

ארכיוני משמר העמק ושומריה 2019

נוהל סריקת מסמכי הארכיון וקיטלוגם

המטרה- לשמר ולהנגיש את מסמכי הארכיון.

שלב 1. פעולות הארכיון אית, כהכנה לסריקות

- א. ממיינת את המסמכים בכל מיכל ותיק:
- מאגדת בכל תיק, מסמכים בעלי נושא זהה, לסריקתם כמסמך אחד.
 - כל תיק יכול עד 30 מסמכים (מסמך יכול להכיל כמה דפים, שייסרקו לתוך קובץ משותף אחד, לדוגמה- חוברת שירון).
 - מחלקת תיקים 'שמנים', לתיקי מישנה, שכוללים עד 50 מסמכים בכל אחד, (כל תיק מישנה כזה יסומל באות, לדוגמה- תיק ילדים א, ילדים ב).
 - כתיבת שם הקובץ, שפה, והסיווג על כל תיק (פומבי/פנימי/חסוי).
 - כתיבת מילת הכללה באנגלית על כל תיק. (לדוגמה- Children, Kibbutz, Agriculture, Hashomer Ha Tzair movement, Pioneers, Culture, Festivals, Members, Work).
 - תוודא שהמיכלים והתיקים ממוספרים בסדר נכון.
 - תנפה כפילויות, ואי-רלוונטיות.
- ב. מעבירה את המיכל הממוין לסריקה.
- ג. יוצרת חד פעמית, לפני הפיילוט, פריט דוגמא בתכנה, מסוג החומר של הסריקה, ששמו 'מסמכי ארכיון סרוקים'. נעזרת לכך באיש התכנה או באיש המחשוב.

שלב 2. פעולות הצורך סורק:

- א. יוצר במחשב תיקייה ניפרדת לכל מיכל ארכיוני. ונותן שם לתיקייה זו, בצורה של קוד:
- המספר הראשון מייצג את החטיבה, ואחריו נקודה.
 - המספר השני הוא מספר המיכל, ואחריו נקודה.
 - שם החטיבה. נקודה. (לודא קודם ששם החטיבה כתוב בשם התיקייה, בדיוק כמו שבתוכנת הארכיון וכמו בארכיון עצמו).
 - שם המיכל בשתי מילים עם מקף-אמצעי, ונקודה,
 - שנת פתיחת המיכל, נקודה. (לדוגמה: 0.1.חינוך.ילדי-הגן.1930)
- ב. בתוך תיקייה של מיכל ארכיוני (שנוצרה כאן למעלה, בשלב 2א')- הסורק יוצר תיקיית תיק מסמכים סרוקים, לכל תיק ארכיוני. ואז נותן לה שם, בצורה של קוד:
- המספר הראשון הוא מספר החטיבה, ואחריו נקודה.
 - המספר השני הוא מספר המיכל, ואחריו נקודה.
 - המספר השלישי הוא מספר התיק, ואחריו נקודה.
 - לאחר מכן יקליד את שם המיכל בשתי מילים, עם מקף-אמצעי, נקודה.

- לאחר מכן יקליד את שם התיק בשתי מילים, עם מקף-אמצעי, נקודה.
- לאחר מכן יקליד את שנת פתיחת התיק, נקודה.
- לאחר מכן יקליד את הסיווג, (פומבי/פנימי/חסוי), נקודה.
- לאחר מכן יקליד מילה באנגלית (מעתיק את המילה האנגלית שהארכיונאית כתבה על התיק), ונקודה.

(לדוגמה: 0.1.2.ילדי-הגן.עמוס-כותב1930פנימי.children)

ג. סורק את המסמך בתוכנה הייעודית (במקרה שלנו ABBYY FineReader 12), ונתן שם לקובץ הסריקה.

שם הקובץ, יכלול את פרטי המסמך שניסרק, בצורה של קוד:

- מספר החטיבה, נקודה.
- מספר המיכל, נקודה.
- מספר התיק, נקודה.
- מספר הדף.
- שם התיק בשתי מילים עם מקף-אמצעי, ונקודה.
- שם המסמך בשתי מילים עם מקף-אמצעי, ונקודה. (או לחלופין את השפה אם היא לא עברית).
- שנת יצירת המסמך, ונקודה.
- סיווג (פומבי/פנימי/חסוי)
- לאחר מכן יזין את המילה באנגלית (מעתיק אותה משם התיק), ונקודה.

(שם קובץ לדוגמה: 0.1.2.3.עמוס-כותב.הטיול-בשדה1939פנימי.children).

הערה- שמות קבצים צריכים להיות בלי מרווחים, אלא עם נקודות לחציצה בין הנתונים השונים).

ד. חשוב- במקרים שבהם יש חוברת, או כמה מסמכים בנושא אחד- אז שומרים אותם לקובץ מסמך אחד משותף. (מגיע כך מהארכיונאית שמיינה בשלב א1)

ה. הסורק שומר כל קובץ ממסמך שסרק, בתוך תיקיית תיק מסמכים סרוקים שלו, (זו התיקייה שהוא יצר בשלב ב').

ו. סורק ברזולוציה 300 dpi

ז. בודק ויזואלית שהסריקות הן קריאות. (זו בקרת איכות- משימה שאפשר ל"סבלט").

ח. שומר את הסריקות שבוצעו ב FineReader Documents (Ctrl+Shift+S)

ט. מעביר לאיש המיחשוב את תיקיות הקבצים שסרק.

י. מחזיר את מיכל המסמכים לארכיונאית, כדי שהיא תחזירו למקום.

שלב 3 פעולות איש האחסון, כהכנה לקליטת הסריקות אל תכנת הארכיון

- א. מבצע חילוץ OCR על הקבצים הסרוקים אם אפשר (כתב יד לא ניתן)
- ב. הופך את קבצי ה-FineReader Documents לקבצי PDF. (ואם אפשר לקבצי PDF/a שניתנים לחיפוש) ומקפיד להשאיר את השמות שהסורק נתן לקבצים.
- ג. שומר את הטקסט שחילץ ב-OCR, אל קובץ טקסט, כדי שיכנס ל'דפית מידע נוסף' לאחר מכן.
- ד. (סעיף זה מוקפא והוא יופעל רק אם יוחלט אחרי הפיילוט, שצריך אכן לקצר את שמות הקבצים) גוזר מתוך שם של כל קובץ סרוק- את הסיווג, ומדביק אותו למטהדאטה שלו.
- ה. (סעיף זה מוקפא והוא יופעל רק אם יוחלט אחרי הפיילוט, שצריך אכן לקצר את שמות הקבצים) גוזר מתוך שם של כל קובץ סרוק- את המילה באנגלית, ומדביק אותה למטהדאטה שלו.
- ו. מפרק מכל המידע שנוצר (תיקיית מיכל, תיקיית מסמכים, שמות קבצים סרוקים) ויוצר קובץ XML שאותו ניתן לקלוט באידאה באופן נוח.
- ז. מעתיק את תיקיות תיקי המסמכים הסרוקים, מתוך תיקיות המיכלים- אל תיקיה אחת ב-IDEA FTP שקוראים לה 'ארכיון מסמכים סרוקים'.
- ט. מעביר את קובץ XML אל איש התכנה.

שלב 4. פעולות איש התכנה

- א. קולט את הקובץ XML אל תוכנת הארכיון, הקליטה יוצרת פריטים חדשים. ככה שלכול פריט בתכנה, מקושרים הקבצים הסרוקים, מתיק מסמכי נייר אחד.
- ב. בקובץ XML נמצא כבר הטקסט המיודעת לדפית "מידע נוסף" כך שהטקסט נכנס באופן אוטמטי לדפית.
- ג. במידה ויש בעיה עם קליטה של מידע נוסף מקובץ XML צריך להריץ שאילתה שתתקן את הטקסט הנקלט.
- ג. מעביר לבדיקת הארכיונאית ואיש המיחשוב שבארכיון.

שלב 5. פעולת הארכיון אית ואיש האחשוב

- א. יבדקו בתכנת הארכיון, האם הפריטים שנוצרו בתכנה פועלים ומאפשרים התמצאות בארכיון.
- ב. אם יש הערות אז ממשיכים מול איש התכנה.

סוף נוהל סריקות.

הערות:

משימה ניפרדת להמשך: כתיבת תיאור מורחב בתכנה, למסמכים הסרוקים. אפשר לבצע על ידי מתנדבים, גם נערים בצמד עם וותיקים.