

ירושלים כ"ט סיוון תשס"ט
21 ביוני 2009

מערכת חדשה לעבודה עם קובצי MUC מרחוק

בלשכה המרכזית לסטטיסטיקה פותחה מערכת טכנולוגית חדשה שתאפשר לחוקרים לעבוד על קבצים מסוג MUC בגישה מרחוק. המערכת פותחה בהתאם להמלצות ועדת מומחים בראשותו של פרופ' צבי אקשטיין בעניין סביבת עבודה בטוחה עם קובצי פרט. המלצות הועדה אושרו על ידי המועצה הציבורית לסטטיסטיקה ואומצו על ידי הסטטיסטיקן הממשלתי, פרופ' שלמה יצחקי.

הלשכה המרכזית לסטטיסטיקה (הלמ"ס) מעמידה לרשות חוקרים נתוני פרט (microdata) בלתי-מזהים (anonymized) לשם מחקר אמפירי בשלוש צורות השונות זו מזו ברמת פירוט הנתונים ובצורת העבודה עמם.¹

- קובץ PUF (Public Use File), שמכיל פירוט מינימלי, מתאים יותר לחוקרים מתחילים (סטודנטים) ולניתוח בסיסי. קובץ PUF מופץ על גבי תקליטור והוא נגיש לכל הן מבחינת הרשאות השימוש והן מבחינת המחיר (100 ש"ח).
- קובץ MUC (Microdata Under Contract) מספק פירוט רב יותר של הנתונים ונועד למחקר מתקדם. שימוש בקובץ MUC מוגבל לחוקרים מורשים בלבד, תחת הסכם אישי המגביל את החוקר באופן השימוש בנתונים ומבטיח שמירת סודיות סטטיסטית.
- נתונים בלתי-מזהים המפורטים ביותר, זמינים למחקר שאושר על-ידי המדען הראשי של הלמ"ס במסגרת חדר המחקר של הלמ"ס. עבודה עם נתונים אלה מתבצעת בבית הלמ"ס, תחת המעטפת המשפטית של פקודת הסטטיסטיקה ובהסדרי אבטחת מידע מחמירים, כמקובל בלשכות סטטיסטיות בעולם עבור נתוני פרט מפורטים.

המערכת החדשה באה להחליף את שיטת העבודה הקיימת, שבה קובצי MUC מופצים על גבי מדיה נתיקה לחוקרים המורשים, לשם הרחבת חוג המשתמשים והגברת אבטחת המידע.

כיצד תפעל המערכת החדשה?

כל חוקר המשתייך לצוות מחקר מורשה יקבל גישה לקובץ MUC, שיתוחזק ויופעל בסביבה בטוחה באתר הלמ"ס, בשטח נפרד לכל מחקר. הגישה תהיה באמצעות הזדהות אישית מאובטחת, דרך האינטרנט. החוקר יוכל לבצע עיבודים סטטיסטיים על קובץ MUC כרצונו בכל עת. את תוצאות העיבודים (פלט) ואת דוחות המחקר אפשר יהיה להוציא מהמערכת באופן מבוקר, בהתאם לכללי אבטחת מידע שיתפרסמו בנפרד.

אילו תוכנות יימצאו במערכת?

המערכת תעמיד לרשות החוקרים את התוכנות הסטטיסטיות הנפוצות: SAS, Stata, SPSS. לשם נוחיות העבודה, אפשר יהיה לייבא אל תוך המערכת תוכניות עיבוד בתוכנות אלה ולשמור אותן להמשך העיבודים. כמו כן, בכל סביבת העבודה יהיו תוכנות של MS Office: Word, Excel.

¹ פרטים בקישור: http://www.cbs.gov.il/mucpuf/muc_puf_h.htm

מי יורשה לעבוד במערכת החדשה?

על פי הנוהל הקיים, מוגדר כחוקר מורשה מי שנמצא במסלול מחקרי באוניברסיטה בישראל במעמד של מרצה ומעלה, או בתפקיד אקדמי במסלול האקדמי של מכללה אקדמית שאין בה קביעות, או בעל מינוי או ראש מוסד למחקר שהוכר על-ידי הסטיסטיקן הממשלתי.² בנוסף לאלה, תורשה עבודה במערכת החדשה לתלמידי מחקר (דוקטורנטים) במוסדות להשכלה גבוהה בישראל וכן לעוזרי מחקר כחברי צוות המחקר. החוקר יידרש לחתום על הסכם לשמירת סודיות והתנהגות בהתאם לכללים הנדרשים, כפי שקיים הסכם חוקר MUC היום.

אילו נתונים יהיו זמינים בקובצי MUC?

הלמ"ס תעמיד לרשות החוקרים מבחר נתונים, ברמת פרט (רשומות בלתי-מזוהות) מסקרי משפחות/נפשות שוטפים, ממפקדי האוכלוסין וממקורות מנהליים הנמצאים בידי הלמ"ס. נתוני עסקים אי אפשר לחקור באמצעות קובצי MUC אלא רק באמצעות חדר המחקר של הלמ"ס. **קובצי MUC סטנדרטיים** – מוכנים מראש זמינים מיידית – יכללו את נתוני המפקדים והסקרים הבאים: הסקר חברתי, סקר הכנסות, סקר הוצאות משק-הבית, סקר כוח-אדם ונתוני תאונות דרכים עם נפגעים. קישור רשומות והכנת קבצים לפי דרישות החוקר ייעשו במסגרת **קובצי MUC מורכבים**

מהי עלות השימוש במערכת החדשה?

עלות השימוש תכלול דמי שימוש ועלות הכנת קובץ MUC. דמי השימוש (לצוות מחקר) הם 4,800 ₪ לשנה (או 400 ₪ לחודש). המחיר הוא למינוי שנתי מ-1 באוקטובר ועד 30 בספטמבר. יצוין, כי מדובר במערכת חדשה והשנה הראשונה להפעלתה (עד 30 בספטמבר 2010) תחשב כשנת הרצה, ולכן שלושת החודשים הראשונים של המינוי בתקופה זו בלבד יינתנו בחינם.

האם אפשר להמשיך לעבוד על קובצי MUC שבידי החוקרים?

כן, חוקר שקיבל קובץ MUC על גבי מדיה נתיקה כדין יכול להמשיך להשתמש בו עד סיום תקופת תוקף רשיון השימוש בקובץ. לאחר מכן, במקרה הצורך, אפשר להאריך את תקופת השימוש לפי הנוהל הקיים. יודגש, כי קובצי MUC חדשים – אלה שיופקו לראשונה אחרי מועד הפעלת המערכת החדשה – יהיו זמינים אך ורק לעבודה מרחוק.

מתי תופעל המערכת החדשה?

המערכת תופעל בהדרגה בחודשיים הקרובים. לרישום ולהזמנת קובצי MUC יש לפנות ליחידת המידע של הלמ"ס בטל" 2056 / 02-6592666 או בדוא"ל: info@cbs.gov.il

² http://www.cbs.gov.il/mucpuf/muc/muc_regulations/brian-23_7_2006h.pdf