

‘וקרב אותם אחד אל אחד’: זיהוי כתיבות יד וצירוף קטעי גניזה באמצעות מחשב

אהרן שויקה, יעקב שויקה, ליאור וולף, נחום דרשוביץ

אחת הבעיות הקשות העומדות בפני כל חוקר גניזה היא קיטועם של כתבי היד המקוריים לדפים בודדים ולחלקי דפים. קשה יותר הוא פיזורם של אותם קטעים באוספים שונים ובספריות שונות, תחת מספרי מדף רבים, עד שכתב יד אחד נחלק לשבע ארצות. כך תיאר בעיה זו י' זוסמן בדבריו על שרידי תלמוד שבגניזה:

חסרונם הגדול – **היותם קטעי קטעים**, שהרי תורתך נעשית קרעי קרעים, אם אינך יודע להבחין בין קטע לקטע, לסווגם ולמיינם ולשקול את משקלם מבחינת מוצאם וטיבם. מטרתו של טיפול טקסטואלי פילולוגי קפדני אינו איסוף ורישום שינויי נוסחאות וחילופי גרסאות מכל הבא ליד, כאשר כל הרוצה נוטל את שלבו חפץ או שעומד על המניין: כך וכך נוסחאות גורסות כך וכך וכך גורסות אחרת. תפקידו של החוקר לשקול ולבחון, כאשר הוא יודע מה לקרב ומה לרחק. אין תועלת מרובה בקולציות מדוקדקות כביכול ובאפרטים כבדים, כאשר אינך יכול ליצור סטמה כלשהי לכתבי־היד המרובים שאת גרסאותיהם אתה רושם. והנה קטעים אלה של הגניזה, בהיותם קרעי קרעים מפורדים, קשה מאוד לעמוד על משקלם הסגולי מבחינת מוצאם וטיבם ולשייכם למסורת נוסח זו או אחרת.¹

* מאמר זה הוא פרי מחקר משותף של יחידת 'גנזים' בפרויקט פרידברג לחקר הגניזה ושל המחלקה למדעי המחשב באוניברסיטת תל אביב. החוקרים הראשיים מיחידת גנזים הם ראש היחידה פרופ' יעקב שויקה ובנו ד"ר אהרן שויקה. החוקרים הראשיים מאוניברסיטת תל אביב הם פרופ' ליאור וולף ופרופ' נחום דרשוביץ. את תכניות המחשב כתבו תלמידים מתקדמים במדעי המחשב: רותם ליטמן, מאשה זלדין, נעמה מאייר, טניה גרמן וליזה פוטיחה. בבדיקת הצירופים וביהוי הקטעים השתתף נסים לוק. המחקר נתמך בידי פרויקט פרידברג לחקר הגניזה ויד הנדיב.

¹ הרצאה בכנס לחקר הגניזה הקהירית, אוניברסיטת תל אביב, 24 במרס 1976. ההרצאה פורסמה בתעודה, א (תש"ם), עמ' 21–31, ובנוסח מעודכן בכרך המבואות לאוצר כתבי היד התלמודיים בעריכת י' זוסמן [בדפוס].

בין הראשונים שעמדו על החשיבות הגדולה של קיבוץ הקטעים שהתפזרו יחד כשלב מקדים של המחקר היה פ' קהלה.² הוא גיבש שיטה לרישום כתבי היד ולצירוף הקטעים וסלל את הדרך לבאים אחריו בחקר מסורות הלשון והמקרא, עד לחיבורו המסכם של י' ייבין על הניקוד הבבלי, המבוסס על למעלה מאלף קטעי גניזה שהוא שייך לכתבי היד השונים.³ גם במהדורות של מסכתות המשנה והתלמוד הבבלי, רובן ככולן, ניתנה תשומת לב לצירופם של קטעי הגניזה הרבים.⁴ יש לציין במיוחד את כרכי המשנה סדר זרעים ואת התלמוד הבבלי סדר נשים שהוציא לאור מכון התלמוד הישראלי השלם. אכן, מלאכת צירופם של קטעי הגניזה של המשנה והתלמוד הגיעה קרוב לשלמות באוצר כתבי היד התלמודיים שבעריכת יעקב זוסמן, שבו נרשמו אלפי צירופים בין כששת אלפים וארבע מאות קטעי הגניזה התלמודיים הרשומים באוצר.⁵

מלבד דוגמאות מופת אלה, ישנם בידינו כיום כמה וכמה מחקרים המציגים רשימות בהיקף קטן יותר, שמחבריהם השתדלו למצות את הצירופים של קטעי הגניזה שהביאו ברשימותיהם.⁶ כמו כן נרשמו צירופים רבים שלא באופן שיטתי בקטלוגים שונים⁷ ובפרסומים אחרים. עם זאת, ישנן דוגמאות לא מעטות לחוסר מודעות של החוקרים

2 P. Kahle, 'Die hebräischen Babeli', *Zeitschrift für die alttestamentliche Wissenschaft*, 46 [1928], pp. 113–137 ובמאמרו על כתבי יד בניקוד בבלי של המשנה (P. Kahle & J. Weinberg, 'The Mishna Text in Babylonia — Fragments From the Geniza', *Hebrew Union College Annual*, 10 [1935], pp. 185–223). ראה גם: אוצר כתבי היד התלמודיים, שם, מבוא, הערה 77.

3 י' ייבין, מסורת הלשון המשתקפת בניקוד הבבלי, ירושלים תשמ"ה. רבים מכתבי היד המתוארים בחיבור הוא פרסם קודם לכן בפקסימיליות: י' ייבין, אוסף קטעי הגניזה של המקרא בניקוד ובמסורה הבבלי, ירושלים תשל"ג; הנ"ל, אוסף קטעי הגניזה של המשנה בניקוד בבלי, ירושלים תשל"ד.

4 לרשימה של חיבורים אלה ראה: ד' גולניקין, גנזי ראש השנה, ניו יורק וירושלים תש"ס, עמ' 2], הערה 11.

5 אמנם גם שם לא נבדקו צירופים באופן שיטתי בכתבי יד ובקטעים של סידורים, במשנה אבות, במשנה עם פירוש הרמב"ם ובהעתיקים מן הדפוס. ראה: אוצר כתבי היד התלמודיים (לעיל, הערה 1), מבוא, הערה 79.

6 נציין כאן מספר דוגמאות בלבד: נ' דנציג, מבוא לספר הלכות פסוקות עם תשלום הלכות פסוקות, ניו יורק וירושלים תשנ"ג, עמ' 625–636 (94 קטעים); מ' כהנא, אוצר כתבי היד של מדרשי ההלכה, ירושלים תשנ"ה (249 קטעים); א' אלמגור, כתבי היד של מדרשי ר' דוד הנגיד, ירושלים תשנ"ה; א' שויקה, עיונים בספר הלכות גדולות: נוסח ועריכה, עבודת דוקטור, האוניברסיטה העברית בירושלים, תשס"ח, עמ' 30–51 (306 קטעים).

7 כגון: נ' דנציג, קטלוג של שרידי הלכה ומדרש מגניזת קאהיר באוסף א"נ אדלר שבספריית בית המדרש לרבנים באמריקה, [ניו יורק וירושלים תשנ"ח. ראה שם, מבוא, עמ' לא, והערה *105. בביקורת של ע' שבט על הקטלוג (E. Chwat, 'Danzig's Catalogue of Halakhah and Midrash Fragments in the E. N. Adler Collection and Its Usage as a Research Tool', *Jewish Quarterly Review*, 90 [2000], pp. 405–415) הוא רושם עשרות צירופים נוספים לקטעים שנרשמו בקטלוג. כמו כן נרשמו הרבה צירופים בקטלוג 'אלף', הקטלוג הממוחשב של המכון לתצלומי כתבי יד בספרייה הלאומית, רובם בידי ע' שבט.

לצירופם של קטעי הגניזה. לעתים המחקר חסר הואיל והוא נשען על קטע בודד, שניתן היה לצרף לו קטעים מאותו כתב יד שאינם ידועים לחוקר.⁸ ולעתים נמצאים לפני החוקר קטעים שונים מאותו קודקס אלא שהוא אינו מצרפם ולומד כל קטע וקטע בנפרד, והדבר פוגם בתוצאות המחקר.⁹

אכן, האפשרות לצרף קטעי קודקס שהתפזרו לאוספים שונים הייתה מצומצמת ביחס כל זמן שהגניזה הייתה מפוזרת בספריות שונות ברחבי העולם ולא היה לחוקר אלא מה שענינו רואות לפניו. משום כך התאפיין מחקר הגניזה מראשיתו בשיתוף פעולה בין חוקרים היושבים במקומות שונים, וכל חוקר הוסיף למחקר קטע חדש שהצטרף לקטע שפרסם חברו. ישנן דוגמאות רבות לחיבורים שהתגלו בגניזה ופורסמו פסקי פסקי, ואף כך היה גורלו של אחד הראשונים והמיוחדים שבהם, הוא ספר 'בן סירא' בנוסחו העברי.¹⁰ תרומה מכרעת לצירופם של קטעי הגניזה תרם המכון לתצלומי כתבי יד בספרייה הלאומית, שכינס בין כתליו מיקרופילמים של כל קטעי הגניזה מן האוספים שבהם התפזרו. ועדיין, מלאכת הצירוף דרשה הכנה של צילומים איכותיים מן המיקרופילמים או שימוש בשני מכשירי הקרנה בו זמנית, דבר שלא תמיד היה אפשרי וזמין.

עידן חדש בחקר הגניזה נפתח עם עליית אתר האינטרנט של 'גנזים', יחידת המחשוב של פרויקט פרידברג לחקר הגניזה.¹¹ מעתה ניתן לצפות באתר בתמונות דיגיטליות איכותיות של עשרות אלפי קטעי גניזה ואף יש בו מידע קטלוגי וביבליוגרפי מגוון שרוכז מעשרות מקורות אל מקום אחד. האתר כולל כיום, בשנת תשע"א, יותר ממאתיים אלף תמונות, ובכלל זה אוספי גניזה שלמים, כגון אוסף בית המדרש לרבנים בניו יורק, אוסף כ"ח בפרזי ואוספים רבים אחרים. ועדיין חסרים אנו תמונות של קטעי הגניזה מן האוספים

8 ראה לדוגמה מה שכתב כהנא (לעיל, הערה 6), עמ' 19.

9 ראה: זוסמן (לעיל, הערה 1), הערה 39*. בחיבורו של ג' בירנבאום, לשון המשנה בגניזת קהיר, ירושלים תשס"ח, הוא מביא תיאור לשוני של קטעי משנה רבים אך ממעט מאוד בצירופים שביניהם. אף שמהדורת המשנה סדר זרעים של מכון התלמוד הישראלי השלם עמדה לפניו, הוא לא נזקק לצירופי הקטעים שהובאו שם כלל, וכך הוא מקדיש פרקים שלמים לתיאורים לשוניים של קטעים בודדים שייכתן מאוד שהיו שייכים לקודקס אחד. על פי אוצר כתבי היד התלמודיים (לעיל, הערה 1), כ"י ג 8 רג 9 הנזכרים אצל בירנבאום מצטרפים לאותו כתב יד, ותיאורם היה צריך לבוא יחד. כך גם כ"י גא 2 רגא 4, כ"י גא 3 רגא 5, כ"י גא 2 רגא 2, כ"י 39 רגא 4, ועוד.

10 לסקירת הפרסומים של הקטעים מכתבי היד של בן סירא ראה: מ"צ סגל, ספר בן-סירא השלם, ירושלים תשי"ח², עמ' 48-49. מאז התפרסמו קטעים נוספים המצטרפים לאותם כתבי יד והאחרון שבהם ממש בשנים האחרונות. ראה: ש' אליצור, 'קטע חדש מהנוסח העברי של ספר בן סירא', תרביץ, עו (תשס"ז), עמ' 17-28.

שבאוקספורד ובמנצ'סטר וכן את אוספי אנטונין ופירקוביץ שבסנט פטרבורג. בשנת תשס"ט יזמה יחידת גנזים מחקר משותף עם פרופ' ליאור וולף ופרופ' נחום דרשוביץ, מן המחלקה למדעי המחשב באוניברסיטת תל אביב, ובמסגרתו אנו מפתחים מערכת לצירוף אוטומטי של קטעי הגניזה. תוצאות ראשוניות של המערכת הוצגו בהרצאתו של פרופ' יעקב שויקה, ראש יחידת גנזים, במסגרת הקונגרס העולמי החמישה עשר למדעי היהדות, ירושלים תשס"ט. בד בבד הציג פרופ' וולף את המערכת בכנסים ובמסגרות מקצועיות של מדעי המחשב. להלן נציע תיאור כללי של מאפייני המערכת בשפה שאינה מקצועית.¹² בחלקו השני של המאמר יוצגו דוגמאות מעטות מבין התוצאות הרבות שהשיגה המערכת עד עתה, על פי בדיקות שערכנו בסיועו של מר נסים לוק.

א. תיאור המערכת

הקלט העיקרי של המערכת הוא תמונות דיגיטליות איכותיות של קטעי הגניזה שברשותנו. המאפיין העיקרי של זיהוי הצירופים הוא כמובן הדמיון בכתיבת היד של שני קטעים. המערכת מנתחת את התמונות ומתרגמת את המידע החזותי למידע נומרי באמצעות טכניקות ואלגוריתמים מתוחכמים של ראייה ממוחשבת ושל עיבוד תמונה, הדומים לאלה המשמשים במערכות זיהוי פנים. המערכת מבצעת השוואה של אותו מידע נומרי לכל שתי תמונות ונותנת ציון המייצג את מידת הקרבה ביניהם. ציון זה מתורגם למדד סטטיסטי, המעריך את ההסתברות שזוג התמונות יכול להיות צירוף אמיתי.

1. צירוף חזק וצירוף חלש

כל חוקר שהתנסה בצירוף קטעי גניזה יודע עד כמה דורשת מלאכה זו חכמה. הרבה פעמים נראה במבט שטחי ששני קטעים משקפים כתיבה אחת ורק עין של פליאוגרף מומחה יכולה להבחין בהבדלים ולקבוע שהם נכתבו בידי סופרים שונים. לעומת זאת, סופר יכול לשנות את כתב ידו במהלך הזמן ואף באותו טופס עצמו. אין זה נדיר שחל שינוי בצורת האות אל"ף או נעשה שימוש בסגנונות שונים לכתיבת שם הוי"ה מדף לדף ואף באותו עמוד עצמו. ואכן, הצעות צירוף רבות שנרשמו במחקר טעונות תיקון. ואולם

12 לתיאור טכני של האלגוריתמים והתוצאות ולפרסומים קודמים בתחום זה ראה: L. Wolf et al., 'Automatically Identifying Join Candidates in the Cairo Genizah', *International Journal of Computer Vision*, 90 (2010), pp. 1–18

ישנם מקרים הפוכים, שבהם החוקר מבקש להבחין בין כתיבות שונות אף שמדובר בכתיבה אחת, ואף בדפים מאותו טופס, גם כאשר הקטעים עומדים לפניו בגופם. הקושי בצירוף קטעים ואי-הוודאות הכרוכה בזיהוי כזה הובילו את עורכי 'אוצר כתבי היד התלמודיים' להבחין בין שלוש רמות של צירופים:

- א. 'צירוף פיזי' – צירוף של קטעים השייכים לאותו דף או לאותו גיליון (ביפוליום).
- ב. 'ראה' – צירוף של קטעים או של דפים הנראים שייכים לאותו טופס כאשר הם מאותו סדר משנה או מסכת תלמוד.
- ג. 'ראה גם' – צירוף כמו בסעיף הקודם ביחס לקטעים מסדר משנה אחר או ממסכת תלמוד אחרת.

הרמה הבטוחה ביותר היא כמובן 'צירוף פיזי'. צירוף 'ראה' הוא ברמת ודאות קטנה יותר, אלא אם כן הדפים המצטרפים ממשיכים זה את זה. הרמה הפחות בטוחה היא זו של דפים מסדרי משנה שונים או ממסכתות תלמוד שונות.

במהלך עבודתנו בפיתוח המערכת ובבדיקת התוצאות שעלו ממנה היינו צריכים לעבור במהירות על אלפי זוגות של קטעים שהתוכנה סימנה כצירופים ולהחליט אם מדובר בצירוף אמיתי אם לאו. לשם כך הגדרנו שתי רמות של צירופים:

- א. **צירוף חזק** – צירוף של שני קטעים השייכים לאותו טופס.
- ב. **צירוף חלש** – שני קטעים הכתובים באותה כתיבה ובידי אותו סופר אבל הם שייכים לטפסים שונים.

למעשה, צירוף חלש במובנו הרגיל אינו צירוף כלל. אין הוא מצרף שני קטעים לאותו טופס אלא משייך אותם לסופר אחד. עם זאת, לפי הכללים של התוכנה, האמורה לזהות כתיבות דומות, אין הבדל בין צירוף חזק לצירוף חלש. אם כתיבת היד דומה, היא אמורה לצרף את הקטעים. אף על פי כן הבחנו בין צירוף חזק לצירוף חלש, כי רמת הביטחון בצירוף חלש נמוכה בהרבה מזו של צירופים חזקים (שבהם שני הקטעים שייכים לאותו טופס), וצירוף כזה זוקק בדיקה מקצועית של פליאוגרף מומחה. כמובן, ישנה חשיבות גדולה גם לצירופים מעין אלה, שכן יש בהם כדי ללמד על זמנו ועל מקומו של טופס אחד במקרה שחברו מעיד עליו.

2. מאפיינים נוספים לזיהוי צירופים

דמיון בכתיבת היד יכול להעיד על כך שסופר אחד כתב את שני הקטעים הנבדקים. ואולם, כדי לזהות צירוף חזק, היינו שני קטעים מאותו טופס, יש צורך במאפיינים נוספים. העיקרי שבהם הוא כמובן זיהוי החיבור. כאשר יש דמיון בכתיבת היד של שני קטעים ושניהם שייכים לאותו חיבור, רב הסיכוי שלפנינו קטעים מאותו הטופס. ולהפך, אם הקטעים הם מחיבורים שונים, מסתבר שמדובר בטפסים שונים, וייתכן מאוד שאין לפנינו אפילו צירוף חלש אלא דמיון מסוים בכתיבה בלבד. אכן, מאפיין זה דורש זיהוי מוקדם של הקטע, ולכן השימוש בו מוגבל לקטעים שנרשמו בקטלוגים, ואין הם אלא מיעוט קטעי הגניזה. בכל מקרה, המערכת שתשתמש במאפיין זה תתבסס על מידע חיצוני, שאיכותו תלויה באדם שרשם אותו.

ואולם ישנם מאפיינים של הקטע שאינם קשורים בתוכנו ואינם מחייבים קריאה שלו, וניתן להיעזר בהם לקביעת הצירופים. המאפיינים הללו כוללים נתונים פיזיים של הקטע: החומר שעליו הוא נכתב (קלף או נייר), אורכו ורוחבו, ומאפיינים של כתיבת הטקסט: טור אחד או כמה טורים, מספר השורות, גודל השוליים וכדומה. נתונים מעין אלה רשומים בדרך כלל בקטלוגים השונים, ואחת המטרות של רישום זה היא להקל על זיהוים של צירופים אפשריים.

והנה, גם בנתונים אלה אין אנו תלויים בקטלוגים. באמצעות אלגוריתמים של עיבוד תמונה פיתחנו מערכת המסוגלת למדוד נתונים שונים של כל קטע וקטע, ובמקרה של ביפוליום – של כל דף ודף מן הגיליון. ואלה הנתונים שבדקת התוכנה: גודל הדף, שוליים עליונים ותחתונים ושוליים ימניים ושמאליים. כמו כן, המערכת מזהה תמונות של ביפוליום ומפרידה בין הדפים, ואף מזהה כתיבה בטור אחד או בשני טורים. בכל תיבת טקסט (טור) היא מזהה את גודל השטח הכתוב, סופרת את מספר השורות ומודדת את צפיפותן: גובה שורה ממוצע ומרווח ממוצע בין שתי שורות. התוכנה אף מודדת צפיפות ממוצעת של אותיות ליחידת מידה קבועה (עשרה ס"מ) בדף. המערכת גם מזהה את שלמות הדף: האם חסרות בו פינות או שמא הטקסט נוגע בקצה הדף ללא שוליים, דבר המלמד על דף קרוע או חתוך. באופן כזה יכולה המערכת לזהות אם הנתונים שהיא מדדה, כגון גודל הדף ומספר השורות, הם נתונים מלאים או חלקיים.

מדירות אלה דורשות מספר שלבים מקדימים, כגון:

א. סגמנטציה – הפרדת הקטע מן הרקע שעליו הוא מונח וזיהוי רכיבים

בתמונה שאינם שייכים לקטע הגניזה (כגון סרגל, מהדקי נייר, כריות משקולת וכדומה);

ב. **כיול התמונה** (זיהוי רמת ההפרדה המדויקת שלה) באמצעות זיהויו של סרגל המצולם עם הקטע ומדידתו, או מדידת הרשת המילימטרית שעליה הקטע מונח;

ג. **בינריזציה** – הפיכת התמונה הצבעונית לתמונה בשחור-לבן;

ד. **יישור התמונה** – אם הקטע צולם בזווית לא ישרה או שהוא מונח על צדו.

כל השלבים האלה נעשים באופן אוטומטי במערכת והקלט היחיד שלה הוא התמונה הדיגיטלית. המערכת בונה אפוא בסיס נתונים ענק המכיל את נתוניהם של כל קטעי הגניזה, ללא תלות בקטלוגים או במידע חיצוני אחר. בסיס נתונים זה חשוב כשלעצמו, אך הוא גם משמש מערכת תומכת לזיהוי הצירופים, להבחנה בין צירוף חזק לצירוף חלש ולשקלול ההסתברות של הצירוף בין שני קטעים.

3. מרחב הנתונים וביצועי המערכת

פעולת ההשוואה בין שני מסמכים נעשית במהירות גבוהה מאוד מכיוון שהיא אינה משתמשת בתמונות עצמן אלא מתבצעת על וקטורים נומריים המייצגים את סגנון האותיות בכל מסמך ומסמך. אמנם, כדי להבטיח אחזור מלא של כל הצירופים בגניזה, לפחות ביחס למסמכים שיש לנו תמונות דיגיטליות שלהם, אנו צריכים להשוות כל אחד מן המסמכים לכל המסמכים האחרים. חשבון פשוט יוכיח כי מרחב ההשוואות שצריך לבצע לשם מיצוי כל אפשרויות הצירופים בגניזה עצום בגודלו. באופן עקרוני, די היה לבצע השוואות בין מסמכים שונים המאופיינים במספרי מדף שונים. כמו כן, לכאורה היה מספיק לבחור מכל מסמך תמונה מייצגת אחת (היינו עמוד אחד). ואולם נראה שאין הדבר כך. רמת ההשתמרות של הטקסט שונה במקרים רבים מעמוד לעמוד, ואם בעמוד אחד הטקסט בהיר וברור הרי שבעמוד שכנגדו הטקסט דהוי, מוכתם ובלתי קריא. בחירה של העמוד הפחות טוב תקטין אפוא את הסיכויים לזיהוי צירופים של המסמך. כמו כן, פעמים רבות חל שינוי בכתיבת היד מדף לדף ומעמוד לעמוד. כדי להבטיח מיצוי של אפשרויות הצירוף יש לבצע אפוא את ההשוואה עבור כל תמונה ותמונה ואין להסתפק בדף אחד מתוך מסמך.

הטבלה הבאה מסכמת את מספר ההשוואות הנדרש כדי למצות את כל קורפוס הגניזה:¹³

מספר הפרטים	מספר ההשוואות	
~230,000	26.5×10^9	מספרי מדף
~650,000	211.2×10^9	תמונות

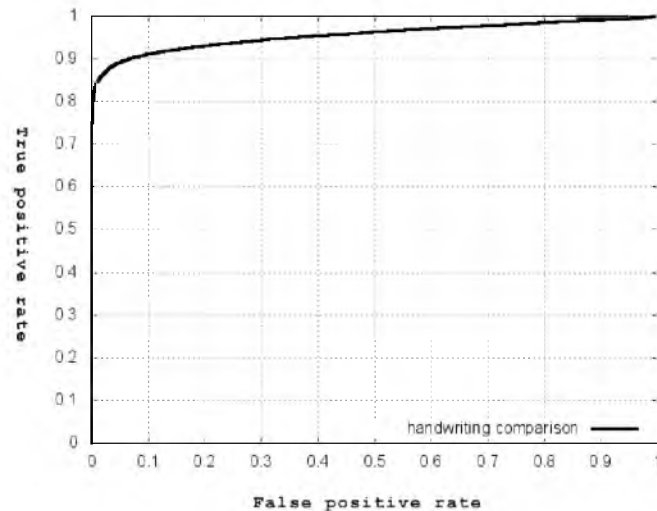
אם כן, מדובר בסדר גודל של עשרות ומאות מיליארדים של פעולות השוואה. כמו כל מערכת אחזור, ביצועי המערכת נמדדים באמצעות גרף ROC (Relative Operating Characteristic Curve), המציג את היחס בין שני פרמטרים: רמת האחזור ורמת השגיאות. רמת האחזור נמדדת על פי אחוז התוצאות הנכונות שאוחזרו מסך התוצאות הנכונות הקיימות. כלומר, במקרה שלנו: אחוז הציורפים הנכונים שהמערכת איתרה מסך הציורפים הקיימים במרחב. מקובל לייצג את הפרמטר בערכים שבין 0 ל-1, כאשר 1 פירושו אחזור מלא. כמובן, השאיפה היא שהפרמטר הזה יהיה קרוב ככל האפשר ל-1. רמת השגיאות נמדדת באמצעות אחוז התוצאות החיוביות השגויות שהמערכת מחזירה מתוך סך הזוגות, דהיינו אחוז הזוגות שהמערכת מסמנת כצירוף, אף שאינם כאלה, מתוך מספר הזוגות שנבדקו. תוצאה כזאת נקראת false positive, והשאיפה היא כמובן שיהיו כמה שפחות תוצאות מעין אלה. גם נתון זה נמדד בערכים שבין 0 ל-1, כאשר 1 פירושו שהמערכת מחזירה תוצאה חיובית עבור כל זוג בין אם הוא צירוף ובין אם לאו, ואילו 0 פירושו שאין כל שגיאות.

ככלל, בכל אלגוריתם שנפעיל אנו יכולים לשלוט ברמת האחזור ובמספר השגיאות. ואולם שני פרמטרים אלה תלויים זה בזה: ככל שרמת האחזור תהיה גבוהה יותר כך גם יגדל אחוז השגיאות, וככל שאחוז השגיאות יהיה נמוך יותר כך תקטן רמת האחזור. ההבדלים בין האלגוריתמים השונים יתבטאו ביחס שבין שני הפרמטרים הללו בנקודות שונות. על כן עלינו לקבוע מהו המחיר שנרצה לשלם כדי להעלות את רמת האחזור או להוריד את רמת השגיאות. לשם בחינת ביצועי המערכת יש לנסות אותה על מרחב

13 נכון לנובמבר 2010 נמצאים באתר מעל מאתיים אלף תמונות.

נתונים שבו הצירופים האמיתיים ידועים מראש. ואכן, מדגם כזה של מספר אלפי מסמכים נבנה לצורך בחינת המערכת.

להלן יובא גרף ROC המציג את רמת הביצועים הנוכחית של המערכת לזיהוי כתב יד, ללא התחשבות במאפיינים אחרים.¹⁴ בנובמבר 2010 עומדת רמת האחזור של המערכת על 90% (true positive rate), ואחוז השגיאה עומד על 0.1% (false positive rate). מעבר לנקודה זו, ככל שאנו עולים ברמת האחזור כך גדלה והולכת רמת השגיאה, והמחיר שאנו צריכים לשלם עבור שיפור רמת האחזור אינו משתלם יותר.



4. הרצה ראשונה

הרצת הגרסה הראשונה של המערכת התבצעה במעבדות החוג למדעי המחשב של אוניברסיטת תל אביב ביולי 2009. בהרצה זו נכללו כל האוספים שהיו בידינו צילומים דיגיטליים שלהם באותו הזמן. רשימת אוספים זו חופפת לרשימה הנמצאת כיום באתר פרויקט פרידברג לחקר הגניזה, למעט אוסף קמברידג' ואוסף הספרייה הבריטית, שצילומיהם עדיין לא היו מוכנים לשימוש, וכן מספר אוספים קטנים שנוספו בשלב מאוחר יותר. לפיכך אין נכללים בהצעות הצירוף קטעים מספריות אלה. בהרצה זו נכללו בסך הכול קרוב לשמונים אלף תמונות, והיא התבצעה בעשרים וארבעה מחשבים שהרצו

14 גרסה רחבה יותר של גרף זה נמצאת במאמר הנזכר לעיל, הערה 12.

את המערכת במשך כשלושה שבועות. בהרצה זו נבדק רק הדמיון בין כתיבות היד ולא נעשה שימוש במאפיינים אחרים של הקטע שצוינו לעיל.

בדיקה ראשונית של כחמשת אלפים זוגות, מתוך אלפים רבים של זוגות שהתוכנה סימנה כצירופים, העלתה כחמש מאות וחמישים צירופים חזקים וכשלוש מאות עשרים וחמישה צירופים חלשים. הואיל והתוכנה אמורה לזהות כתיבות דומות ואינה מבדילה בין צירוף חזק לצירוף חלש, הגדרנו את שני הסוגים כתוצאה חיובית נכונה (true positive). מכאן שכשבעה עשר אחוזים של הזוגות מתוך המדגם שנבדק היו תוצאות חיוביות נכונות (true positive), והשאר היו תוצאות חיוביות שגויות (false positive).

מתוך מאות הצירופים שהתוכנה זיהתה ואושרו על ידינו כצירופים נכונים, אנו מפרסמים כאן מבחר מצומצם של מאה ושלושה עשר צירופים המורכבים מכשלוש מאות ועשרים קטעי גניזה. ככל הידוע לנו, כל הצירופים המפורסמים להלן לא היו רשומים עד עתה. קטלוג צירופים זה מוכיח באופן משכנע את יכולתה של המערכת לאתר צירופים חדשים של קטעי גניזה. יש לציין כי היכולת המוכחת של התוכנה אינה תלויה בסגנון הכתב, ובין הצירופים שזוהו ניתן למצוא את כל סגנונות הכתב: מרובע, בינוני או רהוט, ובכתב מזרחי, ספרדי, איטלקי או אשכנזי.

5. המשך הפיתוח

פיתוח המערכת עדיין לא הסתיים. בימים אלה אנו מכינים את ההרצה הבאה של המערכת, שתכלול יותר ממאתיים אלף תמונות (20×10^6 השוואות). בזמן שחלף מן ההרצה הראשונה שיפרנו את אלגוריתם ההשוואה בין כתבי היד שיפור של ממש, פיתחנו אלגוריתמים נוספים לגזירה אוטומטית של מאפייני הקטעים, ושילבנו את בסיס הנתונים באופן שיאפשר למערכת לשקלל לצורך ההשוואה פרמטרים נוספים של הקטעים, כמו מידות הדף, מספר השורות ותוכן הקטעים. כמו כן הוספנו למערכת יכולת להתמודד עם תוצאות שגויות שיוצרים קטעים בעייתיים, והתחלנו לגבש דרך לעיבוד חכם של התוצאות המתקבלות כדי לאתר בהן במהירות תוצאות חיוביות שגויות.

אנו מעריכים שבהרצות הבאות יהיו התוצאות טובות יותר במידה ניכרת, ואנו מקווים שהן יגיעו לרמה משביעת רצון ונוכל להעמיד אותן בעתיד לרשות החוקרים במסגרת האתר של פרויקט פרידברג לחקר הגניזה. בכך נתקרב מאוד לקראת מה שנראה כרגע כחלום רחוק: העמדה חדשה של כל הגניזה הקהירית, על כל אוספיה, משוחזרת לטפסים המקוריים ככל ששרדו בזמן גניזתם. מימושו המוצלח של חזון זה עדיין מחייב עבודה רבה, ובין השאר יש צורך בבדיקות מדגמיות רחבות של התוצאות, בעדכונים של האלגוריתמים

ובהרצות חוזרות ונשנות עד לייצוב המערכת סביב רמת אֶחזור הקרובה למאה אחוז עם אחוז שגיאות מזערי. מכיוון שעבודה זו עוד תימשך זמן רב, ראינו לנכון לפרסם כבר עתה מבחר מתוך התוצאות שבידינו כפי שעלו בהרצה הראשונה, כדי לאפשר לקהילת החוקרים להתרשם מיכולותיה של המערכת ומן הברכה הטמונה בה.

ב. קטלוג צירופים

את כל הצירופים שבקטלוג זה הציעה התוכנה, והם נבדקו בידי אהרן שויקה יחד עם נסים לוק. השתדלנו לא לרשום צירופים שכבר צוינו בקטלוגים ובמחקרים, או כאלה שכבר רשמו חוקרי הפרויקט באתר פרידברג לחקר הגניזה. אף לא רשמנו צירופים של הספרות התלמודית מכיוון שהם רשומים ב'אוצר כתבי היד התלמודיים' הנמצא בדפוס. במקרים רבים זוהו הקטעים לראשונה על ידינו. בסוף הקטלוג הובא מפתח קטעי הגניזה וציון הסעיף שבו הקטע הוזכר.

מקרא

1. Geneva 43 (שני דפים) – בראשית ה, יא–כד; ח, יג–כב; AIU I A 59 (דף אחד) – בראשית ו, ד–טז.
2. JTS Krenzel 14a–b (שני דפים) – בראשית יב, י–טו; טו, יז – טז, ג; JTS Lutzki 519/3–8 – בראשית יג, ג – יד, ג; יד, טו – טו, ט.
3. AIU I A 158 (שני דפים) – שמות ג, טו – ד, ט; Strasbourg 4068/27 (שני דפים) – שמות כ, יח – כא, יח.
4. Geneva 101 (שני דפים) – שמות ג, כב – ד, כב; ו, כו – ז, כז; JTS Lutzki 51/16–17 – שמות יב, לא – יג, טו; כב, כט – כג, לג.
5. Geneva 99 (שמונה דפים) – שמות יא, א – טז, יג; JTS Lutzki 593/52 – שמות טז, לא – יז, יד.
6. Geneva 33 (דף אחד) – שמות כב, י – כד, ג; JTS Lutzki 530/58–61 – במדבר כג, כד – כז, א; לג, טז – לד, כד; דברים ד, א – לו.
7. AIU I A 170 (דף אחד) – ויקרא ח, כז – ט, ה; ENA 2976/1–2 – ויקרא כג, ד – לז; Strasbourg 4068/40–48 – דברים א, יא – ה, א.
8. JTS Lutzki 593/62 – במדבר ה, יד–ל; Geneva 108 (דף אחד) – במדבר ו, כא

- ז, יח; JTS Lutzki 593/36 – במדבר י, יז-כט; JTS Lutzki 515/34-35 – במדבר יח, יז – יט, ה; כג, ז – כד, ג.
9. – JTS Lutzki 503/34-35 – במדבר כ, כד – כב, טו; JTS Lutzki 530/28-29 – במדבר כו, מז – כז, יא; כט, יד – ל, ד; Geneva 146 (דף אחד) – דברים ב, לא – ג, יח.
10. Geneva 131 (שני דפים) – חבקוק א, יא – ג, ו; זכריה א, ו – ב, טו; JTS Lutzki 594/58 – זכריה ז, ז – ח, יט.
11. ENA 3756/1-2 – זכריה יא, ז – יב, ז; מלאכי ב, טו – ג, ד; והמשכו: JTS Lutzki 596/49 – מלאכי ג, ד-יב.
12. JTS Lutzki 594/74 (שני דפים) – תהלים י, יד – יז, ג; Geneva 89 (שני דפים) – תהלים פט, כו-נג; קב, טז – קג, יז.
13. שני טפסים בכתיבה זהה: Geneva 46 (מסורה); AIU X 13 (מקרא, מלכים ב' יז, ג – יח, ז).

הפטרות

14. הפטרות חד-שנתיות: Geneva 18 (שני דפים) – הפטרות מקץ, ויגש, ויחי; JTS Lutzki 515/74-75 – הפטרות כי תשא, תזריע.
15. הפטרות במחזור התלת-שנתי עם תרגום ארמי: Geneva 147 (דף אחד) – הפטרות לסדרי בראשית; (ENA 2103/5-6) JTS Lutzki 241/5-6 – הפטרות לסדרי בראשית; Vienna H 12 (דף אחד) – הפטרות לסדרי ויקרא. את הצירוף בקטלוג ז'נבה לדף מספריית וינה מסרנו אנו ל' עופר. דף זה אינו רשום ברשימתו של עופר לקטעי הגניזה של הפטרות במחזור התלת-שנתי.¹⁵ בקטלוג ז'נבה מצרף עופר לדפים מז'נבה ומווינה עוד שני דפים מספריית אוקספורד (MS Heb. b 2), דפים 12-13; מספר 2606.7 בקטלוג נויבאוואר-קאלי, ואנו מצרפים להם כאן שני דפים נוספים מספריית ניו יורק. הדפים מספריית ניו יורק רשומים ברשימה הנזכרת, אך עופר לא צירף בה באופן שיטתי את קטעי הגניזה לטפסים, לבד מהערות בודדות.

15 'עופר, 'סדרי נביאים וכתובים', תרביץ, נח (תשמ"ט), עמ' 186-189. לקטלוג ז'נבה ראה: אוסף הגניזה הקהירית בז'נבה – קטלוג ומחקרים, בעריכת ד' רוזנטל, ירושלים תש"ע.

תרגומים ארמיים

16. תרגום אונקלוס: Geneva 52 (שני דפים) – בראשית כ, ו-טו; כא, כ-לב; JTS Lutzki 605/1 – בראשית כב, כ – כג, ו; Budapest AS 16 (תשעה דפים) – בראשית כד, ח – כה, יח; כז, ה-כט.
17. תרגום אונקלוס (ופסאודריונותן): ENA 3347/3 – שמות ט, כה – י, א; Strasbourg 4017/1-8 – שמות כ, י-יג; תרגום פסאודריונותן כ, א-יג;¹⁶ אונקלוס כה, לא – כח, ח.
18. מקרא ותרגום אונקלוס: JTS Lutzki 503/15-16 – שמות יז, י – יח, יב; יט, כ – כ, יד; Geneva 117 (שני דפים) – ויקרא כב, ג-כא; כג, כ-לח.

תרגום רס"ג ותרגומים אחרים בערבית-יהודית

19. מקרא עם תרגום אונקלוס ותרגום רס"ג: HUC Acc. 961 (דף אחד) – בראשית א, כח – ב, ג; ENA 3186/11 – בראשית כז, יט-כז; ENA 3002/22-23 – בראשית ל, כו – לא, ח; JTS Lutzki 596/27 – שמות יז, ב-ז.
20. מקרא עם תרגום רס"ג: AIU I C 30 (דף אחד) – במדבר ח, כב – ט, ה; והמשכו: ENA 3190/6 – במדבר ט, ה-יב; והמשכו: ENA 2988/5-8 – במדבר ט, יב – י, יח.
21. תרגום רס"ג לתורה: AIU II A 125 (שלושה דפים) – דברים יד-טז; Budapest 204.2 (שני דפים) – דברים כא, ח – כב, ג; כו, טו – כז, יג; ביניהם: ENA 3094/12-13 – דברים כב, ג-כד; כו, ב-טו.
22. תרגום בערבית-יהודית לבראשית (לא מזוהה): ENA 2715/48-49 – בראשית ב, ז – ג, יט; ו, ז – ח, י; והמשכו: Halper 43 (דף אחד) – ח, יא – ט, כו.
23. תפסיר ירמיהו (לא מזוהה): Halper 49 (שני דפים) – ירמיהו ד-ה; ENA 2638/3-4 – ירמיהו ה-ו.
24. תרגום רס"ג לדניאל: Budapest 382 (שני דפים) – ג, כ-כט; ה, י-כא; ENA 3755/10-11 – ד, י-לב.

פירושי מקרא

25. פירוש רס"ג לתורה, ספר בראשית: ENA 2767/7 – פרק א, מהדורת מ' צוקר
- 16 M. L. Klein, *Genizah Manuscripts of Palestinian Targum to the Pentateuch*, Cincinnati 1986, pp. 273-274

- (פירושי רב סעדיה גאון לבראשית, ניו יורק תשמ"ד), עמ' 38–39 (רשום במהדורה); AIU VIII E 35 (דף אחד) – פרק א, מהדורת צוקר, עמ' 41–42 (רשום במהדורה, ללא זיהוי¹⁷).
26. פירוש רס"ג לתורה, ספר בראשית: ENA 3554/5 – פרק ב, מהדורת צוקר, עמ' 55–56 (רשום במהדורה); JNUL 577.5.5 (חמישה דפים) – פרקים ב, ג, מהדורת צוקר, עמ' 61–78 (אינו רשום במהדורה; רשום בקטלוג 'אלף', ושם צירופים לשני קטעים נוספים מספריית קמברידג'); AIU II B 14 (שני דפים) – פרקים ב, ג, מהדורת צוקר, עמ' 81–82, 91–92 (אינו רשום במהדורה ואינו מזהה בקטלוגים).
27. פירוש רס"ג לתורה, ספר ויקרא: ENA 2815/12 – יט, יט (לא זוהה); AIU II B; 79 (שמונה דפים) – יט, כו–כח (לא זוהה); ENA 2931/30–31 – יט, כח–ל; ENA 1696/9–10 – כ, יג–יד; והמשכו: ENA 3202/6 – כ, יד–כ.
28. פירוש רב שמואל בן חפני גאון לבראשית: Halper 47 (שני דפים) – בראשית לא–לב; ENA 2983/5–6 – בראשית לו–לז. קטע הלפר זוהה בקטלוג הלפר B. Halper, *Descriptive Catalogue of Genizah Fragments in Philadelphia*, (Philadelphia 1924), ושם צורפו אליו שני קטעים, מבודפסט ומקמברידג', שא' גרינבאום השתמש בהם במהדורתו (פירוש התורה לרב שמואל בן חפני גאון, ירושלים תשל"ט). גרינבאום לא השתמש בקטע הלפר ובקטע ניו יורק ויש בהם להשלים את המהדורה.
29. פירוש רב שמואל בן חפני גאון לתורה: ENA 2927/2–8 – במדבר כב; ENA 3187/4–5 – דברים ב–ג (זוהה כפירוש רשב"ח באתר גנזים בידי אפרים בן פורת); JNUL 577.5.8 (שישה עשר דפים) – דברים לג–לד. שלושה קטעים אלה כתובים באותה כתיבה ובאותו פורמט. פרשת בלק ופרשת ואתחנן פותחות שתיהן באותו סגנון עיטור.
- קטע נוסף באותה כתיבה ובפורמט זהה הוא פירוש על תהלים: AIU II B 138 (שני דפים).
30. בכתיבה זהה גם מילון מקראי בערבית יהודית: Halper 67 (שני דפים); HUC 1070 (שני דפים).
31. פירוש רס"ג לאיוב (כתאב אלתעדיל): AIU VIII B 49 (דף אחד) – סוף ההקדמה, פרק א; ENA 3515/1–2 – פרקים ח–ט (זוהה בידי אפרים בן פורת).
- 17 לקטע זה מכוון צוקר בעמ' 41–43, הערה 54 ('קטע כ"ח בפריס, דהה ומטושטש'), ללא מספר מדף. הקטע אינו רשום גם במפתח כתבי היד.

32. פירוש ישעיהו ודברי הימים בערבית-יהודית: Budapest 197 (שני דפים) – ישעיהו לז–לח, מ; ENA 3349/7 – ישעיהו מא–מב; ENA 2758/3 – דברי הימים א' ו.
33. פירוש רש"י למקרא: AIU II B 238 (שני דפים) – בראשית א, כ – ג, ח; והמשכו: AIU II A 9–10 (שבעה דפים) – בראשית ג, ח – יב, ח; ENA 1180/55–56 – במדבר כז, יח – כט, לה; דברים יג, ח – טו, א.
34. טופס אחר: AIU II B 248 (שני דפים) – שמות א, ה – ב, א; יב, מ – יג, ט; ENA 960/16–17 – ויקרא י, יד – יא, ח; טו, כג – טז, יד.
35. פירוש רמב"ן לתורה: ENA 1179/51 – ויקרא יח; AIU II B 97 (דף אחד) – ויקרא כו; ENA 1069/36–39 – במדבר לב–לג; ENA 1069/43 – במדבר לה; ENA 1179/42 – דברים (תחילת החומש); ENA 1180/36 – דברים יג; ENA 2905/4 – דברים לג.
36. פירוש רבנו בחיי בן אשר לתורה: ENA 1179/20–21 (דפים קיד–קטו) – בראשית כה, כא–כד; AIU II B 246 (דף קנד) – בראשית לב, יד–כה.
37. טופס אחר: ENA 3612/5–11 (שבעה דפים, קרעים) – פרשות מקץ, ויגש; AIU XI 94 (שני קרעים משני גיליונות כפולים [ביפוליה]) – פרשת ויגש. אחד הקרעים מקטע פריז שייך לדף שממנו נקדע ENA 3612/10.
38. פירוש הר"ד לתהלים: AIU II A 56 (ארבעה עשר דפים) – תהלים ו–מא; ENA 2996/9–18 – תהלים נה–ע.
39. פירוש תהלים (לא מזוהה): ENA 960/21–23 – סד–סה, פב–פג; AIU II B 222 (דף אחד) – קלה–קלו.
40. פירוש שיר השירים ליוסף אבן עקנין: התגלות הסודות והופעת המאורות (כתאב אנכשאף אלאסראר וט'הור אלאנואר): AIU II B 111 (שני דפים) – פתיחה, א, ב (כולל חיסרון שבמהדורת הלקין, עמ' 22, שורה 27); ENA 2993/4–7 – ח, יב; ENA 1069/8 – ח, ד.
41. פירוש רד"ק ליחזקאל פרק א על דרך הנסתר (מקראות גדולות, מהדורת הכתר, בעדיכת מ' כהן, רמת גן תש"ס, עמ' 329–333): ENA 2151/78–79 – א, כ, כו; Halper 57 (שני דפים) – א, כא–כב, כד–כה; ENA 2892/1–2 – א, כב–כד. שישה דפים רצופים בשלושה גיליונות.
42. מראות אלהים לחנוך אל-קונסטנטיני (מהדורת ק' סיראט, ירושלים תשל"ו): ENA 2710/59–60 – עמ' 19–20, 49–51; AIU VI B 121 (שני דפים) – עמ' 22–25, 48–49; ENA 2585/16–17 – עמ' 25–27, 43–45; AIU II B 123 (דף אחד) –

- עמ' 48-49. ארבעה גיליונות רצופים. באחד הגיליונות (AIU II B 123) חסר דף אחד.
43. ספר זכרון לרבי אברהם בקראט (על פירוש רש"י לתורה, מהדורת מ' פיליפ, פתח תקוה תשמ"ה): AIU II A 3 (עשרים ושניים דפים) – בראשית, עמ' צז-קט; קנד-קסא, קע-קעז; שמות, עמ' רכו-רלג, רסד-רסח, רפו-רצא, רצג-רצז; ויקרא, עמ' עה-פ; במדבר, עמ' קי-קכ, קכב-קכו (נרשם על פי ח' מיליקובסקי); ENA 2147/42-44 – בראשית, עמ' פח-צב, קלג-קלו, קסה-קסט; ENA 2572/9A-9B (שני דפים) – בראשית, עמ' קי-קיג, קכא-קכז; ENA 3167/3 – בראשית, עמ' קעז-קפ; ENA 2751/48-49 – שמות, עמ' רנא-רנו, רנח-רס; ENA 2852/36-37 – שמות, עמ' רסט-רעה. מלבד קטע פריז הגדול, שאר הקטעים לא זוהו כספר זכרון.
44. מדרשי התורה לשלמה אשתרוק (ספרד, המאה הארבע עשרה): ENA 3003/13 – פרשת שמות; AIU II A 35 (שני דפים) – פרשות ויקרא, צו, שמיני, תזריע; AIU XI 10 (שני דפים) – פרשות מצורע, אחרי מות. בהוספות רבות על מהדורת ש' עפנשטיין, ברלין תרנ"ט.
45. תורת המנחה לרבי יעקב ב"ר חננאל הסקילי (מהדורת ב"א חפץ, צפת תשנ"א): ENA 2991/10-11 – עמ' 426-432 (תזריע); והמשכם: ENA 2991/8 – עמ' 435-433 (מצורע); ENA 3515/3 – עמ' 442-444 (אחרי וקדושים); והמשכו: ENA 1069/31 – עמ' 447-444 (אחרי וקדושים); והמשכו: ENA 3033/13 – עמ' 450-447 (אחרי וקדושים – אמור); והמשכו: AIU II B 82 (שני דפים) – עמ' 453-450 (אמור), 458-456 (אמור – בהר ובחוקותי); ביניהם: AIU VI B 150 (דף אחד) – עמ' 453-456 (אמור); AIU II B 243 (דף אחד) – עמ' 526-523 (שלח – קרח); AIU II B 93 (דף אחד) – עמ' 569-570 (דברים). אחד עשר דפים בתשעה קטעים. מתוכם רצופים דפים: א-ג, ד-ט. מתוך תשעה קטעים, רק שלושה זוהו נכונה.
46. ילקוט לתורה (בראשית, במדבר): ENA 3002/11-14 – בראשית כו-כז; ENA 1179/27-28 – בראשית כו-כז; Vienna H 169 (שני דפים) – בראשית כז-כח; ENA 2915/3 – בראשית מד; ENA 2989/11 – במדבר כב. דומה לפירוש רש"י ולמדרש לקח טוב.

מדרש רבי דוד הנגיד לתורה

- בחיבורה של אלה אלמגור, כתבי היד של מדרשי ר' דוד הנגיד, ירושלים תשנ"ה, נרשמו מאות קטעים של החיבור שצורפו לעשרות טפסים. להלן מספר השלמות לרשימות הללו על פי צירופים שהציעה התוכנה. רשמנו רק את הקטעים שלא נרשמו אצל אלמגור.
47. עותק יט: ENA 2710/38-39 – פרשות קדושים, בהר;¹⁸ ENA 2987/7 – פרשת שופטים.
48. עותק יט/A: ENA 2710/66-67 – פרשות כי תשא, ויקהל; ENA 2585/6-7 – פרשות כי תשא, ויקהל.
49. עותק כ: ENA 2707/28-29 – פרשת חיי שרה; AIU II B 302 bis (דף אחד) – פרשת נשא; ENA 2638/36 – פרשת שלח לך; ENA 2709/47 – מדרש איכה.
50. עוד בכתיבת אותו סופר, אבל בטופס אחר, בפורמט קטן יותר: AIU II B 269 (דף אחד) – פרשת וירא; AIU XI 257 (שני דפים) – פרשת וירא.
51. עותק כא: ENA 2931/17-22 – פרשת נח, הפטרת נח; ENA 2707/4-7 – פרשות חיי שרה, תולדות; ENA 2707/33-34 – פרשת ויגש, הפטרה לשבת ראשונה של חנוכה.
52. עותק לג: ENA 2123/2-3 – פרשות שופטים, כי תצא; AIU IV B 55 (שני דפים) – ליום שני של ראש השנה. עוד שייכים לעותק זה ככל הנראה מספר קטעים מתוך מדרש איכה: JNUL 577.4.82 (שני דפים); ENA; ENA 2707/9-10; ENA 2709/42-43

ליטורגיקה

53. סידור רס"ג, סליחות לתענית אסתר: Budapest 120 (שני דפים) – עמ' שלו-שלו (במהדורת דודזון, אסף ויואל, ירושלים תש"א); והמשכו: ENA 2923/21 – עמ' שלו-שלו.
54. סידור רס"ג, סדר תעניות וסליחות לתשעה באב: Halper 223 (שני דפים) – עמ' שיז-שיח, שכב; המשך הדף הראשון: HUC 1066 (דף אחד) – עמ' שיח-שיט; והמשכו: ENA 3432/4-5 – עמ' שיט-שכא. קטע ניו יורק רשום במהדורת דודזון.

55. הבקשה השנייה של רס"ג: ENA 3019/14-15; JTS Lutzki 596/29 (שני דפים מחוברים).
56. קובץ תוכחות: ENA 3448/11-12 – מס' (בכתב היד) [ריט], רכ, [רכא]; AIU IV C 463 (דף אחד) – מס' [רלא]-רלג; AIU IV C 349 (שישה דפים) – מס' [רלג]-רמא; AIU IV C 183 (ארבעה דפים) – מס' [רמד]-רמו.
57. קובץ פיוטים: ENA 3452/3-4 – מס' (בכתב היד) [ג-ו], [יד-יו]; ENA 2629/50-51 – מס' פה-פו, קי-קיא; AIU IV C 12 (דף אחד).
58. קדושתאות לר' יהודה: Geneva 85 (דף אחד) – קדושתא לסדר 'והיה כי תבוא'; ENA 2048/1 – קדושתא לפרשת האזינו.¹⁹
59. פיוטים וסליחות ליום הכיפורים: Geneva 76 (שני דפים) – ברכי נפשי לרס"ג; Halper 247 (שישה דפים) – סליחות ליום הכיפורים, קצתן לרס"ג.
60. פיוטים לימים נוראים: Halper 239 (שני דפים); ENA 2675/1; ENA 3424/3-4. דף ENA 3424/4 ממשיך את הדף השני בקטע הלפר ('אמיצי שחקים ממעל').
61. פיוטים: ENA 589/1-4 – ברכי נפשי לרס"ג;²⁰ ENA 2704/16-17; Budapest 109 (דף אחד) – קדושתא של הקליר ליום הכיפורים;²¹ Budapest 113 (דף אחד).
62. בקשה: Halper 326-327 (שלושה דפים); ENA 2924/21; HUC 1060 (דף אחד). ציטוט: 'נפשי מה לך תחני בקברות התאוה...!'
63. פסוקי תחנונים: ENA 3466/13-14; JTS Krenzel 91 (דף אחד).
64. אזהרות 'שמור לבי מענה' לשלמה אבן גבירול: Halper 255-256 (שלושה דפים); HUC 1071 (שני דפים); הדף השני ממשיך את הדף השלישי באוסף הלפר²²;
והמשכו: ENA 3452/6
65. טופס אחד: ENA 1232/29-30; HUC 1203 (שני דפים).
66. פזמונים: ENA 3446/8-9; AIU IV C 478 (דף אחד).
67. סידור בערבית-יהודית. מבוסס על סידור רס"ג אך אינו זהה לו ושונה בעריכתו: HUC 1235 (שני דפים) – סוף ברכות קריאת שמע, שמונה עשרה, קדושה
- 19 ש' אליצור (קטלוג ז'נבה [לעיל, הערה 15], עמ' 78) צירפה את קטע ז'נבה עם קטע קמברידג'. לצירוף של כל הקטעים הידועים מטופס זה, כולל קטע ניו יורק, ראה: 'יהלום', 'בחזרה למחזור ארץ ישראל', קרית ספר, סג (תש"ן-תשנ"א), עמ' 200, הערה 23.
- 20 ראה: 'מרקוס, גנזי שירה ופיוט, ניו יורק תרצ"ג, ס' כז (אבל ENA 960/1-2 הנזכר שם שייך לכ"י אחר); מ' זולאי (האסכולה הפייטנית של רס"ג, ירושלים תשכ"ד, עמ' קי, כ"י 39) לא ראה את הקטע.
- 21 ראה: א' שייבר, גנזי קויפמן, בודפשט תש"ט, עמ' 7 (כ"י ל).
- 22 אכן, הוא פותח בטעות בהעתקה חוזרת של החרוז 'ולא תאכל מעשר דגנך בשערים'.

- דסדרא; ENA 694/20 – תפילות יום הכיפורים ותשעה באב (השווה: סידור רס"ג, עמ' שיז-שיט), ולאחריהן תחילת הלכות ברכות; והמשכו: AIU IV B 40 (דף אחד) – ברכות הנהנין (השווה: סידור רס"ג, עמ' פב-פד); ENA 907/4 – ברכת המזון (השווה: סידור רס"ג, עמ' קג-קט).
68. סידור בערבית־יהודית: JTS Krengel 87 (דף אחד) – ערבית של שבת; AIU IV A 257 (דף אחד) – הגדה של פסח; ENA 695/10-11 – הגדה של פסח.
69. פרק שירה ותפילות: AIU IV A 194a (ארבעה דפים); ENA 2485/1-2

פרשנות תלמוד והלכה

70. פירוש לספרא, פרשת בהר: AIU III C 52 (דף אחד); והמשכו: ENA 1418 (שישה דפים).

פירוש המשנה לרמב"ם: קטעי משנה עם פירוש המשנה לרמב"ם כבר צורפו ב'אוצר כתבי היד התלמודיים', אך למרות העבודה הרבה שהושקעה באוצר, התוכנה תרמה צירופים חדשים שלא נרשמו בו. בין השאר, התברר כי שנים עשר קטעים שהיו רשומים באוצר בשלושה טפסים שונים – טופס של שבעה קטעים, טופס של שלושה קטעים וטופס של שני קטעים – שייכים כולם לטופס אחד. מכיוון שצירופים אלה עודכנו באוצר, העומד להתפרסם בקרוב, לא רשמנו אותם כאן שנית.

71. פירוש רשב"ם לתלמוד, מסכת בבא בתרא: AIU II B 305 (דף אחד) – דפים צ-צא; ENA 1180/42-43 – דפים קלב-קלג; והמשכו: AIU III C 8 (דף אחד) – דף קלג. בקטלוג 'אלף' רשומים תחת הקטע האחרון צירופים לשני קטעים נוספים בספריית מנצ'סטר.
72. ספר המצוות לרס"ג: ENA 3552/12-13 – פרקים א-ב, ה-ו; Geneva 38 (שני דפים) – פרקים טז-יז, יט.
73. משנה תורה לרמב"ם: ENA 2632/27-30 – עבודת יום הכיפורים ה, ט-כג; מעילה ב, יד – ג, ו; Strasbourg 4104/23-24 – פסולי המוקדשין ח, י – ט, ה; י, ג – יא, ח.
74. משנה תורה לרמב"ם, כתיבת שלמה הלוי בן הדיין שמואל בן סעדיה הלוי: ENA 3098/1-2 – איסורי ביאה כא, שחיטה ז; Vienna H 101 (שמונה דפים) – איסורי ביאה טז-כב (לא רצופים, מלבד דפים 5-6), שחיטה ו-ח (לא רצופים).

75. מגיד משנה, ספר קניין וספר משפטים: ע' שבט רושם בקטלוג 'אלף' בשני מקומות (AIU III B 128, AIU III C 47) צירופים לטופס זה המסתכמים בעשרים וארבעה קטעים: חמישה מאוסף ENA, שניים מאוסף AIU, ארבעה עשר מספריית מנצ'סטר, שניים מספריית קמברידג' ואחד מן הספרייה הבריטית. התוכנה צירפה את שני הקטעים מאוסף AIU ושלושה קטעים מאוסף ENA שרשם שבט, והוסיפה להם שלושה קטעים מאוסף זה, אחד מהם שאף אינו רשום בקטלוג דנציג (לעיל, הע' 7) (ENA 3034/3-4).²³ טופס זה כולל אפוא לפחות עשרים ושבעה קטעים. ואלה הקטעים שלא רשם שבט בקטלוג 'אלף': ENA 1179/43-50 – טוען ונטען, נחלות; ENA 2120/11 – מלווה ולווה; ENA 3034/3-4 – מלווה ולווה.
76. טופס אחר של מגיד משנה: AIU III B 284 (דף אחד) – הקדמה לספר נזיקין; ENA 3039/15 – שכירות א, א – ב, א. בקטלוג דנציג נרשם קטע נוסף המצטרף לטופס: ENA 3001/2 (גזילה ז).
77. שרידיו של טופס אחד של תשובות הרא"ש הגיעו אלינו בקטעי גניזה רבים, ראה קטלוג 'אלף', ENA 2631. קטע זה כולל גם קולופון של הסופר: 'אני יעקב קורקוס כתבתי... בשיבת ה'ה'ה' ר' יצחק אבוהב יצ"ו וסיימתיו ביום ל"ו לעומר שנת חמשת אלפים ומאתים וחמשים ואחד לפרט היצירה בוואדאלחאג'ארה'. מכתבת ידו של סופר זה יש בידינו קטעים של כתב יד נוסף, פירוש הרא"ש או הר"ש משאנץ למשנה סדר זרעים: ENA 3307/9; AIU III A 97 (ארבעה דפים). ע' שבט²⁴ רושם דף נוסף מספריית מנצ'סטר המצטרף לקטע ניו יורק. כמו כן רשומים בקטלוג 'אלף' תחת קטע כ"ח עוד שני קטעים המצטרפים אליו: דף באוקספורד ודף בספרייה הבריטית. סך הכול הגיעו לידינו מטופס זה שמונה דפים בחמישה קטעים. את הצירוף של קטע ניו יורק לקטע כ"ח ואת חיבורם לטופס תשובות הרא"ש עשתה התוכנה.
78. רבנו ירוחם, ספר מישרים: דנציג רשם בקטלוג תחת הקטע ENA 1179/59 שלושה קטעים נוספים המצטרפים אליו, סך הכול שמונה דפים מספר מישרים. יש להוסיף להם: ENA 3034/2 – מישרים, נתיב א, חלק ג; AIU III C 55 (שני דפים) – נתיב א, חלק יא-יב; נתיב ב, חלק ח; נתיב ג, חלק א.
79. תשובות הרדב"ז: AIU III D 35 (דף אחד) – תשובות [תתפג]-תתפו (בכתב
- 23 מתוך שמונה קטעי ENA, דנציג בקטלוג (לעיל, הערה 7) מצרף רק שני קטעים (ENA 1069/21, ENA 1180/57).
- 24 לעיל, הערה 7.

- היד). עזרא שבט פרסם שתי תשובות מקטע זה,²⁵ וצירף אותו לכתב יד מאוסף פירקוביץ EVR II A 80 (דפים 49) ולדף נוסף באוסף זה, EVR II A 746. התוכנה צירפה שני דפים נוספים מאותו טופס: ENA 2989/12-13 – [תקמה]–תקמז. טופס אחר באותה כתיבה: AIU III D 20 (שני דפים) – [תמב]–תמו; ENA 2639/18 – תשובה בעניין גטין, לא נדפסה.
80. באותה כתיבה יש גם פירוש על משנה תורה לרמב"ם הלכות עבודה זרה: ENA 3309/1 – עבודה זרה ז, ה-ז; AIU XI 288 (דף אחד) – עבודה זרה ט, ה-ו. מסתבר שזהו פירוש הרדב"ז. רמז לדבר הביטוי הנמצא בקטע ניו יורק 'והנכון אצלי...!', ביטוי הרגיל במיוחד אצל הרדב"ז.²⁶

קבלה

82. שערי אורה לרבי יוסף ג'יקטיליה: AIU VI B 131 (שמונה דפים) – שער שני; ENA 1918/1-6 – שער חמישי. בכתיבת אותו סופר נמצא גם קטע ממערכת האלהות המיוחס לר' פרץ בן יצחק גירונדי: ENA 2974/1-4 – פרק שמיני.
83. זוהר: AIU VI A 35 (שלושה דפים?) – שמות; ENA 1180/31-32 – במדבר.
84. זוהר במדבר: AIU VI A 43 (שבעה דפים); ENA 3310/5.
85. אדרא זוטא: AIU VI A 9 (שמונה דפים) – מס' דפים (בכתב היד) ה-ו, ט-י, יג, יח, כא-כב; ENA 3175/5-8 – דפים כז, כח-כה; ENA 2985/3-4 – דפים מב-מג.
86. כתבי רבי חיים ויטאל (מאמרים מעץ חיים, שער מאמרי רז"ל, שער הגלגולים): AIU VI B 134 (ארבעה דפים); ENA 2692/33-36.
87. 'פירוש האדרא קדישא היא הקדמת האדרא זוטא' לרבי חיים ויטאל (כנדפס בסוף ספר אדם ישר, קראקא תרמ"ה): ENA 3175/1-4 – דפים רכה-רכח; AIU XI 210 (ארבעה דפים) – דפים רלג-רלו; והמשכו: ENA 2585/60-63 – דפים רלז-רמ.
88. כוונות הנענועים בהלל: ENA 2207/1-4; AIU VI B 47 (דף אחד).

25 ע' שבט, 'קובץ תשובות הרדב"ז', ירוק, יב (תשס"ג), עמ' פ-פג (מצוין שם: 'אוסף כ"ח 53'); יג (תשס"ד), עמ' סה-סט; יד (תשס"ד), עמ' עג.

26 חיפוש באמצעות פרויקט השו"ת מגלה שהביטוי מופיע בתשובותיו עשרים וארבע פעמים מתוך סך שלושים וארבעה מופעים בכל ספרות השו"ת. הביטוי חוזר גם בפירושו לספר זרעים (כלאים י, כב; מתנות עניים א, יא; תרומות י, ג; שמיטה ויובל ח, כ; יב, טו) ולהלכות מאכלות אסורות (יב, א; יד, ט).

89. בקשה 'אשבח למלך', מקור מים חיים לר' חיים העזרי (מחכמי תימן): AIU IV A 199 (עשרה דפים רצופים) – דפים 1-6: בקשה; דפים 6-10: מקור מים חיים; ENA 2964/17 – מקור מים חיים.

חכמת הלשון

90. ספר הערוך: Halper 95-96 (שני דפים) – ערכים אוב-אביי; והמשכו: ENA 1490/3 – ערך אביי;²⁷ ENA 960/5 – ערכים אביונות-אבל-אובלא.
91. אלמרשד אלכאפי לתנחום הירושלמי: ENA 2708/50 – ערך זרע; JNUL 577.6.27 (דף אחד) – ערכים הנם-הותיים. קטע ניו יורק זוהה אצל א' ממון, אוצרות לשון, ניו יורק וירושלים תשס"ו. שני הקטעים לא נרשמו במהדורת ה' שי, ירושלים תשס"ה. נראה שקטעים אלה שייכים לטופס האוטוגרף שרשמה שי. אצל ממון, שם, נרשמו עוד שלושה עשר קטעים מחיבור זה באוסף הגניזה שם, ומתוכם רשום אצל שי קטע אחד בלבד (4-3/3003 ENA). ייתכן שקטעים נוספים מאלה שרשם ממון מצטרפים אל שני הקטעים שרשמנו.
92. גלוסר בערבית-יהודית למשנה סדר זרעים, מועד: JNUL 577.7.4 (שלושה דפים) – ברכות, מעשרות, חלה, שבת; 10-9/2752 ENA – דמאי, כלאיים, שביעית (ראה תיאורו אצל ממון, שם).
93. גלוסר בערבית-יהודית למשנה סדר טהרות: 6-5/2580 ENA – מסכת כלים (ראה תיאורו אצל ממון, שם); 52.4.577 JNUL (שני דפים) – מסכתות פרה-טהרות.
94. חיבור על טעמי המקרא בערבית-יהודית: AIU IX A 10 (שלושה דפים); ENA NS 70/22 (דף אחד).
95. מילון חרוזים: AIU IX A 33 (דף אחד) – ין-יש; 5-1/3042 ENA – לים-לום, מה-מים, מע-מוש, ני-נים; 2-1/3178 ENA – נה, נים (ממשיך את ENA 3042/5); 137 C IV AIU (דף אחד) – נה-ני (ממשיך את ENA 3178/1). ממון, אוצרות לשון, רשם את שני קטעי ניו יורק אך לא צירף ביניהם. עם זאת, הוא צירף לקטע 5-1/3042 ENA את הקטע 19-18/3097 ENA, המכיל את חרוזי לי-לו.

27 הקטע זוהה בקטלוג דנציג (לעיל, הערה 7) כאיגרת רב שרירא גאון השנייה. למעשה, מדובר בקטע מן הערך אביי של הערוך' שבו כלולה איגרת זו.

תאולוגיה ומוסר

96. דלאלה אלחאירין (מורה נבוכים) לרמב"ם: 3-2/3342 ENA – ג, יב; Halper 430 (דף אחד) – ג, יב-יג.
97. כתאב כפאיה אלעאבידין (המספיק לעובדי ה') לרבי אברהם בן הרמב"ם: בקטלוג דנציג (לעיל, הערה 7) רשומים מספר קטעים מחיבור זה השייכים לטופס אחד: ENA 2859/31-33; ENA 3630/5-6; ENA 3774/11; ENA 3789/4 אליהם קטע נוסף מאוסף כ"ח: 78 AIU II B (דף אחד). קטע זה זוהה כ'כתאב אלכפאיה' גם אצל י' פנטון.²⁸ התוכנה רשמה קטע נוסף באותה כתיבה, גם הוא מאותו חיבור, אם כי בפורמט גדול יותר: ENA 2976/13.
98. שערי תשובה לרבי יונה גירונדי: 20 AIU V B (שבעה דפים) – שער א, ד-לא; ENA 2711/61-62 – שער א, מט – שער ב, ב; שער ב, יח-כא; AIU V A 31 (דף אחד) – שער ג, ה-ט.
99. דרשיח בין הגוף לנשמה, וספר יחוד הייחוד (אוטוגרף?): 18 ENA 2585 – קטע מחיבור לא מזוהה, כתוב במליצה מחורזת כדרשיח בין הגוף לנשמה. סיום החיבור נמצא בקטע 49 AIU IV B (דף אחד), ושם קולופון: 'תם ונשלם... שנת הרצ"ג ליצירה על ידי הצעיר יוסף בן החסיד כה"ר יצחק דונדי נ"ע'.²⁹ העמוד השני פותח: 'ספר יחוד הייחוד. מאשר יקרת בעיני נכבדת... כללים קצרים... א' בסוד הקדיש ב' נכלל עמו סוד אמן בכל מקום...! גם הוא כתוב במליצה מחורזת, בסגנון דומה לחיבור הקודם.
100. דרשות תלמודיות(?) בלאדינו: באתר 'גנזים' רשום צירוף של אלדינה קינטנה לארבעה קטעים באוסף JTS: 11/2853 ENA; 35/2716 ENA; 24/2709 ENA; 30-29/2943 ENA. לקטעים אלה צירפה התוכנה קטע נוסף מספריית כ"ח: AIU IX B 1 (שלושה דפים).
101. חיבור תאולוגי בערבית יהודית: 33/2852 ENA; 54/2712 ENA; 291 AIU IV A (שני דפים).
102. חיבור על מוסר ותאולוגיה בערבית יהודית: 25-24/2983 ENA; 7 AIU VII B (שני דפים). את קטע ניו יורק תיארה איילה אליהו באתר 'גנזים'. לפי תיאורה,
- 28 P. Fenton, 'Dana's Edition of Abraham Maimuni's "Kifayat al-Abidin" [Review]', *Jewish Quarterly Review*, 82 (1991), p. 198
- 29 'יוסף בכ"ר יצחק דונדי ס"ט' חתום על איגרת בני הישיבה מירושלים משנת רפ"א כתלמיד הישיבה. ראה: א' יערי, אגרות ארץ ישראל, רמת גן, 1971, עמ' 166.

- ייתכן ששני הדפים שרדו משני חיבורים שונים. הקטע מפריז לא זוהה. מכל מקום, ברור שכל הדפים שייכים לאותו טופס.
103. כלאם מוסלמי: ENA 3166/13-14; Budapest GEN 405 (שני דפים). שני הקטעים מתוארים באתר 'גנזים': איילה אליהו תיארה את קטע ניו יורק, וכריסטינה סילדג' את קטע בודפסט. סילדג' צירפה לקטע בודפסט קטע נוסף מספריית בודפסט השייך לחיבור: GEN 211 (שני דפים).
104. טופס אחר שכתב אותו סופר צירפה התוכנה מסידור תפילות קראי, והדבר מלמד כי חיבור הכלאם המוסלמי מגיע מספרייה קראית דווקא: Halper 225-226 (שני דפים); JNUL 577.4.96 (שני דפים).

ספרות יפה

105. ספר יוסיפון בערבית יהודית: Strasbourg 4038/7-8 – רשום אצל ש' סלע, ספר יוסף בן גוריון הערבי, ירושלים תשס"ט, עמ' 543 (דף 8 אינו רשום שם, אך היא השתמשה בו במהדורה); ENA 3316/12 – אינו רשום אצל סלע.
106. סירת ענתר (סיפור ענתר בערבית יהודית): AIU VII B 1 (תשעה עשר דפים); ENA 2713/27; ENA 2713/44-45; ENA 2852/34-35; Halper 330 (דף אחד).

מחברות (אוטוגרפים)

107. מחברת דרשות (אוטוגרף): ENA 3045/12; ENA 2365/27-28; AIU III B 99 (שני דפים, דף אחד חלק).
108. מחברת דרשות, הספדים ועוד (אוטוגרף): AIU II A 76 (שלושים ושניים דפים); ENA 2365/25; ENA 3770/1-2
109. מחברת עם רשימות עזר לדרשן: ילקוט פסוקים, מאמרים, משלים ופתגמים, רשימות ממוספרות של פסוקים ורמזים לדרשות: AIU II A 85 (שלושה עשר דפים); ENA 2709/34-35; ENA 3042/7-8
110. מחברת חידושים לתלמוד מסכתות שבת, כתובות, בבא קמא: AIU III B 105 (שישה דפים) – שבת פרקים א, ג, ד. בראש הקטע תאריך: 'ר"ח שבט שנת רצ"ו' (1586); ENA 2365/49-50 – שבת פרק ט; ENA 3178/5-6 – שבת פרקים ט-יד; AIU III B 131 (שני דפים) – כתובות פרק ג; ENA 3174/19-20 – בבא קמא פרק א.

111. מסופר זה ישנו טופס נוסף, של פירושים לחומש בראשית: AIU III C 13 (ארבעה דפים) – בראשית כב-כג; ENA 3192/5 – בראשית כח-ל.
112. מחברת חידושים לתלמוד בבלי ותשובות? (אוטוגרף): ENA 2645/1-10 – סנהדרין, קידושין; ENA 2692/22-25; AIU IV B 12 (שני דפים); AIU III B 182 (שני דפים); AIU III C 31 (שני דפים); AIU III D 28 (שמונה דפים). בקטע AIU III D 28 יש חתימה: 'הצעיר יעקב בן לא"א יוסף לונדוס(?)'. בקטע זה גם נכתב: 'וז"ל הרב חיים שבתי ז"ל...'. ייתכן שהכותב הוא תלמיד הרב חיים שבתי (מהרח"ש), ראש ישיבה בסלונקי במחצית הראשונה של המאה השבע עשרה.

תעודות

- בין הצירופים הרבים שזיהתה התוכנה יש לציין גם צירופים של קטעי תעודות שהדמיון בכתיבת היד שלהן מזהה אותן כצירופים חלשים. מכיוון שרוב התעודות הללו כבר זוהו במחקר ובקטלוגים, לא ראינו צורך לרשום אותן כאן. נציין, כדוגמה בלבד, מספר תעודות של הסופר חלפון בן מנשה שהתוכנה צירפה ביניהן. אף שלא כל התעודות שזיהתה התוכנה ככתיבת יד חלפון בן מנשה רשומות בעבודתו של ג' וייס,³⁰ הן כבר זוהו בקטלוג 'אלף'. כן נציין שתי תעודות של עמרם בן יוסף שהתוכנה צירפה ביניהן, והן פורסמו ממש לאחרונה בספר הודו.³¹ בתעודות מעטות אחרות שטרם פורסמו, חשבנו שלא יהיה אחראי להסתמך על הדמיון בכתיבת היד בלבד מבלי לפענח אותן בשלמות ולעמוד על עניינן. אף על פי כן נרשום כאן תעודה אחת, הואיל והדמיון בכתב היד נראה לנו מובהק דיו:
113. אפרים בן שמריה: התעודות בכתיבת ידו של אפרים בן שמריה, מנהיג קהילת הירושלמים בפסטאט במחצית הראשונה של המאה האחת עשרה, נרשמו ופורסמו אצל א' ברקת, יהודי מצרים 1007-1055, ירושלים תשנ"ה. התוכנה זיהתה תעודה נוספת באותה כתיבה שאינה רשומה שם: AIU IV C 208. כמדומה שתעודה זו עדיין לא התפרסמה ולא נדונה במחקר.

30 G. Weiss, *Legal Documents Written by the Court Clerk Halfon ben Manasse (Dated 1100-1138)*, Ph.D. Thesis, University of Pennsylvania, 1970

31 S.D. Goitein and M.A. Friedman, *India Traders of the Middle Ages (India Book)*, Leiden 2008, Doc. II,5; II,6

מפתח קטעי הגניזה

Budapest, Hungarian Academy, Kaufmann Genizah Collection

109	61
113	61
120	53
197	32
204.2	21
382	24
AS 16	16
GEN 211	103
GEN 405	103

Cincinnati, Hebrew Union College

1060	62
1066	54
1070	30
1071	64
1203	65
1235	67
Acc. 961	19

Geneva, Public and University Library, Genizah Collection

18	14
33	6
38	72
43	1
46	13

52	16
76	59
85	58
89	12
99	5
101	4
108	8
117	18
131	10
146	9
147	15

Jerusalem, JNUL

577.4.52	93
577.4.82	52
577.4.96	104
577.5.5	26
577.5.8	29
577.6.27	91
577.7.4	92

New York, JTS

ENA 589/1-4	61
ENA 694/20	67
ENA 695/10-11	68
ENA 907/4	67
ENA 960/5	90
ENA 960/16-17	34
ENA 960/21-23	39

ENA 1069/8	40
ENA 1069/31	45
ENA 1069/36-39	35
ENA 1069/43	35
ENA 1179/20-21	36
ENA 1179/27-28	46
ENA 1179/42	35
ENA 1179/43-50	75
ENA 1179/51	35
ENA 1179/59	78
ENA 1180/31-32	83
ENA 1180/36	35
ENA 1180/42-43	71
ENA 1180/55-56	33
ENA 1232/29-30	65
ENA 1418	70
ENA 1490/3	90
ENA 1696/9-10	27
ENA 1918/1-6	82
ENA 2048/1	58
ENA 2103/5-6	15
ENA 2120/11	75
ENA 2123/2-3	52
ENA 2147/42-44	43
ENA 2151/78-79	41
ENA 2207/1-4	88
ENA 2365/25	108
ENA 2365/27-28	107
ENA 2365/49-50	110
ENA 2485/1-2	69
ENA 2572/9A-9B	43

ENA 2580/5-6	93
ENA 2585/6-7	48
ENA 2585/16-17	42
ENA 2585/18	99
ENA 2585/60-63	87
ENA 2631	77
ENA 2632/27-30	73
ENA 2638/3-4	23
ENA 2638/36	49
ENA 2639/18	80
ENA 2645/1-10	112
ENA 2675/1	60
ENA 2692/22-25	112
ENA 2692/33-36	86
ENA 2692/50-51	57
ENA 2704/16-17	61
ENA 2707/4-7	51
ENA 2707/9-10	52
ENA 2707/28-29	49
ENA 2707/33-34	51
ENA 2708/50	91
ENA 2709/24	100
ENA 2709/34-35	109
ENA 2709/42-43	52
ENA 2709/47	49
ENA 2710/38-39	47
ENA 2710/59-60	42
ENA 2710/66-67	48
ENA 2711/61-62	98
ENA 2712/54	101
ENA 2713/27	106

ENA 2713/44-45	106
ENA 2715/48-49	22
ENA 2716/35	100
ENA 2751/48-49	43
ENA 2752/9-10	92
ENA 2758/3	32
ENA 2767/7	25
ENA 2815/12	27
ENA 2852/33	101
ENA 2852/34-35	106
ENA 2852/36-37	43
ENA 2853/11	100
ENA 2859/31-33	97
ENA 2892/1-2	41
ENA 2905/4	35
ENA 2923/21	53
ENA 2924/21	62
ENA 2927/2-8	29
ENA 2931/17-22	51
ENA 2931/30-31	27
ENA 2943/29-30	100
ENA 2964/17	89
ENA 2974/1-4	82
ENA 2976/1-2	7
ENA 2976/13	97
ENA 2983/5-6	28
ENA 2983/24-25	102
ENA 2985/3-4	85
ENA 2987/7	47
ENA 2988/5-8	20
ENA 2989/11	46

ENA 2989/12-13	79
ENA 2991/8	45
ENA 2991/10-11	45
ENA 2993/4-7	40
ENA 2996/9-18	38
ENA 3001/2	76
ENA 3002/11-14	46
ENA 3002/22-23	19
ENA 3003/3-4	91
ENA 3003/13	44
ENA 3019/14-15	55
ENA 3033/13	45
ENA 3034/2	78
ENA 3034/3-4	75
ENA 3039/15	76
ENA 3042/1-5	95
ENA 3042/7-8	109
ENA 3045/12	107
ENA 3094/12-13	21
ENA 3097/18-19	95
ENA 3098/1-2	74
ENA 3166/13-14	103
ENA 3167/3	43
ENA 3174/19-20	110
ENA 3175/1-4	87
ENA 3175/5-8	85
ENA 3178/1-2	95
ENA 3178/5-6	110
ENA 3186/11	19
ENA 3187/4-5	29
ENA 3190/6	20

ENA 3192/5	111
ENA 3202/6	27
ENA 3307/9	77
ENA 3309/1	81
ENA 3310/5	84
ENA 3316/12	105
ENA 3342/2-3	96
ENA 3347/3	17
ENA 3349/7	32
ENA 3424/3-4	60
ENA 3432/4-5	54
ENA 3448/11-12	56
ENA 3452/3-4	57
ENA 3452/6	64
ENA 3466/13-14	63
ENA 3515/1-2	31
ENA 3515/3	45
ENA 3552/12-13	72
ENA 3554/5	26
ENA 3612/5-11	37
ENA 3630/5-6	97
ENA 3755/10-11	24
ENA 3756/1-2	11
ENA 3770/1-2	108
ENA 3774/11	97
ENA 3789/4	97
ENA NS 70/22	94
Krengel 14a-b	2
Krengel 87	68
Krengel 91	63
Lutzki 241/5-6	15

Lutzki 503/15–16	18
Lutzki 503/34–35	9
Lutzki 515/16–17	4
Lutzki 515/34–35	8
Lutzki 515/74–75	14
Lutzki 519/3–8	2
Lutzki 530/28–29	9
Lutzki 530/58–61	6
Lutzki 593/36	8
Lutzki 593/52	5
Lutzki 593/62	8
Lutzki 594/58	10
Lutzki 594/74	12
Lutzki 596/27	19
Lutzki 596/29	55
Lutzki 596/49	11
Lutzki 605/1	16

Paris, AIU

I A 59	1
I A 158	3
I A 170	7
I C 30	20
II A 3	43
II A 9–10	33
II A 35	44
II A 56	38
II A 76	108
II A 85	109
II A 125	21

II B 14	26
II B 78	97
II B 79	27
II B 82	45
II B 93	45
II B 97	35
II B 111	40
II B 123	42
II B 138	29
II B 222	39
II B 238	33
II B 243	45
II B 246	36
II B 248	34
II B 269	50
II B 302bis	49
II B 305	71
III A 97	77
III B 99	107
III B 105	110
III B 128	75
III B 131	110
III B 182	112
III B 284	76
III C 8	71
III C 13	111
III C 31	112
III C 47	75
III C 52	70
III C 55	78
III D 20	80

III D 28	112
III D 35	79
IV A 194a	69
IV A 199	89
IV A 257	68
IV A 291	101
IV B 12	112
IV B 40	67
IV B 49	99
IV B 55	52
IV C 12	57
IV C 137	95
IV C 183	56
IV C 208	113
IV C 349	56
IV C 463	56
V A 31	98
V B 20	98
VI A 9	85
VI A 35	83
VI A 43	84
VI B 47	88
VI B 121	42
VI B 131	82
VI B 134	86
VI B 150	45
VII B 1	106
VII B 7	102
VIII B 49	31
VIII E 35	25
IX A 10	94

IX A 33	95
IX B 1	100
X 13	13
XI 10	44
XI 94	37
XI 210	87
XI 257	50
XI 288	81

Philadelphia, University of Pennsylvania

Halper 43	22
Halper 47	28
Halper 49	23
Halper 57	41
Halper 67	30
Halper 95–96	90
Halper 223	54
Halper 225–226	104
Halper 239	60
Halper 247	59
Halper 255–256	64
Halper 326–327	62
Halper 330	106
Halper 430	96

Strasbourg, National Library

4017/1–8	17
4038/7–8	105
4068/27	3

4068/40-48	7
4104/23-24	73

Vienna, National Library, Reiner Collection

H 12	15
H 101	74
H 169	46

