד לטובת מטרה זו.

המודל החישובי שנציג מתייחס לעדות כאירוע שיחני נרטיבי: כהתרחשות מתגלגלת הנפרשת בזמן ונמסרת בשיחה בין עד/ה ומראיינ/ת. עדות היא דיאלוגית, גם אם ברובה צד אחד מספר והאחר מקשיב. ההקשבה מעניקה הכרה לעדות ולעדים, ובהיותה מוקלטת, מזמינה אופנים נוספים של הקשבה והכרה על ידי נמענים עתידיים. התוקף המוסרי של העדות מתכונן אפוא כבר באירוע מסירתה, והוא שריר כל עוד עקרון ההקשבה וההכרה מתקיים. אם לתרגם את הדברים למונחים של מודלים חישוביים, הרי שהכרה בייחודיות של כל עדות, ובחשיבות של פעולת העדות עצמה, מחייבת גישה שהיא מנוגדת לדרך שבה ניגשים בדרך כלל למידע במחקר נתוני עתק (big data) טיפוסיים. עצם ההתייחסות לעדות כאוסף נתונים בלבד היא בעייתית משום שבכך נעקר ממנה הרצף הנרטיבי שבו נמסרה ונשמעה, ובכך נפגמת ההכרה בתוקפה המוסרי. המטרה, אם כן, היא לאלץ את המודל החישובי לפעול בניגוד ללוגיקה המקובלת בתחום: לעקוב אחר הנרטיב העדותי במקום רק לחלץ מתוכו מידע מייצג, "להקשיב" לאופן שבו הוא נמסר ולא רק למה שנאמר בו – ולבצע כל זאת ביחס לעדויות רבות במקביל.

ההתייחסות למכלול עצמו, מעבר לכל עדות בפני עצמה, נוגעת גם היא לשאלה מוסרית. ניתן לראות במאמץ המשותף של מוסדות וארגונים לאסוף, לארכב ולהנגיש עדויות רבות ככל האפשר כניסיון להתמודד עם ההיקף העצום של הטרגדיה האנושית של השואה. התמודדות עם פשע המוני מחייבת עדות המונית. ריבוי העדויות לנוכח פשעי זוועה המוניים ממלא מטרת ראייתיות פורנזיות של הוכחת הפשעים, אך לצד זאת הריבוי מהווה גם כלי חברתי, אפיסטמי, להתמודדות עם היקף הזוועה ולהבנת אופייה.¹ גם אם לעולם לא יהיה ניתן לתפוס את האסון במלואו, עצם המחויבות להמשיך ולאסוף, לשמוע ולהשמיע עדויות, מבקשת לתת מענה – בהכרח חלקי וחסר – לעוצמת האסון. במובן זה, כלים חישוביים עשויים להתאים יותר מהתמקדות בעדויות ספציפיות ובחינתן כמטונימיות לאובדן בכללותו. למעשה, שתי המחויבויות – ההקשבה לעדות היחידה מחד גיסא, וההתייחסות למכלול ככזה מאידך גיסא – משלימות זו את זו. הגישה והמודל שיוצגו להלן מבקשים להעמיד אמצעי למימוש מחויבות כפולה זו, על ידי שילוב בין ניתוח כמותני והתמקדות בנרטיב ספציפי – וברתימת הכוח החישובי להקשבה לעדויות שייתכן שלא היו זוכות לכך בצורה אחרת. כיוון שמדובר במאמץ הקשבה מתמשך, ובתהליך שמעצם טיבו לא יוכל להגיע לעולם לכדי מיצוי, יש לראות את המהלך החישובי כאקט מחזורי שיש להפעילו שוב ושוב על מנת לגלות בכל פעם עדויות חדשות. מעשה החזרתיות מייצר בכל פעם מפגש ייחודי בין עדות מסוימת לבין מכלול העדויות, ובכך יוצר הנגדה פרודוקטיבית שמביאה לידי ביטוי הן את המאפיינים הספציפיים של אותה עדות מסוימת והן את המאפיינים של המכלול.

המאמר יתקדם באופן הבא: בפרק הבא נסקור את הפרדיגמות הרווחות בחקר עדויות שואה ואת האתגרים הנובעים מהמפגש שבין שיטות ניתוח נרטיביות מסורתיות ובין ניתוח עדויות בכלים חישוביים. בהמשך נציג את המסגרת התאורטית ואת מטרת המחקר, שנועד לפתח כלי חישובי המאפשר לעקוב אחר הרצף הנרטיבי של קורפוס עדויות שואה ולחשוף באמצעותו את המבנה הנרטיבי הטיפוסי של הקורפוס, ועל בסיסו לאתר עדויות החורגות מן המבנה הטיפוסי. בפרק שלאחריו נציג את המודל החישובי המבוסס על מידול נושאי של אוסף העדויות שפיתחנו לצורך הניתוח הנרטיבי ואת הקורפוס שאותו ניתחנו. הפרק המרכזי יציג את המהלך המשולב של חישוביות והקשבה הנפרש על פני ארבעה שלבים: (1) זיהוי הנושאים המרכיבים את כלל אוסף העדויות; (2) השוואה בין העדויות השונות על מנת לזהות דפוסים נרטיביים; (3) זיהוי חריגות מהדפוס הנרטיבי הטיפוסי – לשם הדגמה, לפי קטגוריה מגדרית; (4) איתור עדויות חריגות (במקרה זה, מקרב תת-הקבוצות המגדריות). לבסוף, נציג את מסקנותינו וכיווני מחקר עתידיים.

ב. עידן העד(ים): בין קול היחיד למקהלת הרבים בחקר עדויות שואה

הפרדיגמה הרווחת בחקר עדויות שואה בעשורים האחרונים היא של ניתוח ופרשנות של מקרים ספציפיים בכלים ספרותיים. במסגרת זו הציעו חוקרי ספרות ותאוריות ספרותיות תובנות תאורטיות ופרשניות חשובות לניתוח עדויות שואה, בעיקר בהתייחס לעדויות

Renana Keydar, *Mass Atrocity, Mass Testimony, and the Quantitative Turn in* 1
International Law, 53 LAW & SOC'Y. REV. 554 (2019)

וידאו.² תובנות אלה חורגות מעבר לגבולות השיח הספרותי ומזינות דיונים המאתגרים את ההבחנה שבין היסטוריה לזיכרון.³ תובנות ודיונים אלה מאפשרים לבחון שיקולים מוסריים הנוגעים לתהליך איסוף העדויות,⁴ מעוררים אימוץ פרספקטיבות השוואתיות ביחס לשפה, לזמן ולמקום שבו ניתנה העדות,⁵ למדיום העדות עצמו,⁶ ומעוררים שאלות בנוגע להקשר

- 2 LAWRENCE L. LANGER, HOLOCAUST TESTIMONIES: THE RUINS OF MEMORY (1991) (להלן: SHOSHANA FELMAN & DORI LAUB, TESTIMONY: CRISES OF WITNESSING IN ;(LANGER CATHY ;(FELMAN & LAUB :להלן: LITERATURE, PSYCHOANALYSIS, AND HISTORY (1991) CARUTH, UNCLAIMED EXPERIENCE: TRAUMA, NARRATIVE, AND HISTORY (1996); KELLY SHOSHANA FELMAN, ;(OLIVER :להלן: OLIVER, WITNESSING: BEYOND RECOGNITION (2001) THE JURIDICAL UNCONSCIOUS: TRIALS AND TRAUMAS IN THE TWENTIETH CENTURY (2002); Dori Laub, *Testimonies in the Treatment of Genocidal Trauma*, 4 J. APPLIED PSYCHOANAL. Stud. 63 (2002) (להלן: Laub).
- 3 OLIVER, לעיל ה"ש 2; JAMES EDWARD YOUNG, WRITING AND REWRITING THE HOLOCAUST: NARRATIVE AND THE CONSEQUENCES OF INTERPRETATION (1988); GIORGIO AGAMBEN, REMNANTS OF AUSCHWITZ: THE WITNESS AND THE ARCHIVE (1999); Aleida Assmann, *History, Memory, and the Genre of Testimony*, 27 POETICS TODAY 261 (2006); ANNETTE WIEVIORKA, THE ERA OF THE WITNESS (2006). הגישה שאנו נוקטים במאמר זה כלפי עדויות שואה נוטה במובהק לעיסוק בזיכרון השואה והנצחתה ולא דווקא בכל הקשור להפקת נתונים או ידע היסטוריים מתוך העדויות. עם זאת, הרברים אינם מוציאים זה את זה.
- 4 JAMES EDWARD YOUNG, THE TEXTURE OF MEMORY: HOLOCAUST MEMORIALS AND MEANING (1993); DOMINICK LACAPRA, WRITING HISTORY, WRITING TRAUMA (2001); OREN BARUCH STIER, COMMITTED TO MEMORY: CULTURAL MEDIATIONS OF THE HOLOCAUST (1st ed. 2003); NOAH SHENKER, REFRAMING HOLOCAUST TESTIMONY (2015).
- 5 AMOS GOLDBERG, TRAUMA IN FIRST PERSON: DIARY WRITING DURING THE HOLOCAUST (2017); HANNAH POLLIN-GALAY, ECOLOGIES OF WITNESSING: LANGUAGE, PLACE, AND HOLOCAUST TESTIMONY (2018); NICHOLAS CHARE & DOMINIC WILLIAMS, THE AUSCHWITZ SONDERKOMMANDO: TESTIMONIES, HISTORIES, REPRESENTATIONS (2019).
- 6 JEFFREY SHANDLER, WHILE AMERICA WATCHES: TELEVISIONING THE HOLOCAUST (1999); Geoffrey H. Hartman, *Tele-Suffering and Testimony in the Dot Com Era*, in VISUAL CULTURE AND THE HOLOCAUST 111 (Barbie Zelizer ed., 2001); Simone Gigliotti, *Technology, Trauma and Representation: Holocaust Testimony and Videotape*, in TEMPORALITIES, AUTOBIOGRAPHY, AND EVERYDAY LIFE 204 (Jan Campbell & Janet Harbord eds., 2002); Amit Pinchevski, *Archive, Media, Trauma*, in ON MEDIA MEMORY: COLLECTIVE MEMORY IN A NEW MEDIA AGE 253 (Mordechai Neiger, Oren Meyers & Eyal Zandberg eds., 2011); Amit Pinchevski, *The Audiovisual Unconscious: Media and Trauma in the Video Archive for Holocaust Testimonies*, 39 CRIT. INQUIRY 142 (2012); Caroline Wake, *Regarding the Recording: The Viewer of Video Testimony, the Complexity of Copresence and the Possibility of Tertiary Witnessing*, 25 HIST. & MEMORY 111 (2013).

המשפטי של מתן העדות,⁷ ולהיבטים היסטוריים-ביקורתיים נוספים.⁸ יתרה מכך, כפי שמעידים מחקרים רבים שנערכו לאורך השנים בתוך שבין ספרות, זיכרון ופסיכואנליזה, אותם חלקים שנותרו "בלתי מדוברים" (the unsaid) הם אלה שחושפים את הממד הטראומטי שנותר מעבר ליכולת הביטוי של העדים, ודווקא הם בעלי משקל מיוחד בהבנת העדות.⁹ הממד הטראומטי בא לידי ביטוי באופן פרדוקסלי בחוסר היכולת לבטא, ועשוי להגיה, בין השאר, דרך שתיקות, היסוסים, מחוות, וסתירות במהלך מתן העדות. לפיכך נדרשת תשומת לב יתרה לדקויות של התהליך העדותי ולפיתולי הסיפור הנפרש בו. גישה זו מודרכת על ידי מחויבות מוסרית: מעשה העדות מעיד לא רק על מה שקרה אלא גם על אי-האפשרות להעיד באופן מלא על מה שקרה. האימה של השואה תישאר תמיד מעבר להבנה ולתפיסה של נמעני העדות בהווה. כל עדות מכילה בתוכה לא רק סיפור אישי-אנושי אלא גם חסר מובנה, כזה שמסמן – על דרך השלילה – את כל מה שלא ניתן לבטא. גישה פרשנית זו נעה מן הפרט אל הכלל, מהקשבה קרובה לעדויות בודדות ומשם מנסה לרמוז על המון עדויות פוטנציאלי אך כזה שלעולם לא יישמע. זאת כיוון שסוג כזה של "הקשבה קרובה" מוגבל בהכרח לניתוח מספר מצומצם של עדויות. מטרתנו במחקר זה היא לאמץ את המחויבות המוסרית שבהקשבה לעדות אך לנוע בכיוון ההפוך – מן הכלל אל הפרט: קודם כול היררשות לריבוי והקשבה לו – באמצעים חישוביים – ומשם אל העדות הספציפית, אל קולות וסיפורים שלא ניתן היה להגיע אליהם אחרת.

נכון להיום נעשו ניסיונות מספר להרחיב את יריעת הניתוח ולבחון קורפוסים גדולים של עדויות שואה, לרוב באמצעות שיטות כמותניות של תיוג וקידוד טקסטים והמרתם לנתונים שניתן לנתח בשיטות סטטיסטיות. אולם תהליך המרה זה מאבד את המהלך הסיפורי של העדות, את התגלגלותו מעניין אחד לעניין שני ולשלישי, בחזרה לעניין הקודם, וקפיצה לעניין אחר. העדות נפתלת ומהלכה משמעותי לא פחות מהנאמר בה. זאת ועוד, כל אותם היבטים "בלתי מדוברים" בעדות אובדים גם הם בהמרה לקטגוריות שאינן רגישות למהלך הטמפורלי של העדות. כל אלה מציבים אתגר מיוחד בפני ניתוח אוטומטי, ואין זה מן הנמנע שקשיים כגון אלה מנעו פיתוח של שיטות חדשות חרף העושר הרב של כלים דיגיטליים וחישוביים המאפשרים לפרוץ את מגבלות יכולת הקריאה והניתוח האנושיים.¹⁰ חריג לכך

Renana Keydar, *Rethinking Plurality: On Ethics and Storytelling in the Search for Justice*, 1 DIBUR LITER. J., <http://arcade.stanford.edu/dibur/rethinking-plurality-ethics-and-storytelling-search-justice> (last visited Jan. 23, 2024) 7

WULF KANSTEINER, IN PURSUIT OF GERMAN MEMORY: HISTORY, TELEVISION, AND POLITICS AFTER AUSCHWITZ (1st ed. 2006); THOMAS TREZISE, WITNESSING WITNESSING: ON THE RECEPTION OF HOLOCAUST SURVIVOR TESTIMONY (1st ed. 2013); MICHAL GIVONI, THE CARE OF THE WITNESS: A CONTEMPORARY HISTORY OF TESTIMONY IN CRISES (2016) 8

Michael 2; לעיל ה"ש 2, CARUTH; לעיל ה"ש 2, LANGER; לעיל ה"ש 2, FELMAN & LAUB 9
Bernard-Donals & Richard Glejzer, *Between Witness and Testimony: Survivor Narratives and the Shoah*, 27 COLL. LITER. 1 (2000) 2.

Todd Samuel Presner, *The Ethics of the Algorithm: Close and Distant Listening to the Shoah Foundation Visual History Archive*, in PROBING THE ETHICS OF HOLOCAUST CULTURE 175 (Claudio Fogu, Wulf Kansteiner, & Todd Samuel Presner eds., 2016) (להלן): 10

נמצא בעבודתו של פרסנר, שמציע לעשות שימוש בעקרונות של "קריאה רחוקה", מושג שטבע חוקר הספרות פרנקו מורטי,¹¹ לטובת ניתוח חישובי של מספר רב של נרטיבי עדות.¹² בשונה מהקריאה הצמודה המסורתית, אשר בוחנת לעומק מבחר מצומצם של טקסטים נבחרים, קריאה רחוקה ממנפת את היכולות החישוביות על מנת לנתח מספר רב של טקסטים במקביל, ובכך שואפת לזהות דפוסים וקשרים על פני מגוון מסדי נתונים.

פיתוחים עכשוויים בתחום של עיבוד שפה טבעית (natural language processing), תת-תחום בלמידת מכונה ובינה מלאכותית, שעוסק במידול חישובי של השפה האנושית, למידת השפה והבנתה) מאפשרים ניתוח שיטתי של רצפים נרטיביים במקורות טקסטואליים. במסגרת זו נעשו לאחרונה ניסיונות מספר לניתוח נרטיב וזיהוי מושגי יסוד נרטולוגיים כגון עלילה, זיהוי אירועים, ויחסים בין דמויות.¹³

נרטיבים של עדות מציינים אתגר משמעותי לבחינה מחקרית שיטתית, זאת בשל אורכן, מורכבותן, וייחודיות הסוגה. תמלולי שיחות מוקלטות (אודיו או וידאו) בנויים סביב חילופי דברים בין מראיין ומרואיין (הניצול), גם אם רובה הגדול של העדות הוא של האחרון. עדות היא דיאלוג, נרטיב שנוצר במפגש בין שניים בנוכחות אמצעי הקלטה. לפיכך מה שמוקלט אינו סיפור לינארי וכרונולוגי אלא סיפור בעל פה הנע קדימה ואחורה בזמן, ולעיתים קרובות תוך קפיצות אסוציאטיביות בין נושאים, משפטים קטועים (כאשר למשל המרואיין מאבד את חוט המחשבה), פערים בין שאלה לתשובה, תגובות מודעות ושאיןן מודעות. לכך יש להוסיף את הקצב והאינטונציה של הדיבור, ובעדויות וידאו גם את הבעות הפנים ומחוות הגוף. לבסוף, יש לציין את החשש שהסתכלות כוללנית, "חישובית", אשר מבקשת לגת

-
- JEFFREY SHANDLER, HOLOCAUST MEMORY IN THE DIGITAL AGE: SURVIVORS' STORIES; (Presner and New Media Practices (2017); Amit Pinchevski, Transmitted Wounds: Media Renana Keydar, *Listening*; (להלן: PINCHEVSKI) AND THE MEDIATION OF TRAUMA (2019) *from Afar: An Algorithmic Analysis of Testimonies from the International Criminal Courts*, 20 J.L. Tech. & Pol'y 55 (2020); Stephen Naron & Gabor Mihaly Toth, *Let Them Speak: An Effort to Reconnect Communities of Survivors in a Digital Archive*, in MASS VIOLENCE AND MEMORY IN THE DIGITAL AGE: MEMORIALIZATION UNMOORED 71 (Eve Monique Zucker & David J. Simon eds., 2020).
- Franco Moretti, *Conjectures on World Literature*, 1 New Left Rev. 54 (2000); FRANCO 11
MORETTI, *Distant Reading* (2013).
- Presner 12
10. לעיל ה"ש.
- David Bamman, Ted Underwood & Noah A. Smith, *A Bayesian Mixed Effects Model of* 13
Literary Character, in 1 PROCEEDINGS OF THE 52ND ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS 370 (2014); Maria Antoniak, David Mimno & Karen Levy, *Narrative Paths and Negotiation of Power in Birth Stories*, 3 Proc. ACM Hum.-Comput. Interact. 88:1 (2019), <https://dl.acm.org/doi/10.1145/3359190> (last visited Dec. 12, 2020); Matthew Sims, Jong Ho Park & David Bamman, *Literary Event Detection*, in PROCEEDINGS OF THE 57TH ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS 3623 (2019); Andrew Piper, Richard Jean So & David Bamman, *Narrative Theory for Computational Narrative Understanding*, EMNLP 298 (2021).

אל המכלול מעבר לפרטים, עלולה להביא לדה-פרסונליזציה ולמחיקת הייחוד האנושי שמשקף בעדויות.

כאמור, ביסוד הגישה שאנו מציעים נמצאת הטענה כי לא כל התייחסות כוללת היא בהכרח רדוקטיבית מבחינה מוסרית. אדרבה: דווקא התייחסות למכלול עשויה לשרת מחויבות מוסרית מחודשת מתוך ניסיון לכונן עמדת הקשבה שלא ניתן היה לקיימה בעבר, בין השאר משום שלא היו אמצעים טכנולוגיים מתאימים. לפיכך אנו מתעקשים על קיומה של אפשרות שלישית, כזו שמצד אחד איננה מבוססת על התמקדות בעדות ספציפית כמטונימיה (גם אם חלקית) לכלל, ומן הצד השני גם אינה פונה למכלול במקום ובנפרד מעדויות ספציפיות. למעשה מדובר בניסיון לרתום את המתח המובנה בין פרט לכלל על מנת לכונן, באופן מחזורי, בסבב חוזר ונשנה, צורת הקשבה חדשה לעדויות מתוך המכלול ובאמצעותו. לפיכך, המודל שפיתחנו אינו מכוון למיצוע או ליצירת סינתזה אלא לבחינה דינמית ומבוססת-קונטקסט לגבי מהו הסיפור הטיפוסי באוסף עדויות נתון, כאשר המטרה היא לאתר את הסיפורים שהרצף הנרטיבי שלהם חורג מהטיפוסי באותה קטגוריה. המטרה אפוא אינה להגיע למבנה של עדות טיפוסית כשלעצמו, אלא לעשות שימוש במבנה זה על מנת להגיע בכל הרצה של המודל לעדויות חריגות חדשות. מכאן שאין מבנה טיפוסי אחד, אלא כזה שמשתנה בהתאם לפרמטרים המוזנים, ובהתאם תוצאות המידול החישובי שתנינה בכל הרצה ובכל פעם מחדש. המבנה הטיפוסי, אם כן, משמש כנקודת התייחסות ארעית למציאת החריגה ולזיהויו של יוצא הדופן. צורת ההקשבה שנוצרת בתהליך זה היא המשכית, ולמעשה אינסופית. באופן זה, הלוגיקה החישובית והמחויבות המוסרית משלימות זו את זו לטובת הקשבה מחזורית והמשך הגילוי שטמון בריבוי.

ג. מתודולוגיה ודאטה

המודל החישובי שפיתחנו מבוסס על אלגוריתם מידול נושאי (topic modeling). מידול נושאי מתייחס לקבוצה של מודלים סטטיסטיים המאפשרים לזהות בתוך אוסף מסמכים גדול קבוצות מילים אשר מבחינה סטטיסטית נוטות להופיע יחד במסמכים המרכיבים את האוסף.¹⁴ קבוצות או צבירי המילים (clusters) הללו מכונים "נושאים". הנחת העבודה היא שכל מסמך מורכב ממספר נתון של נושאים חבויים (latent) אשר המודל, באמצעות ניתוח סטטיסטי, מסוגל לזהותם. מטרת המודל הסטטיסטי היא כפולה: (1) לזהות את הנושאים החבויים המתבטאים בקורפוס המסמכים, ו-(2) לזהות עבור כל מסמך מהו החוזק שבו כל נושא בא לידי ביטוי (אם בכלל).

לדוגמה, בהיתן אוסף גדול של כתבות עיתונות, אלגוריתם מידול נושאי עשוי לחשוף נושאים שהקורא יזהה כעוסקים ב"פוליטיקה", "טכנולוגיה", "ספורט", "כלכלה" וכו', כאשר כל נושא יכיל סט של מילים בהסתברויות שונות, אשר ביחד מהוות את השדה הסמנטי שמרכיב את אותו נושא. כך, למשל, הנושא "פוליטיקה" עשוי להכיל מילים כמו: "ממשלה", "שר", "כנסת", "אוצר", "ישיבה", "דיון", "ביטחון" וכו', כאשר כל אחת מהמילים תקבל משקל שונה לפי מידת השתייכותה לנושא האמור. בנוסף, עבור כל כתבה באוסף הכתבות המודל

David M. Blei, Andrew Y. Ng & Michael I. Jordan, *Latent Dirichlet Allocation*, 3 J. 14
 (Blei, Ng & Jordan: להלן: MACH. LEARN. RSCH. 993 (2003)).

ינבא מאילו נושאים היא מורכבת. למשל, דיווח חדשותי על אישור תקציב המדינה עשוי להיות מורכב מ-60% נושא "פוליטיקה" ו-40% נושא "כלכלה". המשמעות היא שהדיווח החדשותי האמור מכיל בהסתברות גבוהה את צבירי המילים הקשורים לכל אחד מנושאים אלה. בשונה משיטות חישוביות מונחות (supervised), שבהן החוקרת מגדירה מראש את הנושאים באוסף ואז מאמנת את המודל לחלק את המסמכים בהתאם (תהליך המכונה "קלסיפיקציה"), מידול נושאי הוא מודל בלתי מונחה. המשמעות היא שהאלגוריתם הוא שמזהה את הדפוסים באוסף המסמכים ומקבץ מתוכם צבירי מילים שנמצאו קשורות זו לזו מבחינה סטטיסטית. לאחר שהמודל מייצר את צבירי המילים, ניתן בהליך פרשני המבוסס על היכרות עם הקורפוס והקונטקסט להגדיר כמה אותו נושא עוסק. כך בדוגמה לעיל, הפלט של המודל יהיה צביר מילים במשקלים שונים (למשל "ממשלה", "שר", "כנסת", "אוצר", "ישיבה", "דיון", "ביטחון") והחוקרת היא שתכנה את צביר המילים כנושא "פוליטיקה". מספר הנושאים המבוקש נקבע מראש, וקביעה זו נעשית באופן המשלב היכרות עם הקורפוס ומדר חישובי המנבא את מספר הנושאים האופטימלי עבור הקורפוס הנתון. נבהיר כי כל נושא שהמודל מזהה מבוסס על חישוב התפלגותו על פני כלל אוצר המילים של הקורפוס, כאשר מילים שונות מקבלות משקל שונה בכל נושא. לפיכך, אותה מילה עשויה להופיע בכמה נושאים, אך בכל נושא היא תקבל משקל שונה ותופיע בסביבה של מילים אחרות. מידול נושאי אינו מצריך התערבות אנושית בשלב הראשוני.¹⁵ בהיותו אלגוריתם בלתי מונחה, מידול נושאי הופך למעשה את סדר המחקר האיכותני המקובל: בשלב הראשון מתבצע זיהוי הנושאים המרכיבים את הטקסטים בקורפוס – וזאת מבלי שהחוקרת מגדירה את הקטגוריות הרצויות מראש (אלא רק את מספר הנושאים המבוקש). רק לאחר שהמודל זיהה את הנושאים המרכיבים את הקורפוס, החוקרת ניגשת לפרש את הנושאים ולהציג את התוצאות בהקשרן.¹⁶ באופן זה, המודל החישובי אינו מבוסס על הטיות מחקריות מוקדמות בשאלה מה צפויים להיות הנושאים המרכיבים את הטקסטים. בשונה ממחקר כמותני מסודתי, מודל למידת המכונה אינו אדיש למהלך העדות אלא פועל על הנרטיב עצמו. פלט המודל, אם כן, הוא תוצאה של ניתוח אורגני של השפה והמילים של העדים, ולא חלוקה או סיווג שנכפו מבחוץ.

לצורך המחקר, בנינו מערך נתונים המכיל תמלולים של 1,000 עדויות ניצולי שואה ממאגר העדויות של USC Shoah Foundation (להלן: SF).¹⁷ כל העדויות נערכו בשפה האנגלית, בצורת ראיונות בעל פה עם מראיינים מטעם ה-SF. מספר המילים נע בין 2,609

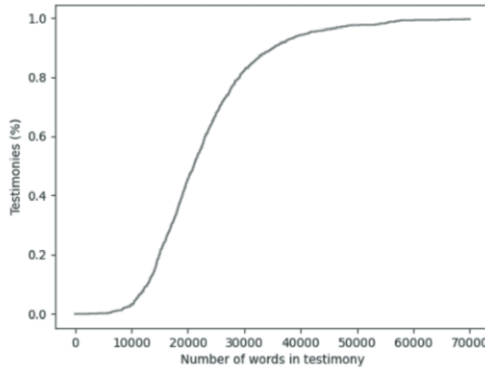
David S. Law, *The Global Language of Human Rights: A Computational Linguistic Analysis*, 12 LAW & ETHICS HUM. RTS. 111 (2018) 15

J. B. Ruhl, John Nay & Jonathan Gilligan, *Topic Modeling the President: Conventional and Computational Methods*, 86 GEO. WASH. L. REV. 1243 (2018) 16

העדויות נבחרו USC SHOAH FOUNDATION, <https://sfi.usc.edu> (last visited Jan. 23, 2024) 17
 בידי ארכיון השואה שמסר לנו את התמלילים. למיטב ידיעתנו, העדויות נבחרו באופן אקראי מתוך אותו מקבץ העדויות שעברו תמלול.

ל-88,105 מילים, כאשר האורך הממוצע של עדות הוא 23,536 מילים, כפי שניתן לראות באיור 1, והאורך החציוני הוא 21,424 מילים.¹⁸

איור 1: התפלגות העדויות באוסף לפי מספר המילים הכולל בכל עדות

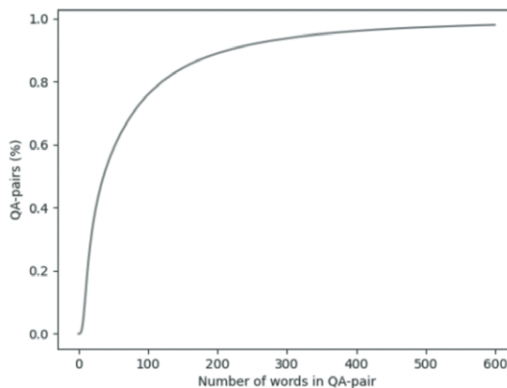


ד. מהכלל אל הפרט וחוזר חלילה – מהלך המידול החישובי

1. שלב ראשון – מידול נושאי של קורפוס עדויות השואה

השלב הראשון, כאמור, הוא מידול אוסף העדויות. על מנת לבצע ניתוח נרטיבי של העדות, ניצלנו את המבנה הפנימי של העדות בהיותה ריאיון בעל פה. כל עדות חולקה לזוגות של שאלה-תשובה, ועל יסוד החלוקה חישבנו את האורך הטיפוסי של צמדי שאלה-תשובה (בין 250-450 מילים), כפי שניתן לראות באיור 2.

איור 2: התפלגות אורכי המסמכים (צמדי שאלות-תשובות) באוסף



18 אורך העדות אינו משפיע על המודל הסטטיסטי, שכן המודל קיבל כנתון את העדויות מחולקות לקטעים שלא בסדר המקורי שבו הופיעו בעדות. המשמעות היא שהמודל אינו "לומד" איזה קטע שייך לאיזו עדות ומה האורך המקורי של כל עדות שממנה נגזרו הקטעים.

כל צמד שאלה-תשובה באורך הטיפוסי מהווה מסמך עבור אימון המודל הנושאי. על יסוד המסמכים שייצרנו, כלומר צמדי שאלה-תשובה, הפעלנו אלגוריתם מידול נושאי המבוסס על למידה עמוקה (Contextualized Topic Models (CTM), שאומן על אוסף המסמכים.¹⁹ לאחר תהליך הרצה לפי מספר משתנה של נושאים, קבענו את מספר הנושאים על 35.²⁰ טבלה 1 מציגה את רשימת הנושאים שייצר המודל, כאשר עבור כל נושא מצוינות 20 המילים הבולטות ביותר מבחינה סטטיסטית באותו נושא.²¹ סדר המילים בכל נושא מבטא את מידת הבולטות של המילה, כלומר עד כמה היא שייכת לשדה הסמנטי שבו היא מופיעה. כפי שניתן לראות, אותה מילה יכולה להופיע בנושאים מספר, ואף במידת בולטות דומה (כאשר בולטות נמדדת תמיד בהשוואה למילים אחרות בתוך אותו נושא). סדר הנושאים [34-0] הוא רנדומלי.

19 CTM מתייחס למשפחה של מודלי שפה גדולים מבוססי רשתות נוירונים (דוגמת BERT). בניגוד למידול נושאי מסורתי (דוגמת LDA), שפועל ביחס לשק המילים המרכיב את הקורפוס (Bag-of-Words), לומד גם מההקשר של המילים, ומכאן שמו: "מודל נושאי מבוסס קונטקסט". הקלט של מודל CTM מורכב מייצוג קונטקסטואלי של המסמך באמצעות מודל RoBERTa וייצוג Bag-of-Words של המילים של המסמך לאחר תהליך עיבוד וניקוי סטנדרטי (pre-processing) שכולל למטיזציה, ניקוי מספרים וסימני פיסוק וסינון מילות עזירה. להרחבה ראו: Federico Bianchi, Silvia Terragni & Dirk Hovy, *Pre-Training Is a Hot Topic: Contextualized Document Embeddings Improve Topic Coherence*, in 2 PROCEEDINGS OF THE 59TH ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS AND THE 11TH INTERNATIONAL JOINT CONFERENCE ON NATURAL LANGUAGE PROCESSING 759 (2021); Federico Bianchi et al., *Cross-Lingual Contextualized Topic Models with Zero-Shot Learning*, in PROCEEDINGS OF THE 16TH CONFERENCE OF THE EUROPEAN CHAPTER OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS 1676 (2021); Jacob Devlin et al., *Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding*, in PROCEEDINGS OF THE 2019 CONFERENCE OF THE NORTH AMERICAN CHAPTER OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS: HUMAN LANGUAGE TECHNOLOGIES 4171 (2018); Blei, Ng & Jordan; (2018), לעיל ה"ש 14.

20 מספר הנושאים נקבע בהליך של ניסוי וטעייה, המבוסס על שיטה מעורבת (mixed method) שמשלבת בין הערכה פרשנית של חוקרת בעלת מומחיות בתחום התוכן שבו עוסק אוסף המסמכים ומדרד עזר חישובי המנבא מהו מספר הנושאים האופטימלי עבור הקורפוס הנתון. לאחר שדגמנו מספר משתנה של נושאים, זיהינו שב-35 נושאים המודל מתייצב ומוציא תוצאות דומות מדי הרצה. לצורך התיקוף, ביצענו את כל המדידות גם על 40 ו-45 נושאים.

21 כפי שציינו, כל נושא מהווה התפלגות על פני כלל אוצר המילים בקורפוס. במילים אחרות, כל נושא מורכב מכלל המילים שיש בכלל העדויות. עם זאת, רק מילים מועטות בולטות מבחינת משקלן, ויתר המילים מופיעות בחלקיקי אחוזים זניחים המהווים את הזנב הארוך של ההתפלגות. מבריקה שערכנו עולה כי לאחר המילה ה-20 משקל המילים יורד ירידה משמעותית (פי שלושה וארבעה) בהשוואה למשקל המילים הראשונות.

**טבלה 1: 35 הנושאים המרכיבים את קורפוס העדויות
(20 המילים הבולטות). סדר הנושאים רנדומלי.**

0	['say', 'know', 'school', 'want', 'tell', 'go', 'mean', 'jewish', 'year', 'think', 'really', 'child', 'talk', 'speak', 'never', 'ask', 'good', 'teacher', 'book', 'come']
1	['person', 'child', 'think', 'family', 'life', 'feel', 'really', 'never', 'happen', 'survive', 'know', 'time', 'even', 'always', 'parent', 'many', 'thing', 'live', 'lot', 'die']
2	['day', 'person', 'camp', 'barrack', 'food', 'time', 'take', 'march', 'bread', 'water', 'soup', 'put', 'little', 'die', 'give', 'body', 'train', 'work', 'eat', 'dead']
3	['russian', 'border', 'bomb', 'run', 'russia', 'river', 'horse', 'train', 'warsaw', 'away', 'forest', 'stay', 'army', 'poland', 'german', 'village', 'soldier', 'start', 'town', 'city']
4	['father', 'go', 'mother', 'leave', 'know', 'come', 'money', 'take', 'train', 'say', 'time', 'parent', 'happen', 'family', 'live', 'person', 'think', 'find', 'home', 'stay']
5	['camp', 'know', 'auschwitz', 'work', 'gas', 'number', 'man', 'barrack', 'prisoner', 'woman', 'person', 'chamber', 'kapo', 'transport', 'mean', 'selection', 'birkenau', 'block', 'mengele', 'see']
6	['school', 'grade', 'teacher', 'public', 'hebrew', 'education', 'teach', 'gymnasium', 'learn', 'class', 'study', 'cheder', 'high', 'class', 'language', 'subject', 'student', 'elementary', 'clas', 'private']
7	['work', 'job', 'new', 'company', 'year', 'college', 'york', 'business', 'marry', 'married', 'university', 'month', 'husband', 'start', 'move', 'study', 'meet', 'become', 'rent', 'money']
8	['wedding', 'son', 'picture', 'daughter', 'david', 'law', 'photograph', 'wife', 'photo', 'right', 'ago', 'crew', 'michael', 'husband', 'peter', 'baby', 'next', 'grandson', 'mine', 'joe']
9	['person', 'state', 'germany', 'israel', 'time', 'year', 'war', 'jew', 'united', 'go', 'country', 'jewish', 'german', 'leave', 'want', 'come', 'camp', 'poland', 'also', 'make']
10	['jewish', 'jew', 'community', 'antisemitism', 'anti', 'population', 'organization', 'synagogue', 'religious', 'zionist', 'belong', 'party', 'hitler', 'gentile', 'orthodox', 'catholic', 'friend', 'polish', 'antisemitic', 'nazi']
11	['name', 'date', 'english', 'interview', 'bear', 'birth', 'spell', 'please', 'today', 'conduct', 'survivor', 'language', 'november', 'state', 'december', 'october', 'united', 'nickname', 'interviewer', 'january']
12	['unit', 'army', 'british', 'military', 'soviet', 'officer', 'italian', 'general', 'force', 'nazi', 'communist', 'war', 'french', 'france', 'join', 'american', 'court', 'german', 'fight', 'germany']
13	['gas', 'auschwitz', 'dog', 'car', 'cattle', 'march', 'see', 'body', 'guard', 'chamber', 'wagon', 'dead', 'shoot', 'hit', 'crematorium', 'throw', 'scream', 'gate', 'walk', 'shower']
14	['bike', 'zloty', 'swedish', 'tram', 'tefillin', 'convent', 'stamp', 'nun', 'hostel', 'arab', 'taxi', 'daddy', 'laundry', 'parcel', 'krakow', 'key', 'cafe', 'bicycle', 'spy', 'split']
15	['father', 'mother', 'sister', 'brother', 'grandfather', 'family', 'old', 'grandmother', 'name', 'bear', 'grandparent', 'die', 'young', 'live', 'parent', 'marry', 'child', 'uncle', 'year', 'aunt']
16	['barrack', 'bunk', 'coffee', 'shoe', 'shift', 'factory', 'block', 'count', 'stone', 'cold', 'cement', 'number', 'soup', 'bed', 'assign', 'bread', 'coal', 'prisoner', 'condition', 'soap']
17	['letter', 'cousin', 'israel', 'write', 'find', 'uncle', 'palestine', 'send', 'meet', 'stay', 'united', 'brother', 'dp', 'aunt', 'relative', 'affidavit', 'visa', 'state', 'paris', 'italy']
18	['say', 'go', 'german', 'come', 'see', 'know', 'soldier', 'american', 'russian', 'gun', 'camp', 'day', 'officer', 'walk', 'tank', 'guy', 'prisoner', 'look', 'take', 'army']
19	['remember', 'think', 'train', 'must', 'sort', 'really', 'station', 'recall', 'obviously', 'journey', 'suitcase', 'trip', 'long', 'ship', 'quite', 'parent', 'holland', 'boat', 'time', 'whether']
20	['come', 'work', 'person', 'use', 'give', 'go', 'want', 'kill', 'russian', 'polish', 'say', 'tell', 'take', 'make', 'jewish', 'jew', 'away', 'bring', 'good', 'know']
21	['school', 'use', 'go', 'jewish', 'home', 'house', 'remember', 'time', 'town', 'little', 'day', 'come', 'live', 'father', 'jew', 'big', 'call', 'year', 'street', 'start']
22	['go', 'come', 'say', 'year', 'sister', 'husband', 'child', 'want', 'brother', 'old', 'take', 'mother', 'see', 'israel', 'stay', 'day', 'daughter', 'family', 'time', 'marry']
23	['room', 'house', 'bedroom', 'apartment', 'live', 'kitchen', 'dining', 'floor', 'flat', 'big', 'building', 'bathroom', 'garden', 'maid', 'living', 'describe', 'large', 'street', 'upstairs', 'small']

24	['holiday', 'shul', 'seder', 'friday', 'synagogue', 'shabbos', 'shabbat', 'pesach', 'passover', 'candle', 'saturday', 'use', 'bar', 'sing', 'mitzvah', 'celebrate', 'religiou', 'kosher', 'always', 'yom']
25	['know', 'little', 'say', 'use', 'eat', 'put', 'give', 'something', 'shoe', 'look', 'make', 'bread', 'water', 'see', 'always', 'nothing', 'everything', 'remember', 'go', 'come']
26	['know', 'german', 'war', 'radio', 'russian', 'hear', 'hitler', 'jew', 'happen', 'poland', 'polish', 'mean', 'germany', 'speak', 'army', 'see', 'start', 'come', 'thing', 'soldier']
27	['child', 'life', 'experience', 'grandchild', 'holocaust', 'love', 'proud', 'message', 'raise', 'wonderful', 'hope', 'son', 'daughter', 'generation', 'future', 'introduce', 'marry', 'survivor', 'affect', 'happy']
28	['say', 'go', 'ticket', 'ship', 'give', 'want', 'come', 'money', 'visa', 'paper', 'train', 'tell', 'boat', 'passport', 'hotel', 'station', 'take', 'guy', 'ask', 'office']
29	['picture', 'take', 'mother', 'right', 'sister', 'photograph', 'father', 'brother', 'leave', 'grandmother', 'grandfather', 'photo', 'uncle', 'side', 'next', 'aunt', 'old', 'thank', 'name', 'row']
30	['say', 'tell', 'go', 'come', 'kill', 'brother', 'run', 'see', 'sister', 'want', 'away', 'mother', 'look', 'house', 'door', 'hide', 'let', 'ghetto', 'father', 'hear']
31	['god', 'feel', 'message', 'question', 'holocaust', 'think', 'world', 'answer', 'future', 'human', 'never', 'experience', 'affect', 'important', 'generation', 'way', 'believe', 'dream', 'happen', 'talk']
32	['ghetto', 'person', 'police', 'judenrat', 'jew', 'star', 'factory', 'work', 'liquidate', 'place', 'hide', 'german', 'street', 'gestapo', 'yellow', 'escape', 'labor', 'policeman', 'area', 'deportation']
33	['go', 'person', 'take', 'come', 'see', 'mother', 'ghetto', 'say', 'know', 'walk', 'child', 'day', 'father', 'place', 'happen', 'german', 'house', 'remember', 'leave', 'run']
34	['food', 'buy', 'money', 'egg', 'meat', 'bread', 'sell', 'flour', 'eat', 'ration', 'potato', 'milk', 'sugar', 'market', 'cook', 'chicken', 'make', 'use', 'soup', 'piece']

בחינה של צבירי המילים, כלומר הנושאים, מלמדת כי המילים בכל נושא אינן קשורות בהכרח זו לזו מבחינה סמנטית. כלומר, לא מדובר במונחים אשר באופן טבעי שייכים לאותו תחום משמעות (כמו רשימת צבעים, פירות או חיות). עם זאת, המודל מצביע על כך שמילים אלו שייכות לאותו שדה מבחינת סיכוייהן להופיע יחדיו באותו מסמך. המטרה שלנו בתהליך הפרשנות של התוצר הסטטיסטי של המודל היא לחלץ את המשמעות החיצונית, בהקשר של נושא השואה, של קבוצות מילים אלו על סמך ידע בתחום. הקריאה בנושאים מציירת תמונה מוכרת של זיכרון השואה: מושגים כגון "kapo", "judenrat" או "gestapo" לצד מושגים כמו "potato", "gas", או "star" אשר משמעותם מתגלה מן ההקשר הרחב של הקורפוס, כמו גם מן המילים שבסביבתן. או כמו במקרה של המילים "sister", "brother" או "mother" בנושא 30, אשר יוצרות ביהד שדה סמנטי של משפחה, אך נצבעות בצבע שונה לחלוטין כאשר הן נקראות לצד מילים כמו "kill" ו"run", אשר המודל מזהה שנוטות להופיע ביחד איתן במסמכי הקורפוס.

פירוש ושיום של הנושאים מצביעים על תמות מוכרות בעדויות שואה: נושא 32 עוסק במציאות החיים בגטו, נושאים 2 ו-16 כוללים מילים המתייחסות לחיים במחנה ריכוז, נושא 34 עוסק בסחר במוצרי מזון בגטו או במחנות, ונושא 13 עוסק בהשמדת היהודים. לצד אלה, ניתן למצוא גם נושאים הקשורים לחיי דת וקהילה, כמו למשל נושא 6 שעוסק בחינוך היהודי, נושא 10 שמורכב ממילים הנוגעות לארגונים יהודיים דתיים, או נושא 24 שכולל מושגים הקשורים למנהגי הדת היהודיים. בנוסף, מספר נושאים עוסקים במלחמה עצמה, דוגמת נושא 26 המורכב ממילים הקשורות לפרוץ המלחמה והשמעות עליה, נושא 3 שעוסק בהפגזות ובבריחה, ונושאים 12 ו-18 שעוסקים בחיילים וביחידות הצבאות השונים.

לצד נושאים אלה, ניתן לזהות נושאים הקשורים למעשה העדות ולשיחה בין מראיין למרואיין בעת מתן העדות: כך, למשל, נושא 11 כולל מילים הנוגעות להצגה העצמית של העד, ומאזכרות את ההתחלה המקובלת בראיונות של ה-SF, שבה העד מתבקש להציג את עצמו ואת חייו לפני השואה. נושא 31 מהדהד אלמנט נוסף בראיונות העדות של ה-SF, והוא סיום הריאיון בהצגת שאלה לעד בדבר המסר שהוא מעוניין להעביר לדור העתיד.

2. שלב שני – זיהוי הדפוס הנרטיבי הטיפוסי

לאחר בחינת הממצאים בשלב הראשון של זיהוי הנושאים המרכיבים את מכלול אוסף העדויות, עוברים לשלב השני והוא זיהוי דפוסים נרטיביים משותפים לכלל העדויות. מכיוון שכל עדות בקורפוס היא בעלת אורך שונה, ביקשנו ליצור יישור (alignment) בין העדויות, על מנת שנוכל להשוות את רצף הנושאים המרכיב אותן. לצורך כך, חילקנו כל עדות ל-15 מקטעים.²² עבור כל מקטע חישבנו את התפלגות הנושאים עבור אותו מקטע ואת הנושא הבולט ביותר בכל מקטע.²³

22 בבחירה לכמה מקטעים לחלק את העדויות נדרשנו לאזן בין שני גורמים: הראשון, ככל שהרזולוציה גבוהה יותר (קרי מחלקים כל עדות ליותר מקטעים) כך המקטעים הסמוכים זה לזה דומים יותר זה לזה וישנה חזרתיות בנושאים. החזרתיות של אותו סט נושאים דומיננטיים לאורך רצפי מקטעים מקשה על זיהוי של דפוסים והסקת מסקנות בעלות משמעות. השני, ככל שהרזולוציה נמוכה יותר (חלוקה למספר קטן יותר של מקטעים) מתקבל מגוון קטן יותר של נושאים דומיננטיים שונים שהיו עשויים לתרום לניתוח הנרטיבי, כך שהעושר הנושאי שמרכיב את העדות ממוסך בשל תמצות יתר.

23 על מנת לאמוד את התפלגות הנושאים עבור כל מקטע, סכמנו את התפלגויות הנושאים במסמכים שמרכיבים כל מקטע ונרמלנו את התוצאות. לצורך ביסוס ותיקוף בכלים סטטיסטיים, הנחנו שכל אחד מהסגמנטים מתפלג כהתפלגות דיריכלה (מודל פשוט וסטנדרטי עבור התפלגות על פני התפלגויות). אמדנו את הפרמטרים של כל אחת מהתפלגויות הסגמנטים באמצעות אומד נראות מרכיב. הפרמטרים של כל התפלגות דיריכלה מתאימים לכל אחד מן הנושאים, ומבטאים את החוזק של נושא בתוך הקטע (באופן מדויק יותר, התוחלת של ההסתברות של נושא מסוים תהיה בדיוק ערך הפרמטר המתאים לו, מנורמל בסכום ערכי הפרמטרים של ההתפלגות). מדרנו את הקורלציה בין הערכים של פרמטרי הדיריכלה שהתקבלו (מנורמלים ל-1 על כל סגמנט) וערכי טבלה 2, המשקפים חישוב ממוצע פשוט, ומצאנו קורלציה גבוהה ביותר (כ-0.95). תוצאה זו, אם כן, מעניקה תוקף לטבלה 2, ולמעשה מראה שהניתוח היה כמעט זהה לו השתמשנו בערכי פרמטרי דיריכלה.

טבלה 2: משקלו של הנושא הדומיננטי ביותר בכל מקטע עדות, באופן ממוצע על פני כלל העדויות

Topic/Part	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	1.87	2.73	2.92	2.87	2.69	2.58	2.56	2.47	2.47	2.49	2.59	2.69	2.89	3.21	2.9
1	1.67	2.53	2.59	2.46	2.5	2.44	2.46	2.44	2.38	2.31	2.27	2.36	2.52	2.82	3.18
2	0.89	1.35	1.58	1.94	2.4	2.9	3.35	3.98	4.33	4.55	4.44	4.16	3.71	3.08	1.82
3	1.01	1.64	2.31	3.13	3.47	3.35	3.19	3.02	3.08	3.28	3.41	3.41	3.21	2.95	1.91
4	1.32	1.94	2.41	3.09	3.38	3.37	3.31	3.34	3.04	2.97	2.95	2.96	2.96	2.97	2.69
5	0.92	1.3	1.57	1.88	2.3	2.8	3.41	3.84	4.13	4.02	3.94	3.59	3.05	2.68	1.85
6	3.84	7.85	7.5	5.29	3.89	2.87	2.34	2.12	2.08	2.04	2.04	2.16	2.24	2.37	2.04
7	1.51	2.15	2.26	2.2	2.24	2.2	2.17	2.13	2.29	2.44	2.65	3.3	3.9	4.63	3.89
8	1.84	2.03	1.94	1.85	1.83	1.78	1.75	1.72	1.72	1.78	1.77	1.89	2.29	3.14	9.16
9	1.68	2.07	2.48	2.6	2.58	2.41	2.38	2.2	2.38	2.47	2.77	2.98	3.35	3.56	3.22
10	3.32	5.58	5.49	4.77	3.67	3.06	2.48	2.16	2.03	1.89	1.78	1.82	1.9	1.96	1.97
11	27.9	3	2.28	2.01	1.87	1.78	1.69	1.67	1.71	1.72	1.77	1.88	2.08	2.35	2.64
12	1.38	2.01	2.52	2.83	3.09	2.94	2.82	2.84	2.95	3.32	3.51	3.59	3.53	3.33	2.51
13	0.87	1.3	1.65	2.08	2.57	3.16	3.91	4.27	4.42	4.65	4.36	3.82	3.36	2.72	1.83
14	1.41	2.14	2.29	2.43	2.5	2.74	2.63	2.63	2.64	2.77	2.67	2.54	2.69	2.57	6.18
15	8.46	5.77	3.93	2.86	2.43	2.22	1.99	1.86	1.84	1.75	1.75	1.85	1.93	2.11	2.19
16	1.03	1.52	1.68	1.89	2.22	2.77	3.51	4.06	4.49	4.39	4.24	3.66	3.19	2.51	1.62
17	1.43	2.01	2.2	2.51	2.58	2.75	2.68	2.7	2.77	2.93	3.25	4.1	4.48	4.7	3.58
18	0.91	1.33	1.64	2.09	2.44	2.66	3.04	3.29	3.61	4.02	4.25	4.22	4.03	3.44	2.09
19	1.72	2.92	3.7	4.21	4.34	4.33	4.19	3.99	3.52	3.36	2.98	2.75	2.59	2.34	1.81
20	1.26	1.83	2.02	2.34	2.57	2.6	2.71	2.71	2.76	2.44	2.28	2.89	3	2.99	2.91
21	2.9	4.36	4.29	4.06	3.67	3.26	2.99	2.64	2.44	2.28	2.16	2.15	2.07	1.97	1.63
22	1.77	2.16	2.12	2.16	2.15	2.27	2.33	2.4	2.5	2.51	2.72	3.14	3.4	3.49	3.31
23	9.83	9.93	6.72	5.22	4.08	3.64	3.23	2.9	2.75	2.41	2.3	2.04	1.96	1.91	1.66
24	5.26	8.48	6.99	4.76	3.21	2.52	2.27	2.16	2.05	1.93	1.9	1.73	1.64	1.63	1.52
25	1.31	2.21	2.27	2.43	2.65	3	3.31	3.46	3.55	3.47	3.43	3.18	2.88	2.51	1.83
26	1.5	2.58	3.58	4.1	4.17	3.65	3.11	2.86	2.68	2.81	2.84	2.83	2.5	2.24	1.75
27	1.68	2.17	2.14	1.94	1.79	1.7	1.77	1.78	1.86	1.97	2.08	2.52	3.23	4.56	6.52
28	0.96	1.45	1.73	2.08	2.39	2.55	2.68	2.8	2.99	3.26	3.43	3.76	3.86	3.93	2.67
29	2.41	2.36	2.18	2.13	2.17	2.1	2.05	2.01	1.93	1.86	1.82	1.86	1.99	2.8	6.88
30	1.03	1.57	1.91	2.39	2.85	3.23	3.37	3.48	3.34	3.13	3.16	2.93	2.82	2.56	1.98
31	1.14	1.78	2.09	2.22	2.29	2.29	2.43	2.43	2.5	2.53	2.56	2.8	3.14	3.61	4.47
32	1.01	1.55	2.18	3.25	4.26	4.62	4.34	4	3.58	3.17	2.86	2.51	2.22	2.02	1.55
33	1.16	1.72	2.14	2.85	3.55	3.89	3.99	4.13	3.75	3.44	3.11	2.9	2.7	2.27	1.65
34	1.71	2.68	2.71	3.07	3.2	3.6	3.57	3.53	3.44	3.32	3.35	2.9	2.7	2.43	1.77

כפי שניתן לראות בטבלה 2, הנושא הדומיננטי ביותר במקטע הראשון של כלל העדויות בקורפוס הוא 11 – נושא ההצגה העצמית שראינו לעיל. שני הנושאים הבולטים הבאים במקטע הראשון עוסקים אף הם באלמנטים של ההצגה עצמית: נושא 15 שכולל מילים הקשורות לתא המשפחתי ונושא 23 שכולל מילים הקשורות למרחב המגורים המשפחתי. ממצא זה – אשר מזהה באופן סטטיסטי את מבנה ריאיון העדות – מסייע בתיקוף ממצאי המודל. גם במקטע השני נשמרת הדומיננטיות של שני נושאים בולטים במיוחד (אם כי במידה פחותה בהרבה מנושא 11 במקטע הפותח): נושאים 23 שראינו גם במקטע 1, ונושא 24 שעוסק בחיי הדת והקהילה, כפי שהתקיימו לפני המלחמה.

בשונה מהדומיננטיות החריגה של נושא יחיד (במקרה זה נושא 11 במקטעים הראשונים של העדויות), ביתר מקטעי העדות ניתן לראות שונות רבה יותר בהתפלגות הנושאים, כאשר בכל מקטע נושאים מספר מקבלים משקל דומה (סביב 4% אחוזים) בהתפלגות הנושאים באותו מקטע. כך, למשל, במקטעים 7 ו-8, חלקה האמצעי של העדות, ניתן לראות שהרצף הנרטיבי מתרכז סביב נושאים שונים: נושאים 2 ו-16 שמתארים את מציאות מחנה הריכוז, נושא 13 שעוסק בהשמדה, נושא 19 שעוסק במסע ברכבות ונושאים 32-33 שמתארים היבטים שונים במציאות החיים בגטו. בחינה של הרכב המקטעים 7-8 מלמדת על נרטיבי העדות השונים. בעוד שאצל חלק מהעדים נמצאים חיי הגטו במרכז סיפור העדות (מבחינת משכה), אצל אחרים מרכז הנרטיב עוסק במחנות ההשמדה.

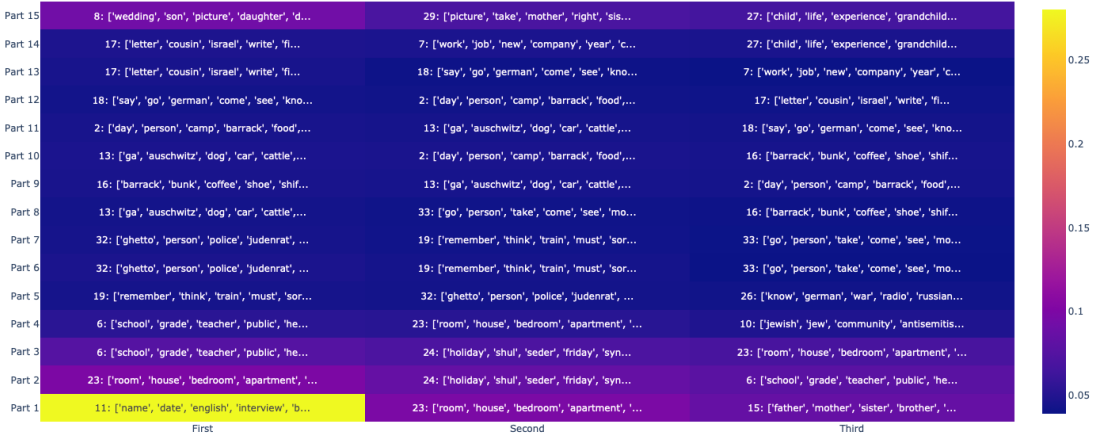
התפלגות הנושאים לאורך רצף העדות מסייעת אף היא בתיקוף המודל, כאשר נושאים שלא היינו מצפים שיופיעו בראשית העדות (למשל מחנות ההשמדה) אכן מופיעים בעיקר במרכז העדות או לקראת סיומה. במקטע האחרון של העדות, מקטע 15, המודל שב ומזהה נושא יחיד שנהנה מדומיננטיות יחסית, נושא 8, שעוסק במילים הקשורות להצגת תמונות. גם כאן, המודל היטיב לזהות חלק מובנה בריאיון העדות והוא ההזמנה של המראיין את המראיין לתאר תמונות משפחתיות עם סיום הריאיון.

לסיכום חלק זה, עד כה ראינו שהמודל מזהה את מבנה העדות הטיפוסי מבחינת ראשית הריאיון וסופו, וכי במרכז העדות ניתן לזהות כמה מסלולים נרטיביים: למשל גטו, מחנה ריכוז ומחנה השמדה. מבט על מאפשר לשרטט את הסכמה הנרטיבית הטיפוסית של עדות שואה דרך ניתוח הנושאים הבולטים ביותר בכל מקטע, כפי שניתן לראות בטבלה 3.

כדי להימנע משרירותיות שעלולה לנבוע מפערים קטנים בין הנושאים הדומיננטיים בכל מקטע, בחרנו להציג את שלושת הנושאים הבולטים ביותר בכל מקטע בממוצע לאורך קורפוס העדויות הכללי. מקטעי העדות בטבלה 3 מסודרים בסדר עולה מהמקטע הראשון (בתחתית התמונה) ועד למקטע האחרון (בראש התמונה). הטור השמאלי מציג את הנושא הבולט ביותר בממוצע באותו מקטע, והטור הימני מציג את הנושא השלישי הבולט ביותר בממוצע באותו מקטע. הצבע של כל תא מסמן את חוזק הנושא באותו מקטע, בהתאם לסולם הצבעים (מופיע בימין התמונה). על מנת להקל על הקריאה, בכל תא מוצגות 4-5 המילים הקשורות ביותר לאותו נושא (רשימת 15 המילים הבולטות ביותר בכל נושא מופיעה בטבלה 1 לעיל). לדוגמה, במקטע מספר 8, באופן ממוצע, הנושא הבולט ביותר הוא נושא 13, שעוסק במחנות ההשמדה וספציפית באשוויץ, הנושא השני הבולט ביותר הוא נושא 33, שעוסק במשפחה ובגטו, והנושא השלישי הבולט ביותר הוא נושא 16, שעוסק במציאות החיים במחנות הריכוז.

טבלה 3: שלושת הנושאים הבולטים ביותר בכל מקטע עדות במבנה הטיפוסי של עדות שואה

All testimonies topics schema

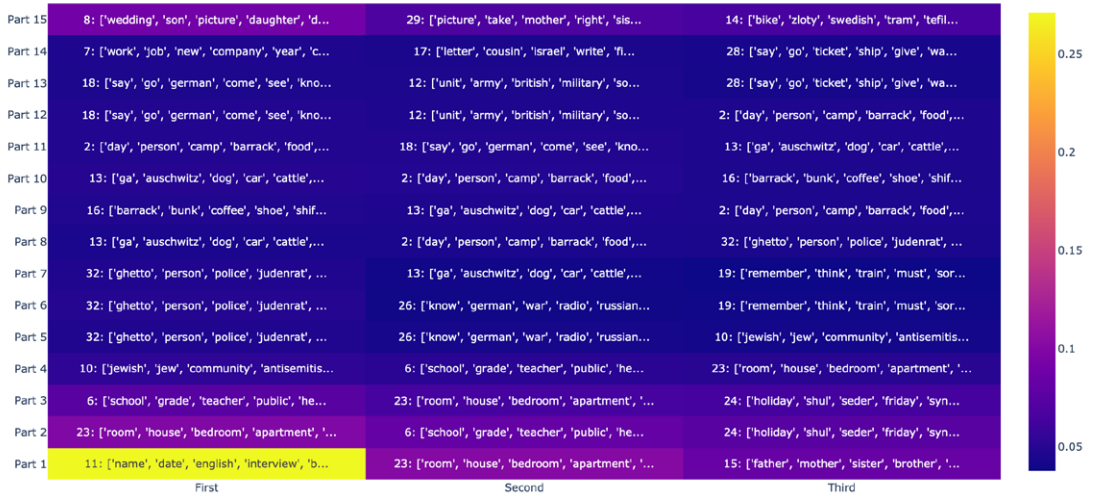


3. שלב שלישי – הטיפוסי ויוצא הדופן בין גברים לנשים

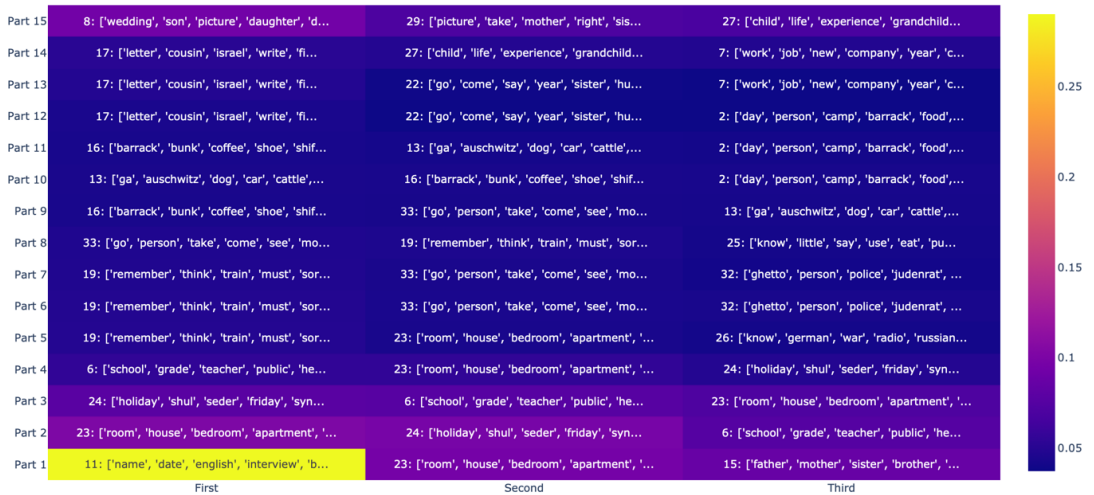
אחרי שזיהינו את המבנה הטיפוסי של עדות שואה בקורפוס, בשלב השלישי נבקש לאתר חריגות מהסכמה הנרטיבית הטיפוסית. לצורך כך, כתוצגת תכלית של המודל, בחרנו להתמקד בחלוקה מגדרית של העדויות ולבחון כיצד מתנהגות עדויות של נשים בהשוואה לעדויות של גברים, וזאת בהתייחס למבנה הטיפוסי הכללי שזיהינו בשלב הקודם. לשם כך, נעזרנו במטא-דאטה של העדויות על מנת לחלק אותן לפי מגדר המרואייין. הקורפוס מורכב מ-531 עדויות גברים (53% מהקורפוס) ו-469 עדויות נשים (47% מהקורפוס). חישבנו את התפלגות הנושאים ואת ממוצע השכיחויות עבור קורפוס הגברים ועבור קורפוס הנשים, וכך זיהינו את הנושאים הבולטים ביותר באופן ממוצע עבור כל מקטע עדות בקורפוס הנשים ובקורפוס הגברים, כפי שניתן לראות בטבלאות 4-א-ב. מקטעי העדות בטבלאות 4-א-ב מסודרים בדומה לטבלה 3.

טבלאות א4-ב: שלושת הנושאים הבולטים ביותר בכל מקטע עדות בעדויות הגברים (א4) ובעדויות הנשים (ב4).

Male testimonies topics schema



Female testimonies topics schema



ראשית, ניתן לראות שהן עדויות נשים והן עדויות גברים שומרות על המבנה הנרטיבי הטיפוסי של ריאיון העדות בפתיחה ובסיום (נושאים 11, 23 בפתיחה ונושא 8 בסיום).²⁴ הדמיון בהרכב

24 כאמור, צבע התא מסמן את מידת הבולטות של הנושא, ואכן ניתן לראות שגם בעדויות נשיות וגם בעדויות גבריות הנושא הבולט ביותר במקטע הראשון – נושא 11 – הוא דומיננטי בצורה

הנושאי בין גברים לנשים נשמר לאורך השליש הראשון של העדות (מקטעים 1-5). עם זאת, הברדל בולט הוא ביחס לנושא 10, שכולל מילים בולטות: "community", "jewish", "antisemitism", "zionist organization". הנושא, שעוסק בחיי הקהילה היהודית, הארגונים היהודיים, כולל התנועה הציונית, ואנטישמיות, מוזכר בעדויות הגברים הן במקטע 4 והן במקטע 5, ואינו מוזכר בעדויות הנשים. בחינה של מקטעים 5-6 בטבלאות 4א ו-ב מלמד שעדויות גברים מעלות בצורה דומיננטית יותר את נושא 26, שכולל מילים כמו "know", "war", "radio", "happen", "army" ו"poland", נושא העוסק בידעיות בדבר הפלישה הנאצית לפולין. עדויות נשים, לעומת זאת, עוסקות בעוצמה רבה יותר ולאורך זמן רב יותר לעומת הגברים בנושא 23, שכולל מילים הקשורות לבית המגורים. במקטעים 5-6 נשים מזכירות בצורה בולטת יותר את נושא 19, שכולל מילים כמו "journey", "train", "suitcase" ו"ship", וגם פעלים הקשורים לזיכרון: "remember", "recall", "think". לפי מיקומו של נושא זה בחצייה הראשון של העדות נראה שהוא עוסק בהגירה/בריחה מארצות המוצא לפני הגירוש למחנות. במקטעים 6-8 נשים עוסקות שוב ושוב בנושא 33, שאינו מוזכר בעדויות הגברים. נושא זה כולל בני משפחה: "mother", "child", "father", וכן פעלים: "take", "come", "walk", ומצטייר כעוסק בקורות בני המשפחה בגטו.

באופן מעניין, נושא 13, שעוסק במחנות ההשמדה, מופיע הן בעדויות הגברים והן באלו של הנשים אבל במיקום שונה. בעוד שגברים מתחילים לדבר על מחנות ההשמדה כבר החל במקטע 7 וממשיכים לעסוק בנושא לאורך רצף ארוך יותר בעדות (מקטעים 7-11), נשים מדברות על מחנות ההשמדה בשלב מאוחר יותר בעדות, רק החל ממקטע 9, ולאורך רצף קצר יותר (מקטעים 9-11 בלבד). מגמה דומה נראית גם בעיסוק בנושא 2, שעוסק במציאות החיים במחנות הריכוז או ההשמדה עם מילים כמו "barrack", "camp", "soup", "march", "die". גברים מתחילים להזכיר את נושא 2 מוקדם יותר לעומת הנשים (מקטע 8 לעומת 10), ובבולטות יחסית גבוהה יותר מאשר בקרב הנשים. עם זאת, מגמה זאת אינה חוזרת על עצמה לגבי נושא 16, שאף הוא עוסק במציאות החיים במחנות הריכוז, עם דגש מסוים על מחנות עבודה, באמצעות מילים כמו "barrack", "bunk", "factory", "shift". בחלקים האחרונים של העדויות נשים עוסקות בנושא 17, העוסק באיתור בני משפחה ("write", "letter", "cousin", "meet", "palestine", "dp" (displaced people) camp או מחנות העקורים) ובנושא 22, שמתייחס אף הוא למשפחה: "brother", "sister", "mother", "child", "husband".

ניתוח הרצף הנרטיבי והחריגה מהסכמה הטיפוסית הכללית מדגים את האופן שבו ניתן לקרוא מחדש עדויות שואה דרך הפריזמה של הטיפוסי ויוצא הדופן. בדוגמה לעיל התמקדנו בחלוקה על יסוד מגדר וזיהינו את המקומות בעדויות שבהם העדות הגברית או הנשית חורגות מן הסכמה הטיפוסית הכללית. בעזרת ניתוח זה, מצאנו שנשים מקדישות חלקים מרכזיים בעדות לפן המשפחתי: בתחילת העדות בזמן ההצגה העצמית ותיאור החיים לפני השואה, בחלקים העוסקים במחנות, וכן לקראת סוף העדות בתיאור החיים לאחר השואה. ראינו גם שנשים מנכיחות בצורה בולטת יותר את אתגר זיכרון השואה. בעדויות הגברים מצאנו דגש רב יותר על הפוליטי והחברתי לעומת האישי, הן במהלך תיאור החיים

לפני המלחמה והן בתיאור החיים בגטו. בעדויות הגברים גם ראינו עיסוק בצבא ובחיילים עם שחרור המחנות. ההתמקדות בפן האישי בעדויות הגברים – לפני המלחמה ולאחריה כאחד – הדגישה יותר היבטים של השכלה ועבודה.

4. שלב רביעי – סיבוב המכונה

שלב הניתוח הרביעי נועד להדגים את המהלך ההמשכי של המודל, בכחינת סיבוב נוסף של המכונה. אחרי שבחנו את העדויות דרך הפרזימה של השונות בהתפלגות הנושאים בכל מקטע בין גברים לנשים, ביקשנו לזהות, שוב כתצוגת תכלית, את העדות הגברית החריגה, קרי זאת שהרכב הנושאים בה מתפלג בצורה הדומה ביותר לעדות נשית, ולהפך – העדות הנשית "הגברית" ביותר. לשם כך, ערכנו השוואה נוספת בין עדויות המגדרים השונים, ובה בחנו את מידת הפער באזכור הנושאים בעדויות הגברים והנשים. בשונה מהשלב השלישי, שבו השוינו את שלושת הנושאים הבולטים ביותר בכל מקטע על בסיס חלוקה מגדרית, כאן המודל החישובי איתר את הנושאים שעבורם מתקיים הפער היחסי הגדול ביותר בין מידת אזכורם בעדויות נשיות ובעדויות גבריות, כפי שניתן לראות בטבלה 25.²⁵

גווני האדום מציינים נושאים שאוזכרו באופן בולט יותר באופן יחסי בעדויות נשים לעומת עדויות הגברים (ככל שהצבע אדום יותר / המספר שלילי יותר, כך הפער גדול יותר לטובת האזכור הנשי) וגווני הירוק מציינים נושאים שהפער באזכורם הוא לטובת עדויות הגברים (ככל שהצבע ירוק יותר / המספר חיובי יותר, כך הפער גדול יותר לטובת האזכור הגברי). כך למשל, נושא 1 במקטע 7 בולט באופן יחסי בעדויות הנשים ואילו נושא 2 במקטע 8 בולט באופן יחסי בעדויות הגברים.

25 טבלה 5 היא תוצאה של מדידת ההפרשים בין התפלגות הנושאים בעדויות הגברים המובאת בנספח 1 להתפלגות הנושאים בעדויות הנשים המובאת בנספח 2. כאמור, אין מדובר בהכרח בנושאים הבולטים ביותר בכל מקטע.

טבלה 5: הפער המגדרי באזכור נושאים בעדויות

Topic/Part	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	-0.16	-0.24	-0.31	-0.34	-0.22	-0.31	-0.24	-0.46	-0.5	-0.44	-0.45	-0.62	-0.65	-0.57	-0.26
1	-0.16	-0.4	-0.36	-0.43	-0.44	-0.51	-0.76	-0.63	-0.53	-0.55	-0.63	-0.9	-0.64	-0.54	-0.42
2	0.07	0.1	0.14	0.12	0.25	0.15	0.2	0.84	0.35	0.55	0.72	0.83	0.53	0.55	-0.02
3	0.15	0.17	0.45	0.11	0.28	0.48	0.54	0.5	0.38	0.53	0.49	0.57	0.44	0.25	0.11
4	-0.04	-0.13	-0.43	-0.58	-0.69	-0.61	-0.43	-0.42	-0.23	-0.28	-0.48	-0.6	-0.6	-0.31	-0.16
5	0.03	0.01	0.09	0.18	0.21	0.07	0.24	0.25	0.24	0.14	0.1	0.07	0.11	-0.07	0.04
6	0.21	0.7	-0.21	-0.39	0.51	0.57	0.16	-0.03	-0.07	-0.01	-0.08	0.01	-0.01	0.13	0.03
7	0.04	-0.03	0.06	0.05	0.07	0.15	0.06	-0.03	-0.09	-0.22	0	0.12	0.34	-0.2	0
8	0.07	0.08	0.14	-0.04	-0.07	0.03	-0.03	-0.02	0	-0.07	0.01	0.08	-0.24	-0.38	-0.36
9	0.47	0.57	0.56	0.63	0.7	0.74	0.66	0.55	0.64	0.69	0.64	0.6	0.58	0.91	0.81
10	0.9	1.59	1.35	1.09	1	1.27	0.6	0.41	0.49	0.32	0.19	0.24	0.42	0.36	0.39
11	-1.88	0.36	0.21	0.1	0.06	0.15	0.15	0.13	0.12	0.11	0.1	0.16	0.1	0.14	0.21
12	0.39	0.84	1.12	1.33	1.43	1.5	1.36	1.47	1.48	1.84	2.18	2.11	1.94	1.73	0.99
13	0.05	0.06	0.14	0.08	0.02	0.22	0.23	0.31	0.37	0.56	0.42	0.28	0.34	0.2	0.05
14	0.08	0.23	0.42	0.31	0.35	0.51	0.31	0.42	0.37	0.45	0.39	0.35	0.71	0.58	0.64
15	-0.34	-0.76	-0.29	-0.33	-0.46	-0.39	-0.24	-0.3	-0.31	-0.33	-0.34	-0.39	-0.32	-0.21	-0.29
16	0.11	0.11	0.23	0.27	0.21	0.34	0.39	0.6	0.37	0.12	-0.09	0.34	0.21	0.04	0.03
17	-0.02	-0.07	0.12	-0.15	-0.08	-0.36	-0.28	-0.37	-0.2	-0.49	-0.53	-0.59	-0.68	-0.76	-0.29
18	0.03	0.13	0.16	0.29	0.5	0.39	0.71	0.88	0.76	0.91	1.07	1.17	0.91	0.64	0.31
19	-0.07	-0.36	-0.41	-0.53	-1.04	-0.86	-0.88	-0.96	-0.6	-0.4	-0.47	-0.48	-0.47	-0.2	-0.26
20	-0.1	-0.09	-0.01	0.11	0.07	0.08	0.18	0.19	0.09	0.05	-0.09	-0.19	0.08	-0.06	0.06
21	0.21	0.22	-0.19	-0.16	-0.25	-0.23	-0.08	-0.24	-0.26	-0.21	-0.21	-0.31	-0.21	-0.03	-0.03
22	-0.24	-0.4	-0.36	-0.47	-0.37	-0.54	-0.52	-0.55	-0.65	-0.6	-0.8	-1.13	-0.89	-0.66	-0.64
23	0.6	-0.23	-0.63	-0.51	-0.46	-0.42	-0.58	-0.4	-0.15	-0.38	-0.12	-0.06	-0.1	-0.15	-0.19
24	-0.46	-2.23	-1.44	-0.29	-0.23	-0.37	-0.35	-0.47	-0.33	-0.44	-0.29	-0.24	-0.19	-0.2	-0.14
25	-0.19	-0.41	-0.43	-0.51	-0.75	-1.04	-0.86	-1.22	-1.1	-1.16	-0.97	-0.83	-0.8	-0.57	-0.32
26	0.16	0.21	0.14	0.28	0.18	0.62	0.51	0.59	0.29	0.22	0.12	0.18	0.13	0.14	0.09
27	-0.21	-0.28	-0.23	-0.29	-0.03	-0.08	-0.28	-0.26	-0.28	-0.26	-0.31	-0.58	-0.84	-0.61	-0.56
28	0.01	0.12	0.21	0.2	0.23	0.25	0.44	0.51	0.61	0.4	0.75	0.86	0.82	0.75	0.55
29	0.18	0.05	0.03	-0.02	-0.14	-0.13	-0.14	-0.18	-0.1	-0.04	-0.1	-0.11	-0.32	-0.31	-0.22
30	-0.14	-0.18	-0.27	-0.58	-0.78	-1.12	-0.88	-0.91	-0.91	-0.72	-0.85	-0.71	-0.53	-0.53	-0.21
31	0.02	0.05	-0.01	0.17	0.02	0.1	-0.24	-0.02	0.12	0.08	0.17	-0.09	0.02	0.26	0.14
32	0.15	0.25	0.48	0.67	0.82	0.58	0.62	0.76	0.66	0.39	0.43	0.45	0.41	0.23	0.17
33	-0.08	-0.1	-0.38	-0.59	-0.71	-0.98	-0.68	-0.88	-0.94	-0.54	-0.62	-0.5	-0.55	-0.25	-0.17
34	0.13	0.05	-0.11	0.22	-0.2	-0.28	0.12	-0.08	-0.09	-0.21	-0.36	-0.12	-0.02	-0.28	-0.07

בחנית טבלת הפערים בהתפלגות הנושאים בין גברים לנשים מאפשרת לעקוב אחרי תמות שנמתחות על פני רצף של מקטעים ולאחר את הדגשים הנרטיביים של הקבוצות השונות. כך, למשל, ניתן לזהות רצפים "נשיים" בנושאים 0 (פעלים המתארים פעולת סיפור או דיבור), 1 (העוסק באיודאות בנוגע לגורל המשפחה), 4 (העוסק אף הוא בהתמודדות המשפחה), 19 (הגירה), 22 (דמויות משפחתיות), 25 (פעלים ושמות עצם שעוסקים בקשיי היום-יום) ו-30 (בריחת המשפחה). לצד זאת, המודל מאתר רצפים "גבריים" בנושאים 2 (שעוסק במחנות), 3 (מתקפות צבאיות), 9 (שכולל אזכורי המדינות גרמניה, ישראל וארצות הברית), 10 (הקהילה היהודית וארגוניה והאנטישמיות), 12 ו-18 (כוחות צבא), 28 (הגירה) ו-32 (גטו).

כפי שצינינו, המטרה של המודל אינה להגיע לזיהוי מבנה עדות טיפוסית מייצגת, אלא זיהוי הטיפוסי כאמצעי להפניית תשומת הלב ליוצא הדופן. בהתאם לכך, הלכנו צעד נוסף וביקשנו לאתר את העדות יוצאת הדופן בקבוצת הגברים – קרי העדות הגברית שמתנהגת מבחינת ההרכב הנושאי שלה כמו עדות "נשית", ולהפך – העדות הנשית שמזכירה מבחינת הרכבה הנושאי עדות "גברית". על יסוד זיהוי הרצפים המגדריים בטבלה 5, המודל איתר את העדויות הנשיות שממקסמות את הערכים הגבריים וממזערות את הערכים הנשיים, וכך גם לגבי העדויות של הגברים, אלה שממקסמות את הערכים הנשיים וממזערות את הערכים הגבריים.²⁶

כך, העדות הנשית שזכתה לציון הגבוה ביותר, קרי, זוהתה כעדות הנשית "הגברית" ביותר, היא עדותה של איירין שפיגל.²⁷ שפיגל נולדה בארצות הברית, נסעה לאירופה לעבוד כאחות ערב המלחמה, נמלטה מספרד לצרפת, ולאחר כיבוש צרפת העבירה את שנות המלחמה בזהויות בדויות ובבריחה מתמדת עם בנה בצרפת, עד לשחרור. קריאה צמודה בעדותה של שפיגל מסייעת להבין מדוע זיהה המודל את עדותה כעדות נשית חריגה כל כך. שפיגל ממעטת לדבר על נושאי משפחה ומקדישה חלק ניכר מעדותה לתיאור חזיתות הקרב שבהן עבדה כאחות. במובן זה, עדותה אכן עוסקת בנושאים שזיהינו קודם לכן כ"גבריים" – צבא ומלחמה – ונוגעת פחות בנושאים "נשיים" של קורות בני המשפחה וגורלם.

מן העבר השני, המודל זיהה את עדותו של הנריק גרינברג כעדות "גברית" חריגה.²⁸ קריאה צמודה בעדותו של גרינברג חושפת עדות המתארת בהרחבה ובפירוט רב את תלאות משפחתו הקרובה והמורחבת. הדבר נובע בין היתר מכך שגרינברג היה ילד קטן עם פרוץ המלחמה והצליח להעביר את מרבית שנותיה במחבוא עם אימו. עדותו רצופה בתיאור היחסים ביניהם ובמאמצי ההישרדות היומיומיים. עדות זו בולטת בפירוט התחושות והרגשות של העד ובחום ובקרבה שעולים מהם. בהינתן שמדובר בעדות של מי שהיה ילד בתקופת המלחמה, בדקנו אם עדותו של גרינברג היא חריגה גם כאשר מחלקים את העדויות לפי גיל. התברר כי עדותו אינה חורגת מהסכמה הנרטיבית הטיפוסית של עדויות ניצולים שהיו ילדים קטנים בזמן השואה. מהלך זה מדגים את האופי האיטרטיבי של הגישה שאנו מציעים, שלפיה ניתן למקם כל עדות על ציר של טיפוסי-חריג כתלות בקטגוריה שדרכה מתבוננים על הנתונים, וזאת על מנת לגלות באופן מחזורי מופעי עדות חדשים.

26 באשר לנושאים באותם מקטעים מאופיינים מגדרית, סכמנו את הערכים המאפיינים גברים והחסרנו מהסכום את סכום הערכים המאפיינים נשים באותה עדות. התוצאה מייצגת את המרחק של העדות הנשית שנבחרה מהעדות הטיפוסית הגברית. ככל שהמרחק גבוה יותר, כך העדות נחשבת כבעלת מאפיינים "גבריים" יותר בערכים המאפיינים מגדרית. אותו תהליך בוצע עבור העדויות של הגברים. עדויות שהמרחק שלהן מהעדות הטיפוסית של המגדר הנגדי הוא הקטן ביותר הוגדרו כעדויות בעלות מאפיינים מגדריים נגדיים. טבלת המקטעים המאופיינים מגדרית מצורפת בנספח 3.

27 USC SHOAH FOUND., *Jewish Survivor Irene Spiegel*, VISUAL HISTORY ARCHIVE (Nov. 9, 1997), <https://vha.usc.edu/testimony/37352?from=search>

28 USC SHOAH FOUND., *Jewish Survivor Henryk Grynberg*, VISUAL HISTORY ARCHIVE (Nov. 25, 1997), <https://vha.usc.edu/testimony/35668?from=search>

כך, באמצעות חזרה אל הטקסט וקריאה קרובה בעדויות שהמודל איתר כחריגות, ניתן לחשוף סיפורים שאינם מצייתים לתבניות נרטיביות טיפוסיות ובאופן זה להתוודע לזוויות חדשות במכלול החוויות שלמדנו להכיר בסיפורי שואה. אפשר להמשיך עוד ולשאול: האם ניתן להצביע על לקחים מיוחדים של הניצולים בעדויות החריגות, לקחים שעליהם הם נשאלים בדרך כלל בסוף הריאיון? האם הם מציגים עצמם באופן שונה בהשוואה לעדויות טיפוסיות? שאלות אלה מתייחסות כמובן לפרמטרים שזוהו כטיפוסיים לנשים ולגברים, ובאותו אופן ניתן לחשוב על ניתוח דומה לגבי גיל, ארץ לידה או מצב משפחתי. בכל פעם המטרה תהיה לזהות דפוסים טיפוסיים ומתוכם חריגות, לבחון את העדויות היוצאות מן הכלל (שעשויות להיות טיפוסיות לגמרי בחיתוכים אחרים), לצלול אל העולם שהן מתארות, להסיק מסקנות, וכך חוזר חלילה.

ה. הזיכרון הבלתי מתקבע: סיכום וכיווני המשך

במאמר זה ביקשנו להדגים את הפוטנציאל הטמון בשימוש במודלים חישוביים לצורך הגדרה דינמית, המשתנה מדי הרצה, של מבנה עדות טיפוסית ויוצא דופן, על מנת להצביע על קולות וסיפורים לא-טיפוסיים ביחס לפרמטרים נתונים. כתהליך המשכי, יש בגישה זו פוטנציאל להציע אלטרנטיבה לסיפורים מוכרים במסגרת זיכרון השואה. במודל החישובי המוצע, הטכנולוגיה אינה אמצעי גרידא – היא חלק בלתי נפרד מתהליך ההקשבה, ומשתתפת בעיצוב דרכי הגישה, החוויה וההבנה של עדויות.

מראשית מלאכת התיעוד של עדויות השואה, כאשר הפסיכולוג החברתי דיוויד בודר החל להקליט עדויות ניצולים במחנות עקורים באירופה, נרתמה הטכנולוגיה לטובת מטרה היסטורית ומוסרית: גילוי העבר, חקר האמת, והנחלת הזיכרון לדורות הבאים. אמצעים טכנולוגיים שונים נהפכו עם השנים לחלק בלתי נפרד מהאופן שבו ניתן להקשיב לעדויות שואה, לחקור אותן וללמוד מהן. המדיה של העדות מעצבים את תנאי האפשרות של העדות למצוא נמענים ולהעביר את מסריה. כלים חישוביים מהווים מדיום חדש המאפשר גישה מחודשת אל עדויות באופן שמקיים את הציווי המוסרי של הנצחת השואה. עם זאת, אין דין זיכרון השואה בהקשר העכשווי של תרבות דיגיטלית כדין ההקלטות הראשונות של עדויות ניצולים. אם בעבר המטרה הייתה הכרה באובדן של הקורבנות והפצת הסיפורים כנגד סכנת הכחשת שואה, הרי שכיום האתגר שונה. דומה שכעת, דווקא הודות להצלחת מאמצי ההנצחה והנחלת הזיכרון, דמות הניצול והז'אנר העדויות הפכו למוכרים, שלא לומר התקבעו סביב תבנית קבועה של נרטיב וסגנון. כיום האתגר הוא דווקא לחתור תחת המקובל והידוע.²⁹ המודל שהצגנו מייצר מבנה טיפוסית של עדות הנתון לשינוי מתמיד. המטרה אינה להגיע לזיהוי צורה טיפוסית קבועה, בבחינת צורה מייצגת, אלא זיהוי הטיפוסי כאמצעי להפניית תשומת הלב ליוצא הדופן. הטיפוסי משתנה תמידית ומשמש כבסיס ארעי להחרגה. כך גם הטיפוסי וגם יוצא הדופן משתנים ללא הרף. באופן זה נמנעת התקבעות הזיכרון סביב נרטיב כזה או אחר, כאשר בכל פעם ניתן להיחשף לסיפורים, לתמות ולדוברים שלא ניתן

LAWRENCE L. LANGER, USING AND ABUSING THE HOLOCAUST ;10 לעיל ה"ש; PINCHEVSKI 29 (2006); GEOFFREY H. HARTMAN, THE LONGEST SHADOW: IN THE AFTERMATH OF THE HOLOCAUST (1996).

היה להגיע אליהם בעבר. תוצאה משתנה תמידית זאת מאפשרת להגיע לתובנות מחקריות חדשות, כפי שהצגנו במקרה הבוחן של עדויות הנשים והגברים. ואולם, לצד הממד המחקרי, יש במודל גם פוטנציאל לשימוש הנצחתי, חינוכי ואוצרותי. חלק מהותי מהציווי המוסרי של הנצחת השואה (או כל זיכרון אחר שיש בו ממד מוסרי) הוא ההכרח לגלות את הזיכרון בכל פעם מחדש. הסכנה הטמונה בכל פרויקט זיכרון, היא של התקבעות הזיכרון לפי נוסחה מסוימת, מסר מסוים, פרספקטיבה של קבוצה מסוימת, או לקח מסוים. ברגע שמתרחשת התקבעות שכזאת, הזיכרון עלול להפוך למונומנט קפוא ולחדול מלהיות זיכרון משמעותי וחי. השינוי המחזורי שמובנה בגישה שהצענו משרת ומדגים בכל פעם מחדש את אי-התקבעותו של הזיכרון.

נספחים

נספח 1: התפלגות הנושאים בעדויות הגברים

Topic/Part	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	1.8	2.62	2.78	2.71	2.59	2.44	2.45	2.25	2.24	2.28	2.38	2.4	2.59	2.94	2.77
1	1.59	2.34	2.42	2.26	2.29	2.2	2.11	2.15	2.13	2.06	1.97	1.94	2.22	2.57	2.98
2	0.93	1.39	1.65	2	2.52	2.98	3.45	4.38	4.5	4.81	4.78	4.55	3.96	3.33	1.81
3	1.07	1.72	2.52	3.18	3.6	3.58	3.45	3.25	3.26	3.53	3.64	3.68	3.42	3.07	1.96
4	1.3	1.88	2.21	2.81	3.05	3.08	3.11	3.14	2.93	2.84	2.73	2.68	2.69	2.54	1.98
5	0.93	1.3	1.62	1.96	2.4	2.83	3.53	3.95	4.24	4.08	3.99	3.63	3.1	2.65	1.87
6	3.94	8.18	7.4	5.11	4.13	3.14	2.41	2.11	2.05	2.03	2.01	2.16	2.23	2.43	2.06
7	1.53	2.14	2.29	2.23	2.28	2.27	2.2	2.11	2.25	2.33	2.65	3.35	4.06	4.54	3.89
8	1.87	2.07	2	1.84	1.8	1.79	1.73	1.71	1.72	1.75	1.78	1.93	2.17	2.96	8.99
9	1.91	2.34	2.74	2.89	2.91	2.75	2.69	2.46	2.68	2.79	3.07	3.26	3.62	3.99	3.61
10	3.75	6.32	6.12	5.28	4.13	3.65	2.76	2.35	2.26	2.04	1.86	1.93	2.1	2.13	2.16
11	27.1	3.16	2.38	2.06	1.9	1.85	1.76	1.73	1.76	1.77	1.82	1.96	2.13	2.41	2.74
12	1.56	2.4	3.05	3.46	3.76	3.64	3.46	3.53	3.64	4.18	4.54	4.58	4.44	4.14	2.99
13	0.89	1.33	1.71	2.12	2.58	3.26	4.01	4.41	4.59	4.91	4.56	3.96	3.52	2.81	1.85
14	1.45	2.24	2.49	2.58	2.66	2.98	2.78	2.82	2.82	2.98	2.85	2.7	3.03	2.84	6.49
15	8.3	5.41	3.79	2.7	2.21	2.04	1.88	1.72	1.7	1.59	1.59	1.67	1.78	2.01	2.04
16	1.08	1.57	1.78	2.02	2.32	2.93	3.69	4.34	4.67	4.45	4.2	3.82	3.28	2.53	1.64
17	1.42	1.98	2.25	2.44	2.55	2.58	2.54	2.52	2.68	2.7	3	3.83	4.16	4.35	3.43
18	0.92	1.4	1.71	2.23	2.68	2.85	3.37	3.71	3.96	4.44	4.75	4.77	4.45	3.74	2.24
19	1.68	2.75	3.51	3.96	3.85	3.92	3.78	3.55	3.24	3.17	2.76	2.53	2.37	2.25	1.69
20	1.21	1.79	2.02	2.39	2.61	2.64	2.79	2.8	2.81	2.81	2.85	2.91	3.02	2.88	2.4
21	3	4.47	4.2	3.98	3.56	3.16	2.96	2.53	2.32	2.18	2.06	2.01	1.97	1.96	1.61
22	1.66	1.97	1.95	1.94	1.98	2.01	2.09	2.14	2.2	2.22	2.35	2.61	2.98	3.18	3
23	10.1	9.82	6.42	4.98	3.87	3.44	2.96	2.71	2.68	2.24	2.25	2.01	1.91	1.84	1.57
24	5.05	7.43	6.32	4.62	3.1	2.34	2.11	1.94	1.89	1.73	1.77	1.61	1.55	1.54	1.45
25	1.22	2.01	2.06	2.19	2.3	2.52	2.9	2.89	3.03	2.93	2.98	2.79	2.5	2.24	1.67
26	1.57	2.68	3.65	4.24	4.25	3.94	3.35	3.13	2.82	2.91	2.89	2.91	2.56	2.31	1.79
27	1.58	2.04	2.03	1.8	1.78	1.66	1.64	1.66	1.72	1.84	1.94	2.25	2.83	4.27	6.25
28	0.96	1.5	1.83	2.18	2.5	2.67	2.89	3.03	3.28	3.44	3.78	4.16	4.24	4.28	2.94
29	2.49	2.38	2.19	2.12	2.1	2.04	1.98	1.93	1.88	1.84	1.77	1.81	1.85	2.66	6.78
30	0.96	1.49	1.78	2.11	2.48	2.71	2.95	3.05	2.91	2.79	2.76	2.6	2.57	2.32	1.88
31	1.15	1.8	2.09	2.3	2.3	2.33	2.32	2.42	2.56	2.56	2.64	2.76	3.15	3.73	4.54
32	1.08	1.67	2.4	3.56	4.65	4.89	4.63	4.36	3.89	3.35	3.06	2.72	2.41	2.13	1.63
33	1.13	1.67	1.96	2.57	3.21	3.43	3.67	3.72	3.31	3.19	2.82	2.67	2.44	2.15	1.57
34	1.78	2.7	2.66	3.17	3.11	3.46	3.63	3.49	3.4	3.22	3.18	2.84	2.69	2.3	1.73

נספח 2: התפלגות הנושאים בעדויות הנשים

Topic/Part	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	1.96	2.86	3.08	3.05	2.81	2.74	2.68	2.71	2.74	2.72	2.83	3.02	3.24	3.52	3.03
1	1.75	2.74	2.78	2.69	2.73	2.71	2.87	2.78	2.66	2.61	2.61	2.84	2.85	3.1	3.39
2	0.85	1.3	1.5	1.88	2.26	2.82	3.25	3.53	4.15	4.26	4.06	3.72	3.43	2.79	1.83
3	0.93	1.55	2.07	3.07	3.32	3.1	2.91	2.75	2.88	3	3.15	3.1	2.98	2.82	1.85
4	1.34	2	2.64	3.4	3.74	3.69	3.54	3.56	3.16	3.12	3.2	3.28	3.29	2.85	2.14
5	0.91	1.29	1.53	1.78	2.19	2.76	3.29	3.71	4	3.95	3.89	3.56	2.99	2.71	1.83
6	3.73	7.47	7.61	5.5	3.62	2.57	2.25	2.14	2.12	2.04	2.09	2.15	2.24	2.3	2.03
7	1.49	2.17	2.23	2.17	2.2	2.12	2.14	2.14	2.34	2.56	2.65	3.23	3.72	4.74	3.89
8	1.8	1.99	1.86	1.87	1.87	1.76	1.76	1.73	1.72	1.81	1.76	1.85	2.42	3.34	9.35
9	1.43	1.77	2.18	2.26	2.21	2.02	2.03	1.91	2.04	2.1	2.43	2.67	3.04	3.08	2.81
10	2.85	4.73	4.77	4.19	3.14	2.38	2.16	1.94	1.77	1.72	1.68	1.69	1.68	1.77	1.76
11	28.9	2.8	2.16	1.96	1.84	1.69	1.61	1.6	1.65	1.66	1.72	1.8	2.03	2.27	2.53
12	1.17	1.56	1.93	2.13	2.33	2.14	2.09	2.06	2.16	2.34	2.35	2.47	2.5	2.42	2
13	0.84	1.27	1.57	2.04	2.55	3.04	3.79	4.1	4.22	4.35	4.14	3.68	3.17	2.61	1.81
14	1.37	2.02	2.07	2.27	2.31	2.47	2.47	2.4	2.45	2.53	2.46	2.35	2.32	2.26	5.85
15	8.65	6.18	4.08	3.04	2.67	2.43	2.12	2.02	2.01	1.92	1.93	2.06	2.1	2.22	2.34
16	0.97	1.46	1.55	1.75	2.1	2.59	3.3	3.74	4.3	4.33	4.28	3.48	3.08	2.49	1.61
17	1.44	2.05	2.14	2.59	2.62	2.94	2.83	2.89	2.88	3.19	3.53	4.41	4.84	5.11	3.73
18	0.89	1.26	1.55	1.94	2.18	2.46	2.67	2.83	3.2	3.53	3.69	3.6	3.55	3.09	1.93
19	1.75	3.11	3.91	4.49	4.89	4.79	4.66	4.5	3.83	3.57	3.23	3.01	2.84	2.45	1.95
20	1.31	1.88	2.03	2.28	2.53	2.55	2.61	2.61	2.72	2.76	2.93	3.11	2.95	2.95	2.34
21	2.78	4.25	4.39	4.15	3.81	3.39	3.03	2.77	2.58	2.39	2.27	2.31	2.18	1.98	1.65
22	1.9	2.38	2.32	2.4	2.35	2.56	2.61	2.7	2.85	2.83	3.14	3.74	3.87	3.84	3.64
23	9.51	10.0	7.05	5.49	4.33	3.86	3.54	3.11	2.83	2.62	2.36	2.07	2.01	1.99	1.76
24	5.51	9.66	7.76	4.92	3.33	2.71	2.46	2.41	2.22	2.17	2.05	1.86	1.74	1.74	1.59
25	1.41	2.42	2.5	2.7	3.05	3.55	3.76	4.11	4.13	4.09	3.95	3.62	3.31	2.81	1.99
26	1.41	2.47	3.51	3.95	4.07	3.32	2.85	2.55	2.52	2.69	2.78	2.73	2.43	2.17	1.7
27	1.8	2.32	2.26	2.09	1.81	1.74	1.91	1.91	2.01	2.11	2.25	2.82	3.67	4.88	6.81
28	0.95	1.39	1.62	1.98	2.27	2.42	2.45	2.53	2.67	3.04	3.03	3.3	3.42	3.53	2.39
29	2.31	2.33	2.16	2.14	2.24	2.16	2.12	2.11	1.98	1.88	1.87	1.92	2.16	2.97	7
30	1.1	1.67	2.06	2.69	3.26	3.83	3.83	3.96	3.82	3.51	3.61	3.31	3.1	2.85	2.09
31	1.13	1.75	2.1	2.13	2.28	2.24	2.55	2.44	2.43	2.48	2.47	2.84	3.14	3.47	4.39
32	0.93	1.42	1.92	2.89	3.83	4.31	4.02	3.59	3.23	2.97	2.63	2.27	2	1.9	1.46
33	1.2	1.78	2.34	3.16	3.93	4.41	4.35	4.6	4.25	3.73	3.44	3.17	2.99	2.4	1.74
34	1.64	2.65	2.77	2.95	3.3	3.75	3.51	3.57	3.49	3.43	3.53	2.97	2.71	2.57	1.8

נספח 3: טבלת הנושאים/מקטעים המאופיינים מגדרית (תאים המוקפים בשוליים אדומים או שחורים), שעל יסודם חושבו העדויות החריגות ביותר בקרב גברים ובקרב נשים

Topic/Part	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	-0.16	-0.24	-0.31	-0.34	-0.22	-0.31	-0.24	-0.46	-0.5	-0.44	-0.45	-0.62	-0.65	-0.57	-0.26
1	-0.16	-0.4	-0.36	-0.43	-0.44	-0.51	-0.76	-0.63	-0.53	-0.55	-0.63	-0.9	-0.64	-0.54	-0.42
2	0.07	0.1	0.14	0.12	0.25	0.15	0.2	0.84	0.35	0.55	0.72	0.83	0.53	0.55	-0.02
3	0.15	0.17	0.45	0.11	0.28	0.48	0.54	0.5	0.38	0.53	0.49	0.57	0.44	0.25	0.11
4	-0.04	-0.13	-0.43	-0.58	-0.69	-0.61	-0.43	-0.42	-0.23	-0.28	-0.48	-0.6	-0.6	-0.31	-0.16
5	0.03	0.01	0.09	0.18	0.21	0.07	0.24	0.25	0.24	0.14	0.1	0.07	0.11	-0.07	0.04
6	0.21	0.7	0.21	-0.39	0.51	0.57	0.16	-0.03	-0.07	-0.1	-0.08	0.01	-0.01	0.13	0.03
7	0.04	-0.03	0.06	0.05	0.07	0.15	0.06	-0.03	-0.09	-0.22	0	0.12	0.34	-0.2	0
8	0.07	0.08	0.14	-0.04	-0.07	0.03	-0.03	-0.02	0	-0.07	0.01	0.08	-0.24	-0.38	-0.36
9	0.47	0.57	0.56	0.63	0.7	0.74	0.66	0.55	0.64	0.69	0.64	0.6	0.58	0.91	0.81
10	0.9	1.59	1.35	1.09	1	1.27	0.6	0.41	0.49	0.32	0.19	0.24	0.42	0.36	0.39
11	-1.88	0.36	0.21	0.1	0.06	0.15	0.15	0.13	0.12	0.11	0.1	0.16	0.1	0.14	0.21
12	0.39	0.84	1.12	1.33	1.43	1.5	1.36	1.47	1.48	1.84	2.18	2.11	1.94	1.73	0.99
13	0.05	0.06	0.14	0.08	0.02	0.22	0.23	0.31	0.37	0.56	0.42	0.28	0.34	0.2	0.05
14	0.08	0.23	0.42	0.31	0.35	0.51	0.31	0.42	0.37	0.45	0.39	0.35	0.71	0.58	0.64
15	-0.34	-0.76	-0.29	-0.33	-0.46	-0.39	-0.24	-0.3	-0.31	-0.33	-0.34	-0.39	-0.32	-0.21	-0.29
16	0.11	0.11	0.23	0.27	0.21	0.34	0.39	0.6	0.37	0.12	-0.09	0.34	0.21	0.04	0.03
17	-0.02	-0.07	0.12	-0.15	-0.08	-0.36	-0.28	-0.37	-0.2	-0.49	-0.53	-0.59	-0.68	-0.76	-0.29
18	0.03	0.13	0.16	0.29	0.5	0.39	0.71	0.88	0.76	0.91	1.07	1.17	0.91	0.64	0.31
19	-0.07	-0.36	-0.41	-0.53	-1.04	-0.86	-0.88	-0.96	-0.6	-0.4	-0.47	-0.48	-0.47	-0.2	-0.26
20	-0.1	-0.09	-0.01	0.11	0.07	0.08	0.18	0.19	0.09	0.05	-0.09	-0.19	0.08	-0.06	0.06
21	0.21	0.22	-0.19	-0.16	-0.25	-0.23	-0.08	-0.24	-0.26	-0.21	-0.21	-0.31	-0.21	-0.03	-0.03
22	-0.24	-0.4	-0.36	-0.47	-0.37	-0.54	-0.52	-0.55	-0.65	-0.6	-0.8	-1.13	-0.89	-0.66	-0.64
23	0.6	-0.23	-0.63	-0.51	-0.46	-0.42	-0.58	-0.4	-0.15	-0.38	-0.12	-0.06	-0.1	-0.15	-0.19
24	-0.46	-2.23	-1.44	-0.29	-0.23	-0.37	-0.35	-0.47	-0.33	-0.44	-0.29	-0.24	-0.19	-0.2	-0.14
25	-0.19	-0.41	-0.43	-0.51	-0.75	-1.04	-0.86	-1.22	-1.1	-1.16	-0.97	-0.83	-0.8	-0.57	-0.32
26	0.16	0.21	0.14	0.28	0.18	0.62	0.51	0.59	0.29	0.22	0.12	0.18	0.13	0.14	0.09
27	-0.21	-0.28	-0.23	-0.29	-0.03	-0.08	-0.28	-0.26	-0.28	-0.26	-0.31	-0.58	-0.84	-0.61	-0.56
28	0.01	0.12	0.21	0.2	0.23	0.25	0.44	0.51	0.61	0.4	0.75	0.86	0.82	0.75	0.55
29	0.18	0.05	0.03	-0.02	-0.14	-0.13	-0.14	-0.18	-0.1	-0.04	-0.1	-0.11	-0.32	-0.31	-0.22
30	-0.14	-0.18	-0.27	-0.58	-0.78	-1.12	-0.88	-0.91	-0.91	-0.72	-0.85	-0.71	-0.53	-0.53	-0.21
31	0.02	0.05	-0.01	0.17	0.02	0.1	-0.24	-0.02	0.12	0.08	0.17	-0.09	0.02	0.26	0.14
32	0.15	0.25	0.48	0.67	0.82	0.58	0.62	0.76	0.66	0.39	0.43	0.45	0.41	0.23	0.17
33	-0.08	-0.1	-0.38	-0.59	-0.71	-0.98	-0.68	-0.88	-0.94	-0.54	-0.62	-0.5	-0.55	-0.25	-0.17
34	0.13	0.05	-0.11	0.22	-0.2	-0.28	0.12	-0.08	-0.09	-0.21	-0.36	-0.12	-0.02	-0.28	-0.07