פרוייקט השו"ת: עבר, הווה — וגם עתיד?

אביעזרי פרנקל, רחובות

התנצלות

מעולם לא הרגשתי צורך להעלות על הכתב את ההסטוריה של פרוייקט השו"ת. לפני מספר שנים אף השלכתי חליפות מכתבים ומסמכים רבים הנוגעים לפרוייקט, כאשר משרדי איים ליהפך למחסן גרוטאות. הדחף לכתוב השתלט עלי לפתע פתאום. הקש ששבר את גב הגמל היה כנראה השמועה שבמחצית השנה האחרונה נמנע מהפרוייקט לעסוק במחקר ופיתוח, עד קבלת התוצאות של עוד דיון ממושך באוניברסיטת בר-אילן על הפרוייקט. (הפרוייקט אכן הביא לבר-אילן הכנסה ניכרת בשנים האחרונות, לרגל המכירה המקומית והבינלאומית של תקליטור השו"ת.) הדברים הבאים אינם הסטוריה מסודרת, אלא שרטוט קוים אחדים לדמותו של הפרוייקט, ייחודו ופעלו; אלה שאני זוכר בעליל ואלה שמבוססים על מסמכים שעדיין מצויים באמתחתי. תמצית הפעילות בשנים האחרונות ודיון קצר על עתיד הפרוייקט מובאים בסיום.

צעדים ראשונים

לידתו של הפרוייקט היתה במיניאפוליס, מינסוטה, אה"ב, בתחילת 1963, לשם נסעתי לפי בקשתו של פרופ' חיים פקריס המנוח, ראש המחלקה למתימטיקה שימושית במכון ויצמן למדע, כדי לפקח על השלבים האחרונים של בדיקת מחשב CDC-1604 שנרכש ע"י המכון. בשבתות נהניתי לעתים קרובות מהכנסת האורחים הלבבית ומאירת הפנים של מר אירווינג קוטוף, שיום אחד שאלני "מה יכול מחשב לתרום ליהדות"? בסופו של הדיון הקצר שלנו חשבנו שייתכן וספרות השו"ת, החובקת משאבי זמן (17 מאות שנה) ומרחב (כל העולם) מרובים, תוכל אולי להיפתח לציבור הרחב באמצעות מערכת מחשב מתוחכמת.

לקחתי את הענין ברצינות, בעיקר בגלל העובדה שבאותו זמן "הפכתי את עורי" ממהנדס חשמל למתימטיקאי, והרגשתי צורך, בנוסף לעיסוק במחקר טהור, אותו אהבתי, לעסוק גם ביישובו של עולם — פעילות בעלת תועלת "מיידית". כבר בשלב ראשוני זה חזיתי בעיני רוחי מערכת שתוכל לענות ביעילות על שאלות המוגשות לה, ע"י הקשה במספר כפתורים.

למרות שאיחזור מידע (או שליפת מידע — Information Retrieval) לא משך אותי במיוחד, נכנסתי לענין בלהט נעורים. למדתי את השטח מיסודותיו, כגון התיזה של בר-הלל [BaH1962], בה הוא פטר איחזור מידע כבלתי מעשי באופן כללי, ועד הפרוייקטים האחרונים באיחזור מידע שהוקמו או תוכננו באותה עת ברחבי העולם. על אף שהסכמתי עם בר-הלל בדבר המורכבות המהותית של השטח, חשבתי שבכל זאת ניתן לתכנן מערכת ממוכנת מעשית.

בקיצור רב, התיזה שלי היתה (ונשארה) פשוטה: מערכת איחזור מידע יכולה "להתקיים מעשית" אם ורק אם החלקים שיכולים להיעשות "ביעילות" ע"י מחשב ימוכנו, אלה בהם האדם מצטיין ייעשו על ידו, ואילו אלה שאינם "מבוצעים ביעילות" ע"י שום צד יבוטלו כליל. בשלב מוקדם הרגשתי אינטואיטיבית שמפתות ומיון לפי נושאים שייכים לקטיגוריה האחרונה. לכן המלצתי על מערכת טקסט מלא, דהיינו, כל הטקסט מאוכסן במחשב ללא מפתות. שמחתי להיוכח שמערכת כזו היתה אז בשלבי הקמה בפיטסבורג

[Hor1962]. אכן מערכות איחזור שאינן מתישבות עם התיזה הנ"ל נעלמו כולן מהשטח בשלב זה או אחר. תוצאות הממצאים שלי פורסמו במאמר ארוך [Fra1968]. אגב, פרופ' בר-הלל המנוח שיגר אלי מכתב ברכה חם על העבודה הזו, אעפ"י שמסקנותיה עמדו בסתירה למסקנותיו הוא. (לאחר שפותחו כלים לבירור מהות המושג "מורכבות חישובית" בשנות ה-1970, הצלחתי להוכיח באופן פורמלי שהחלקים הקלים של מודלים טבעיים למפתוח וסיווג לפי נושאים, הם אכן בעלי מורכבות מותנית גבוהה: [Fra1981] NP-Completeness.)

שמעתי שהאוניברסיטה העברית הקימה בדיוק באותם הימים את "המכון לחקר המשפט העברי" (המלחמ"ע) בראשותו של דר. מנחם אלון, לאחר מכן המשנה לנשיא ביהמ"ש העליון, ובפיקוחה של מועצה מדעית בראשותו של פרופ' אפרים אורבך המנוח. היה זה אפוא טבעי להציג בפניהם את מימצאי, בתקוה שהם יאמצו את השיטה שלי, במגמה לממש את החלום של פתיחת ספרות השו"ת "לציבור הרחב", כך שכל החפץ ליטול כתרה של תורה יוכל לבוא וליטול. בערב ראש השנה התשכ"ד, Sept. 1963, לפני 33 שנים, הגשתי לדר. אלון מזכר, בו המלצתי לאמץ את שיטת הטקסט המלא ללא מפתות. המזכר גם כלל המלצות בדבר מפתוח אוטומטי, כי המלחמ"ע אימץ שיטת מפתוח ידני ואף החל בתנופת מפתוח רחבה בעזרת כ-15 עמיתי מחקר. קויתי שאנו רק נוכיח את תוקפה של השיטה שלנו, שתאומץ ע"י המלחמ"ע, ובזה יסתיים תפקידנו.

מתחילת העבודה היה ברור לי שנזדקק לכלים לשוניים ממוכנים יעילים לטיפול בעברית וארמית, הלשונות העיקריים — אם כי לא הבלעדיים — בהם כתובה ספרות השו"ת. וזאת בגלל האופי המיוחד של שפות אלה, עניות ב"ערכים" ("שרשים"), עתירות בנטיות, ובדרך כלל מחוסרות ניקוד [Fra1968]. למדתי על העבודה החלוצית של דר. יעקב שוויקה, לאחר מכן פרופסור למדעי המחשב באוניברסיטת בר-אילן, ומר מאיר שפירא [ChS1964], על האנליסה הממוכנת של מלים עבריות, המאפשרת לחבור יחדיו כל המלים הגזורות ממילה נתונה. יחד עם זאת הייתי משוכנע שעבור מערכת איחזור מידע מעשית בשיטת הטקסט המלא ידרש לנו גם כלי "הפוך", סינטיזה של מלים: בהינתן מילה, כלי זה ייצור את כל המלים שאפשר לקבל ע"י צירוף אותיות שימוש וכנויים לפי כללי הדקדוק והמורפולוגיה. מר מאיר כהן ממשרד הבטחון שהיה גם סטודנט באונברסיטת תל-אביב, כיום ממנהלי בל"ל, ואנכי, הרצינו על כך בכנס איל"א (האיגוד הישראלי לעיבוד אינפורמציה) באפריל 1968 (קובץ ההרצאות, תקציר, ע' 173). אכן פיתחנו כלי כזה [CCDFS1971]. לכל אחת משתי השיטות האלה חסרונות ויתרונות. לכן מיזגנו את שתיהן לשיטת אנליסה-סינטיזה משולבת: מבוצעת סינטיזה גלובלית על כל בסיס הנתונים, אך ללא אותיות שימוש, כדי לקבל בסיס נתונים בגודל שהוא עדיין סביר; אותיות השימוש נשארות לשלב הסינטיזה. עיין [AFCS1972] לפרטים מלאים.

גם הקימותי "ועדה למיכון בחקר המשפט העברי", שחבריה הראשונים היו, כמדומני, דר. שוויקה, מר אסא כשר, לאחר מכן פרופסור לפילוסופיה באוניברסיטת תל אביב, מר יוסף דוויק, עו"ד צעיר ועמית מחקר במלחמ"ע, ששימש כנציג המלחמ"ע אצלנו ואסיסטנטים. הועדה עסקה בניסוח שיטות לעריכת טקסטים לפני קליטתם, ויצירת אלגוריתמים שונים לעבודה ממוכנת. (רבים אחרים עברו בועדה לאחר מכן, כגון מר ראובן מירקין מהאקדמיה ללשון העברית וסטודנטים לתאר שלישי ושני.) גם הרגשתי צורך לשבץ בצוות מומחים למשפט עברי בעלי עניין במחשבים. באמצע 1966 הצטרף אלינו הרב מנחם סליי, שייצג עבורנו את נקודת המבט של הצרכנים, ומאידך הסביר להם את האפשרויות הגלומות במערכת. הוא תרם רבות גם בבקרתו הקונסטרוקטיבית. שנה לאחר מכן הצטרף הרב יצחק פצ'ניק המנוח, אשר בדרכו השקטה המיוחדת הוסיף רבות לאיכותה של המערכת. שניהם החלו את עבודתם במסגרת מכון ויצמן.

ב-1966 או 1967 כתב מאיר כהן את התכניות הראשונות עבור המחשב המרכזי באוניברסיטת בר-אילן. הן שימשו במשך שנים רבות כשלד של מערכת מעשית ראשונה לאיחזור מידע בטקסטים עבריים מלאים. מאיר הגיש את התוצאות כעבודה לתאר שני באוניברסיטת תל אביב, ואני שימשתי כמדריך חוץ. (לאחרונה מאיר הזכיר לי שארבעה מוסדות היו מעורבים בעבודה זו: מכון ויצמן, אוניברסיטת בר-אילן, האוניברסיטה העברית ואוניברסיטת תל אביב.)

קשיים תחיליים

בעיות תקציביות הטרידו אותנו תמיד, במדה זו או אחרת. בתחילה בקשנו תקציב מינימלי מהמלחמ"ע, אך הבטחותיהם בשטח זה תמיד מולאו רק חלקית, ובאיחור. קבלתי רשות ממכון התקנים האמריקאי, ממנו היה לי מענק מחקר בתורת המספרים, להעביר חלק מהתקציב לעבודותינו באיחזור מידע. (לאחר מכן קבלתי מהם עוד שני מענקי מחקר שהיו מיועדים לאיחזור מידע במילואם.) גם פרופ' פקריס עזר במימון חלקי של שוויקה, סליי, פצ'ניק, וכן ניקוב, על כרטיסי IBM, של שו"ת הריב"ש, הרא"ש והרשב"א. בדיוק עברה תחת ידי חליפת מכתבים מ-1970, ממנה אני למד שאז עדיין שוויקה, סליי וניקוב של שו"ת מומנו ע"י פקריס. גם ב-1973 המכון עדיין מימן את המשכורת של שני עמיתי מחקר, מכונת ניקוב ונקבנית.

בסתיו 1966 נטלתי חל"ת ממכון ויצמן למשך שנתיים, כדי להקים מחלקה למתימטיקה בבר-אילן; גם בשנים רבות לאחר מכן היה לי שם מינוי חלקי. אחד התנאים המוקדמים שלי היה שהמוסד ירכוש מחשב. תנאי זה מולא, אך הפעילות שלנו באיחזור מידע לא זכתה לאהדה. בטויים כמו "פרנקל משוגע" נאמרו באופן גלוי. היוצא מהכלל היה פרופ' א. הראל פיש, ראש המחלקה לאנגלית, שעמד אז בראש ועדת המחקר, ממנו זכינו ליחס אוהד ומבין. הוא עודדני להגיש הצעת מחקר לועדה שלו, במגמה להזרים לנו תקציב קטן. אח"כ הוא סיפר לי שכל הבודקים אמרו שהכלים לאיחזור ועיבוד לשוני לעברית לא קיימים עדיין, ושדרושה עבודת הכנה מוקדמת של שנים אחדות. בנובמבר 1966 קבלתי אפוא מכתב דחיה רשמי, שמנה "עובדות" אלה. (בדומה, אם כי בנושא אחר, כאשר באמצע מלחמת יוה"כ התיצבה במשרדי קבוצת אנשים מאחד המוסדות הבטחוניים ושאלוני על האפשרות להקמת מערכת מסויימת, הצגתי בפניהם, לאחר כמה ימי הרהור, שלד של תכנית להשלמה תוך שנה. כשביקרתי אח"כ במשרדיהם, נוכחתי שכבר כמה ועדות ישבו על מדוכה זו וכתבו שעבודת הכנה בת שנים רבות תידרש לפני שניגשים לתכנון המערכת. אנו השלמנו את המערכת תוך שנה. במשך השנים מערכת זו גדלה אכספוננציאלית תוך התיעלות יפה, והיא ממשיכה גם כיום לתרום רבות לבטחון המדינה.)

באותם ימים זכינו בעיקר לקיטונות של לעג. זכורני יוצא מהכלל: השופט יעקב בזק אמר לנו כמה מלים טובות של הערכה, וזה עזר לנו להתגבר על תחושת חוסר האונים והיאוש.

היחסים שלנו עם המלחמ"ע, שלא היה טובים מעולם, החלו להתדרדר. מצד אחד היינו זקוקים לתקצי-בים נוספים. ב-1967 הגשנו בקשה לסיוע מהאקדמיה הלאומית למדעים, באמצעות מר שלהבת פרייאר המנוח, אז סגן המנכ"ל במכון ויצמן. אך האקדמיה שתמכה בתכנית העבודה של המלחמ"ע ראתה בעצם בקשתנו ערעור על עבודת המלחמ"ע. זומנתי לפגישה באקדמיה, שם קבלתי נזיפה אישית חמורה מראש האקדמיה. (יותר מאוחר קיבלנו מענק מחקר צנוע מהאקדמיה, וכן מ-IBM (ישראל), הסוכנות, בל"ל ועוד.) יחד עם זאת, המלחמ"ע החל להביע את ספקותיו ביחס לגישתנו ביתר תוקף. נאמר לי שהם נועצו במדען מחשב או מומחה אנונימי, והוא פסק שגישתנו פסולה. בקשתי לראות או לקבל רמז כלשהו על התוכן המדעי

של הדו"ח שלו, אך לא נעניתי. מאידך, אני החילותי להבין יותר את האבסורד בגישה שלהם. הצבעתי בפ-ניהם על כך שאם ממירים את הרמות במפתח ההירכי שהם יצרו בכל הצורות האפשריות, מגיעים לגידול אכספוננציאלי, כך שנפח המפתח יגדל מנפח הספרים הממופתחים; ואם אין ממירים בכל הדרכים, הרי המחפש עלול לא למצוא את מבוקשו. על כגון אלה נאמר לי שהם מכינים את המפתחות בצורה "שיטתית ומדעית", אשר על כן הצלחתם מובטחת.

ב-1971 המלחמ"ע ניתק את הקשר אתנו ואף איים לנקוט באמצעים נגדנו. זה גרם לי ליאוש שחור ולמחשבות שאינני מעיז אפילו להעלות על הכתב.

הבעיות שלנו לא הצטמצמו באוניברסיטה העברית; גם עם בר-אילן היו לנו בעיות. הם לא אהבו את הקשר שלנו עם מכון ויצמן. בר-אילן היתה מודאגת מאד בקשר לבעלות על הסרטים המגנטיים של בסיס הנתונים שבנינו, וחיפשה דרכים להוציא את הפרוייקט מהמכון, בעוד שהמכון מעולם לא העלה שאלות מעין אלה, אעפ"י שהמכון תרם יותר כספים בתקופה ההתחלתית, הקריטית. אי-המיומנות שלי בפוליטיקה הכניסה אותי לבעיות חריפות, מה שגרם לי לעתים ליאוש מר. נשענתי יותר ויותר על שוויקה, ממנו קבלתי סיוע רב לא רק בחלק הבלשני שהוא ניהל, אלא גם בענינים פוליטיים, אם כי היה נראה לי שכישוריו הפוליטיים היו טובים משלי באופן שולי בלבד.

לעתים גם שוויקה איבד תקוה ורצה לזנוח את הכל. לדעת שנינו היה זה סיעתא דשמיא שמעולם לא התיאשנו יחדיו, כך שעלה בידינו לעודד איש את רעהו בתקופות של רוב רעה.

התפתחויות נוספות בסגנון טלגרפי

באוקטובר 1968 כתב אלי פרופ' אהרן שרייבר, אז בפקולטה למשפטים של אוניברסיטת מרילנד, על העניין הרב שהוא גילה בפרוייקט השו"ת, עליו הוא קרא באחד מפרסומי מכון ויצמן. במאמרו [Sch1969] הוא תיאר את העבודה שלנו. שרייבר גם השתתף בסדנא-סימפוזיון שאירגנתי במכון ויצמן בקיץ 1969. יותר מאוחר פרופ' שרייבר סייע לנו רבות בקבלת מענק מחקר גדול מ-NEH—National Endowment in the Humanities, בו הוא שימש כחוקר ראשי.

המחקר בפרוייקט שיגשג; תמיד סברנו שהפרוייקט יכול לשרוד רק באוירה של מחקר – עיוני ומעשי. היו לנו פגישות שבועיות בהן דנו בשאלות מדעיות. היו לי מספר סטודנטים לתאר שלישי, משה מור, שמואל קליין (שניהם עתה חברי המחלקה למתימטיקה ומדעי המחשב בבר-אילן), רוני עטר (שייסד את פרוייקט פרח הידוע בהיותו סטודנט) והרבה סטודנטים לתאר שני שעבדו על בעיות שהתעוררו בקשר לפרוייקט. כולם קיבלו את התארים שלהם ממכון ויצמן. בשלב יותר מאוחר היו גם לפרופ' שוויקה סטודנטים לתאר שני בבר-אילן שעבדו על בעיות עם זיקה לפרוייקט. סטודנטים אחרים שעבדו בפרוייקט כללו את נחום דרשוביץ ודוד הראל, שניהם עתה פרופסורים למדעי המחשב באוניברסיטה של אילינוי באורבנה-שמפיין ובמכון ויצמן בהתאמה.

לאחר שהקשר שלי עם בר-אילן הלך ופחת, פוטרתי כמנהל הפרוייקט בידי דיקן בבר-אילן, ושוויקה מונה במקומי. מהלך זה לא גרם לשום משבר, כי המעבר נעשה בצורה מסודרת וחלקה, והמשכתי בעבודה מדעית בפרוייקט. אותו דיקן, לאחר שהוא נהיה נשיא, גרם להתפטרותו של שוויקה בשיאה של מחלוקת רבתי אודות הפרוייקט, והוא גם הפסיק כל עבודת מחקר בפרוייקט, תוך הפיכתו ליחידת שרות. לפני

כן הוא ביקר במשרד הראשי של NEH בוושינגטון והביא להפסקת מענק המחקר שלנו איתם. הפרוייקט מעולם לא התאושש ממכות כבדות אלה.

ההישג העיקרי בעת כהונתו של שוויקה היתה ההחלטה ל"הפריט" את המערכת שעבדה עד אז על המחשב המרכזי של בר-אילן בלבד, באמצעות מסופים. (גם המלחמ"ע רכש מסוף, שנים רבות לאחר ניתוק הקשרים בינינו!) הפירוש המעשי של ההפרטה היה לשנות את צורתה של המערכת כך שהיא תוכל לעבוד על מחשב ביתי באמצעות תקליטור (CD-ROM). פרוייקט התקליטור מומש רק כעבור מספר שנים, בעת כהונתו של פרופ' צבי ארד כרקטור, והודות למאמציו הברוכים; ראיתי שבמרץ 1989 כתבתי לרקטור פרופ' קראוס של בר-אילן, לבקשתו, על הכוונים הרצויים של הפרוייקט, ובמכתב זה כלולה גם המלצה לממש את פרוייקט התקליטור. לאחר המימוש, באפריל 1992, השימוש במערכת גדל במאד. מאז 1990 דר. אורי שילד מנהל את הפרוייקט בכישרון רב; בתשנ"ו, כשאורי היה בשבתון, פרופ' עמיהוד אמיר נשא את המשא בכישרון רב. הרב יעקב ויינברגר משמש שנים רבות כסגן מנהל נאמן ומוכשר, שידע להתמודד עם הבעיות גם בזמנים קשים מאד. בין שאר הדברים, הוא עובר אישית על תלונות המשתמשים (ונהנה מהשבחים), מה שמסייע ליצור מהדורות משופרות של התקליטור. אנשים רבים נוספים סייעו להגיע לאותם הישגים שהושגו עד עתה, ואני מתנצל מאד על השמטות שבטוחני שחטאתי בהן.

תקציר פעילות בעת האחרונה

בנסיבות הנוכחיות מבוצע מעט מאד מחקר. "אינציקלופדיה תנכית" מאוחסנת על דיסקט הוכנה ע"י הפרוייקט עבור הוצאת הספרים סטימצקי. היא מכילה מפתחות שונים ומאפשרת עריכת חיפושים על עותק מנוקד של התנ"ך של מערכת השו"ת. שיווק, באמצעות סטימצקי ופרוייקט השו"ת, החל בסתיו 1994. פעילות משותפת נוספת עם סטימצקי היה מיכון אוסף פירושי המקרא "דעת מקרא", שיצא בהוצאת מוסד הרב קוק, והזכויות על הפירושים נרכשו בידי סטימצקי.

בשנה האחרונה בוצע מיכון, כולל טקסט-על (Hypertext) של הקונקורדציה התלמודית, תחת חוזה עם מפעל הקונקורדציה התלמודית. גם המהדורה הרביעית של התקליטור שוחררה בשנה האחרונה. המהדורה כוללת שכלולי תכנה, כגון אופציה ליצירת כל הכתיבים של מלה, מכתיב מלא ועד כתיב חסר, ביבליוגרפיות של מחברי השו"ת באנגלית (בנוסף לעברית), הגדלה בבסיסי הנתונים, הרצת חיפוש על לא יותר מ-10 בסיסי נתונים בו-זמנית. בסיסי הנתונים בתקליטור הנוכחי הם:

- תנ"ך
- פרשנות תנ"ך: תרגום אונקלוס, רש"י, רמב"ן על החומש; רש"י, רד"ק, רלב"ג, מצודת דוד, מצודת ציון על הנ"ך
- משנה
- תוספתא: מהדורת ליברמן עד בבא בתרא כולל; צוקרמנדל לשאר המסכתות
- מסכתות קטנות: גרים, דרך ארץ, כותים, כלה, כלה רבתי, מזוזה, סופרים, ספר תורה, עבדים, ציצית, שמחות, תפילין

- תלמוד בבלי
- רש"י על תלמוד בבלי
- תלמוד ירושלמי: ונציה, הרפ"ג; חלוקה להלכות עפ"י דפוס ווילנה
- מדרשי הלכה: מכילתא דרבי ישמעאל, מכילתא דרבי שמעון בר יוחאי, ספרא, ספרי (במדבר, דברים), ספרי זוטא, מדרש תנאים לספר דברים
- מדרשי אגדה: הרשימה לא מובאת כאן מפאת ארכה
- רמב"ם: היד החזקה (משנה תורה)
- טור
- בית יוסף
- שולחן ערוך
- קיצור שולחן ערוך
- משנה ברורה
- שו"ת ראשונים (32 מחברים וכן גאונים ובעלי תוספות, 47 כרכים)
- שו"ת המאה ה-16 (28 מחברים, 48 כרכים)
- שו"ת מאות 17 -- 19 (48 מחברים, 87 כרכים)
- שו"ת פוסקי זמננו (15 מחברים, 82 כרכים)

מערכת מוקטנת "אוצרות ישראל" (גם מכונה "מתנת בר מצוה" באה"ב), שוחררה בעת האחרונה. היא מכילה: תנ"ך על פירושיו הנ"ל (למעט דעת מקרא), מדרש רבה, מדרש תנחומא, תלמוד בבלי עם רש"י, רמב"ם, קיצור שולחן ערוך. פרטים נוספים ניתן לקבל מפרוייקט השו"ת, אוניברסיטת בר אילן, רמת-גן 52900, טלפונים: 03-531-8411, 03-531-8625, 03-531-8343; פאקס: 03-534-1850 (עבור "אוצרות ישראל": טל (טכנולוגית למידה), משה חסיד, טל. 03-932-0499). באה"ב: T E S (Torah Educational Software, Inc.), 21 Main Street, Monsey, NY 10952, Tel. 1-800-925-6853, 1-914-356-1485; Fax: 1-914-356-1343. ממדינות אחרות יש להתקשר לאחת משתי הכתובות הנ"ל.

האירוניה הטראגית של המצב הנוכחי

שמנה וסלתה של הקהילייה האקדימית בישראל החליטה על הקמתו של המלחמ"ע. הוא הוקם ב-1963 ותוקצב כיאות. יעודו כממפתח של ספרות השו"ת פורסם ברוב עם. במחצית אוקטובר 1996 שמעתי מפרופ' אליאב שוחטמן, שהיה ראש המלחמ"ע עד לפני שנה, ופרופ' ברכיהו ליפשיץ מהפקולטה למשפטים באוניברסיטה העברית, שלאחר פרסום המפתחות על ראשוני ספרד, המלחמ"ע עוד הכין מפתחות לראשוני אשכנז, עבודה שנסתיימה לפני שנים רבות ועומדת להתפרסם בקרוב. אך לאחר מכן פעילות מפתוח נמשכה רק בצורה מוגבלת מאד. כיום המלחמ"ע מפרסם "שנתון המשפט העברי", אך כנראה אין בו פעילות רבה נוספת.

הפרוייקט נולד בהחבא. הוא לא זכה מעולם בהכרה או תמיכה אקדמית רחבה. היחס אליו היה ונשאר כיחס לבן חורג. בר-אילן דבקה בו, אך לא מאפשרת לו לחיות. היו זמנים בהם הפרוייקט זכה לתמיכה כספית חלקית מאוניברסיטת בר-אילן, ועל כך נתונה לה תודתנו, אך בדרך כלל היה עלינו להשקיע מאמצים אינסופיים ויותר, כדי לזכות בה. על אוניברסיטת בר-אילן גם להיות אסירת תודה לנו על שהבאנו לה את הפרוייקט! הפרוייקט הוא חלון הראווה של האוניברסיטה, היהלום בכתרה. אך היא מענה אותו. הפרוייקט נתון כל העת למצבי הרוח של רקטורים ונשיאים מתחלפים, שכל אחד מהם רואה חובה לעצמו לדוש בעניני הפרוייקט. האוניברסיטה דנה מספר פעמים על הוצאת הפרוייקט מתוכה. חילופין באדמיניסטרציה, או החלטה לדון מחדש בפרוייקט, גרמו לנו בדרך כלל לחוסר שינה והמתנה מתוחה ממושכת להחלטות. שוררת אוירה קבועה של משבר, עם תקופות זמניות של מרפא חלקי, שמזיקה מאד לפעילות. (כרגע נראה שההכנסות שהפרוייקט מכניס יזוקפו לזכותו, אך ההקפאה נמשכת כי טרם התקבלה החלטה סופית. על כל פנים, שום תמיכה כספית של בר-אילן לא ניתנת ולא מתוכננת לפרוייקט היחיד שלה המשלב בתוכו בצורה טבעית תורה עם דרך ארץ, שאמורה להיות הסיבה העיקרית לקיומה של האוניברסיטה.) ייתכן ואני מדמה זאת לעצמי, אך אולי ניכר שיפור מזערי בגישה של בר-אילן, בעטיים של חברי פקולטה בגיל הביניים ובגיל צעיר, שמסוגלים יותר להעריך את הכח הטמון בפרוייקט, ושמגיעים עתה לעמדות בכירות יותר.

פעילות המפתות של המלחמ"ע דעכה בקול דממה דקה. מומחים אומרים לי שהמפתחות שהם הכינו אינם בשימוש רחב, אך ייתכן שזה ענין של שיפוט סובייקטיבי. בתוצרת של הפרוייקט משתמשים בכל העולם. ניתן למצוא את המערכת, לדוגמה, ברחבי אה"ב ואירופה, כולל מזרח אירופה, בקהילות נידחות כגון אוסטרליה המערבית, וכמובן בארץ: במחנה החרדי, בציבור הדתי-לאומי, במספר קבוצים של השומר הצעיר, בבתי משפט ובתי דין, בין סטודנטים ואנשי רוח. סיני דויטש, סטודנט למשפטים בבר-אילן (לאחר מכן: דיקן הפקולטה למשפטים שם) השתמש במערכת עבור התיזה שלו בשלב מוקדם של הפרוייקט. בתחילת 1978 הוא יסד את תת-הפרוייקט "שמע" (שרות משפטי עברי) לשימושם של שופטים ועו"ד, שמסובסד חלקית ע"י לשכת עוה"ד בישראל ומשרדי המשפטים והדתות. פרופ' דויטש והרב מאיר בטיסט מריצים חיפושים במערכת השו"ת, במסגרת שמע, לסיוע במשפטים שוטפים הנידונים בבתי המשפט.

האם נובע מכאן שהפרוייקט השיג את מטרתו לפתוח את ספרות השו"ת לעם היהודי? התשובה, למרבה הצער, היא לאו מוחלט. המשתמשים המדמים שהם שבעי רצון, אינם יודעים מה הם מחמיצים!

בראשית הרשימה הזו השתבחתי בכך שמתחילה ראיתי בעיני רוחי מערכת שתיענה לכל דורשיה ביע-ילות לאחר הקשה על כמה כפתורים. אך מה שלא חזיתי מראש כלל ועיקר היה שמערכות תכנה הן דינמיות

במהותן. מעולם אי אפשר ל"סיימם". "יום תעזבני, יומיים אעזבך" (ירושלמי ברכות, פ' ט ה' ה, בשלהי הפרק). יתר על כן, החלטה שגויה כיום עלולה לגרום להרס המערכות מחר. אין הרבה גופים, אם בכלל, עם גב פיננסי כמו IBM, שמסוגלים לשרוד לאחר החלטה שגויה. גם לו נסיבות פרוייקט השו"ת היו נאותות, היינו צריכים לעמול יומם ולילה למימוש החלום — רק כדי לשמור על מחצית מימושו!

לדוגמה, היינו צריכים לערוך, בדחיפות, את המחקרים והפיתוחים הבאים: איחזור מידע בסביבה של הערות שוליים וטקסט-על [FrK≥1997], שיפורים בניסוח החיפוש (שעדיין פרימיטיבי מאד), קוראים אופטימיים למטרות מיוחדות (כתב רש"י. . .התגבר על הבעייה העיקשת של הקליטה), איחזור עם ניקוד (כמה מהטקסטים, כגון התנ"ך, מנוקדים), השואת כתבי יד (דקדוקי סופרים. . .תיאום טקסטים) הצגה של דף סטנדרטי של התלמוד הבבלי על מסך מחשב, הרחבת הכלים הדקדוקיים, שיטות דחיסה משופרות (אנו בגבול הקיבול של התקליטור), חיפושים בתוך טקסט דחוס, הבטחה והצפנה (המערכת שלנו נפרצה ונמכרת ע"י הפורצים), ופעילויות רבות נוספות. חלק גדול מנושאים אלה מוסברים ביתר פירוט ב-[Ami1996].

כמובן שיש צורך להגדיל את בסיס הנתונים, במיוחד בשו"ת (איחסנו רק כ-10% של השו"ת הידועים — אף כי בעלי המשקל היותר גדול) ובפוסקים. ברור שלמערכת השו"ת הפוטנציאל ליהפך למערכת המידע של כלל אוצרות הרוח העשירים של היהודים. ראשונים ואחרונים, מפרשי התלמוד(ים), נושאי הכלים של הרמב"ם, טור ושולחן ערוך הם רק מספר דוגמאות.

סיכום

פרוייקט השו"ת שייך לעם היהודי. הוא צריך לקבל מעמד של פרוייקט לאומי, בחסות גופים כגון האקדמיה הלאומית למדעים, הכנסת והממשלה, עם תקציב קבוע, בניהולו של אדם צעיר, מוכשר ונמרץ, ובפיקוחה של מועצה מדעית מתאימה. זוהי בקשה טבעית לאחר שנות פעילות רבות בהן הפרוייקט הוכיח את עצמו על אף הקשיים העצומים. העם היהודי יקצור את הפירות ממהלך כזה.

נראה שבמצב אופטימלי הפרוייקט ישכון באחד המוסדות להשכלה גבוהה בארץ, אולי בדומה לאקדמיה ללשון, השוכנת בקמפוס של האוניברסיטה העברית. מוסד זה יוכל אולי להיות בר-אילן, אך רק אם תובטח לחלוטין אי-התלות בהחלטות אדמיניסטרטיביות של בר-אילן. המחיר הנוכחי המוגזם של התקליטור, שנקבע עקב הנסיבות הקשורות לעצם קיומו של הפרוייקט במצב שנוצר, צריך לרדת משמעותית כשהפרוייקט ייהפך למפעל לאומי. (נזכיר שיוזמות מסחריות שקמו לאחרונה לשיווק בסיסי נתונים יהודיים מטרתן רווחים, והרבה מהטקסטים שלהם, למעט אלה שנגנבו מאתנו, לוקים בחסר במובנים שונים.)

נסיים בהבאת חלק מהפרסומים שלנו. כמובן שיש הרבה פרסומים שהמערכת סייעה ביצירתם, לעתים בליווי חסכון של שנות מחקר. יש מהמחברים שהביאו, ויש שלא הביאו גאולה לעולם; כאן לא נזכיר לא את הראשונים ולא את האחרונים. גם חלק מחברי הצוות שלנו ניצלו את המערכת למחקרים ביהדות, מדעי הרוח והחברה. היינו צורכים מספר כפול של דפים לו היינו מביאים כאן רשימה של פירות יצירותיהם. ניתן לעיין ברשימה חלקית [Wei1981]. כאן נזכיר, כדוגמה בלבד, את ספרו של מנחם סליי "הביטוח בהלכה", בחסות איגוד חברות הביטוח בישראל (עברית 1980, אנגלית 1982), שגם ערך את "ספר השו"ת: שמיטה", בית ישרים, הרי ירושלים, 1986, המבוסס על תשובות בעניני שמיטה שנמצאו בעזרת המערכת. אגב, בהרבה ישיבות, תיכונים, ואף בכמה חטיבות ביניים המערכת בשימוש, וצוותים של תלמידים מכינים חוברות יפות על בעיות הלכה ומוסר שלובנו בסיוע מערכת השו"ת.

ביבליוגרפיה

[Ami1996] A. Amir, Bar-Ilan Responsa Project — Research Proposals, Bar-Ilan University Internal Report, 1996.

[ApF1987] A. Apostolico and A.S. Fraenkel, Robust transmission of unbounded strings using Fibonacci representations, *IEEE Trans. Information Theory* **33** (1987) 238–245.

[ACDF1978] R. Attar, Y. Choueka, N. Dershowitz and A.S. Fraenkel, KEDMA — linguistic tools for retrieval systems, *J. Assoc. Comp. Mach.* **25** (1978) 52–66.

[AtF1977] R. Attar and A.S. Fraenkel, Local feedback in full-text retrieval systems, *J. Assoc. Comp. Mach.* **24** (1977) 397–417.

[AtF1981] R. Attar and A.S. Fraenkel, Experiments in local metrical feedback in full-text retrieval systems, *Inform. Process. & Management* **17** (1981) 115–126.

[AFCS1972] R. Attar, A.S. Fraenkel, Y. Choueka and D. Schindler, Linguistic files in information retrieval systems (in Hebrew), Proc. National Conference Inform. Process. Assoc. Israel (Tel-Aviv, November 1972), pp. 218–247.

[Cho1983] Y. Choueka, Linguistic and word-manipulation components in textual information systems, in: The Applications of Mini- and Micro-computers in Information Documentation and Libraries (C. Keren, L. Perlmutter, eds.), Elsevier, North-Holland, pp. 405–417, 1983.

[Cho1990] Y. Choueka, RESPONSA: a full-text retrieval system with linguistic components for large corpora, in: (a) *Computational Lexicology and Lexicography*, volume in honor of B. Quemada (A. Zampolli, ed.), Giardini Ed., Pisa, 1990. Also, with slight variations, in: (b) *Law and Multicultural Societies*, Proc. Intern. Assoc. Law Libraries (IALL) Meeting (E.I. Cuomo, ed.), Jerusalem 1985, The Hebrew University, Jerusalem, pp. 47–82, 1989. Also in: (c) Literary and Linguistic Computing, Proc. 15th Intern. Conf. of Assoc. Literary and Linguistic Computing (ALLC) (Y. Choueka, ed.), Jerusalem 1989, Champion-Slatkine, Paris-Gènève 1990, pp. 51–92. An abridged version entitled "Responsa, a full-text retrieval system with linguistic processing for a 65 million-word corpus of Jewish heritage in Hebrew", in: (d) *Data Engineering* **12**, IEEE Computer Soc., special issue on non-English interfaces to databases, pp. 22–31, 1989.

[ChB1982] Y. Choueka and P. Bratley, Processing truncated terms in document retrieval systems, *Inform. Process. & Management* **18** (1982) 257–266. Also in: Proc. 16th Nation. Conf. Inform. Process. Assoc. Israel (IPA), Jerusalem, Israel.

[CCDFS1971] Y. Choueka, M. Cohen, J. Dueck, A.S. Fraenkel and M. Slae, Full text document retrieval: Hebrew legal texts (report on the first phase of the Responsa Retrieval Project), Proc. Symp. on Information Storage and Retrieval (J. Minker and S. Rosenfeld, eds., University of Maryland, April 1971), pp. 61–79.

[CCDFS1972] Y. Choueka, M. Cohen, J. Dueck, A.S. Fraenkel and M. Slae, Full text case law retrieval: the Responsa Project (a solicited expanded form of paper [CCDFS1971]), Working Papers on Legal Information Processing Series, Pamphlet No. **3** (64 pages), *J. Schweitzer Verlag,* 1972.

[CFK1988] Y. Choueka, A.S. Fraenkel and S.T. Klein, Compression of concordances in full-text retrieval systems, Proc. ACM Conference on Research and Development in Information Retrieval (Y. Chiaramella, ed., Grenoble, France, June 1988), pp. 597–612, Presses Universitaires de Grenoble, 1988.

[CFKS1986] Y. Choueka, A.S. Fraenkel, S.T. Klein and E. Segal, Improved hierarchical bit-vector compression in document retrieval systems, Proc. ACM Conference on Research and Development in Information Retrieval (F. Rabitti, ed., Pisa, Italy, September 1986), pp. 88–96, Assoc. Comp. Mach., New York, NY, 1986.

[CFKS1987] Y. Choueka, A.S. Fraenkel, S.T. Klein and E. Segal, Improved techniques for processing queries in full-text systems, Proc. ACM Conference on Research and Development in Information Retrieval (C.T. Yu and C.J. Van Rijsbergen, eds., New Orleans, LA, June 1987), pp. 306–315, Assoc. Comp. Mach., New York, NY, 1987.

[CFP1981] Y. Choueka, A.S. Fraenkel and Y. Perl, Polynomial construction of optimal prefix tables for text compression, Proc. 19th Annual Allerton Conference on Communication, Control and Computing (Monticello, IL, Oct. 1981), pp. 762–768.

[ChG1979] Y. Choueka and D. Goldberg, Mechanical resolution of lexical ambiguity—a combinatorial approach, Proc. Intern. Conf. on Literary and Linguistic Computing (Zvi Malachi, ed.), Israel, April 1979, The Katz Res. Inst. for Hebrew Literature, Tel Aviv University, pp. 149–165, 1979.

[CKN1983] Y. Choueka, S.T. Klein and E. Neuwitz, Automatic retrieval of frequent idiomatic and collocational expressions in a large corpus, *ALLC J.* **4** (1983) 34–38.

[CKP1985] Y. Choueka, S.T. Klein and Y. Perl, Efficient variants of Huffman codes in high-level languages, Proc. 8th Annual Intern. ACM-SIGIR Conf. Res. & Develop. in Information Retrieval (Montreal, 1985), Assoc. Comp. Mach. New York, NY, pp. 122–130.

[ChL1985] Y. Choueka and S. Lusignan, Disambiguation by short contexts, *Computers and the Humanities* **19** (1985) 147–157.

[ChS1964] Y. Choueka and M. Shapiro, Machine analysis of Hebrew morphology: potentials and achievements (Hebrew), *L'shoneinu* **27–28** (1964) 354–372. English abstract in: *The Finite String* **3**, 1966.

[CSS1980] Y. Choueka, M. Slae and A. Schreiber, The Responsa Project: computerization of traditional Jewish law, in: *Legal and Legislative Information Processing* (B. Erez ed.), Greenwood Press, 1980, Ch. 18, pp. 261–286.

[CSS1979] Y. Choueka, M. Slae and S. Spero, The Responsa Project—revisited, *Proc. AOJS* **5** (1979) 19–66.

[Fra1968] A.S. Fraenkel, Legal information retrieval, in: *Advances in Computers* (F.L. Alt and M. Rubinoff, eds.), Vol. **9**, pp. 113–178, Academic Press, New York, NY, 1968.

[Fra1968a] A.S. Fraenkel, Modern retrieval methods applied to Jewish case law, Proc. 1st Intern. Symposium Soc. Tech. Writers and Publishers (Tel-Aviv, December 9–12, 1968), pp. R1-1–R1-9, Soc. Tech. Writers and Publishers Israel Chapter, Lod Airport, Israel.

[Fra1968b] A.S. Fraenkel, Seventeen centuries on tape, *Law and Computer Technology* **1** (1968) 8–9.

[Fra1969] A.S. Fraenkel, A retrieval system for the Responsa, *Proc. AOJS* **2** (1969) 3–42.

[Fra1969] A.S. Fraenkel, Full text document retrieval, Proc. Symp. on Information Systems in Chemistry (Tel-Aviv, May 14–15, 1969), pp. 86–90, National Council for Research and Development, Center for Scientific and Technological Information, Tel-Aviv, Israel.

[Fra1973] A.S. Fraenkel, Mechanized information retrieval as an auxiliary tool in Jewish law research, Proc. 5th World Congress of Jewish Studies, World Union of Jewish Studies (A. Shinan, ed., Jerusalem, 1969), Vol. **5**, pp. 79–98 (in Hebrew; English abstract: pp. 68–69), Jerusalem Academic Press, Jerusalem, 1973.

[Fra1976] A.S. Fraenkel, All about the Responsa Retrieval Project you always wanted to know but were afraid to ask, Expanded Summary, Proc. 3rd Symposium Legal Data Processing in Europe, Oslo, July 1975, pp. 131–141 (in French edition of Proceedings, pp. 111–121), Council of Europe, Strasbourg, 1976 (invited). Also appeared in *Jurimetrics Journal* **16** (1976) 149–156, and in *Informatica e Diritto* II No. **3**, pp. 362–370, September 1976.

[Fra1981] A.S. Fraenkel, Document classification, indexing and abstracting may be inherently difficult problems, Proc. Theoretical Issues in Information Retrieval, 4th Intern. Conference on Information Storage and Retrieval (C.J. Crouch, ed., Oakland, CA, June 1981), pp. 77–82, Assoc. Comp. Mach., New York, NY, 1981.

[Fra1985] A.S. Fraenkel, Information storage and retrieval—complexities and approximate solutions, Expanded summary, in: The Role of Data in Scientific Progress, Proc. 9th Intern. CODATA Conference (P.S. Glaeser, ed., Jerusalem, June 1984), pp. 391–394, North Holland, Amsterdam, 1985.

[FrK1985] A.S. Fraenkel and S.T. Klein, Novel compression of sparse bit-strings—preliminary report, Proc. NATO Advanced Research Workshop on Combinatorial Algorithms on Words (A. Apostolico and Z. Galil, eds., Maratea, Italy, June 1984), pp. 169–183, Springer-Verlag, Heidelberg, 1985.

[FrK1990] A.S. Fraenkel and S.T. Klein, Bidirectional Huffman coding, *The Computer J.* **33** (1990) 296–307.

[FrK1993] A.S. Fraenkel and S.T. Klein, Bounding the depth of search trees, *The Computer J.* **36** (1993) 668–678.

[FrK1994] A.S. Fraenkel and S.T. Klein, Complexity aspects of guessing prefix codes, *Algorithmica* **12** (1994) 409–419 (special issue on String Algorithmics).

[FrK1996] A.S. Fraenkel and S.T. Klein, Robust universal complete codes for transmission and compression, *Discrete Applied Mathematics* **64** (1996) 31–55.

[FrK≥1997] A.S. Fraenkel and S.T. Klein, Information retrieval from annotated texts with applications to hypertexts, submitted for publication.

[FrM1983] A.S. Fraenkel and M. Mor, Combinatorial compression and partitioning of large dictionaries, *The Computer J.* **26** (1983) 336–343. A version entitled Combinatorial compression and partitioning of large dictionaries: theory and experiments, in which the experimental part has been expanded on, appeared in Proc. 6th Annual Intern. ACM SIGIR Conference on Research and Development in Information Retrieval (M.J. McGill, ed., National Library of Medicine, Bethesda, MD, June 1983), pp. 205–219, Assoc. Comp. Mach., New York, NY, 1983.

[FMP1983] A.S. Fraenkel, M. Mor and Y. Perl, Is text compression by prefixes and suffixes practical?, *Acta Informatica* **20** (1983) 371–389. An earlier version appeared in Proc. Research and Development in Information Retrieval, 5th Intern. Conference on Information Storage and Retrieval (G. Salton and H.-J. Schneider, eds., Berlin, May 1982), Lecture Notes in Computer Science, Vol. **146**, pp. 289–311, Springer-Verlag, 1983.

[FRS1979] A.S. Fraenkel, D. Raab and E. Spitz, Semi-automatic construction of semantic concordances, *Computers and the Humanities* **13** (1979) 283–288. (Extended Abstract in: Databases in the Humanities and the Social Sciences (J. Raben and G. Marks, eds.), pp. 101–103, North-Holland, Amsterdam, 1980.)

[Hor1962] J.F. Horty, Searching statutory law by computer, Interim Report No. 1, Health Law Center, Univ. of Pittsburgh, 1962.

[MoF1982] M. Mor and A.S. Fraenkel, A hash code method for detecting and correcting spelling errors, *Comm. Assoc. Comp. Mach.* **25** (1982) 935–938. An expanded version entitled "Retrieval in an environment of faulty texts or faulty queries", appeared in Proc. 2nd Intern. Conference on Improving Database Usability and Responsiveness (P. Scheuerman, ed., Jerusalem, June 1982), pp. 405–425, Academic Press, New York, NY, 1982.

[Sch1969] A.M. Schreiber, Computerized Storage and retrieval of case law without indexing: the Hebrew Responsa Project, *Law and Computer Technology* **2** (1969) 14–22.

[Wei1981] Y. Weinberger, A list of Hebrew publications (Hebrew), Internal report, The Responsa Project, Bar Ilan University, 1981.