

תפקיד הספרייה בניהול נתוני מחקר: המקרה של ספריית אוניברסיטת אדינבורו

מאת: תמר ישראלי, המחלקה ללימודי מידע, אוניברסיטת בר אילן והאקדמית גליל מערבי

תקציר

ניהול נתוני מחקר היא פעילות המבוצעת לאורך כל שלבי המחקר וכוללת את איסוף הנתונים, שמירתם במקום בטוח, גיבויים ושימורם באופן שיאפשר גישה ושימוש בהם לטווח ארוך. אחד המניעים העיקריים לעיסוק הגובר של אוניברסיטאות בניהול נתוני מחקר בעשור האחרון הוא הדרישה של גופים מממנים ומוציאים לאור לנקוט באמצעים שיבטיחו שמירה בטוחה של הנתונים, גישה לנתונים לצורך אישוש מסקנות המחקר ולשימוש חוזר במחקרים נוספים והעמדת הנתונים לרשות הציבור עד כמה שניתן מבחינה חוקית ואתית. מוסדות להשכלה גבוהה נדרשים לתמוך בחוקרים כדי שיוכלו לעמוד בדרישות הגופים המממנים ובסטנדרטים הנוקשים (במיוחד באירופה) לשמירה ושיתוף הנתונים. לצורך כך מוקצים משאבים כספיים שמושקעים ביצירת תשתיות מתאימות לניהול נתוני המחקר ובהעמדת כוח אדם מיומן לתמיכה בתחום זה. ספריות אקדמיות מנסות להרחיב את המיומנויות שיש לספרנים בניהול וארגון מידע כדי לקחת את התפקיד של תמיכה בניהול נתוני מחקר ולמלא את החלל שנוצר בין דרישות חדשות והתמחות מקצועית. אחת הספריות הפעילות בתחום זה היא ספריית אוניברסיטת אדינבורו המעמידה לרשות החוקרים צוות מיומן לתמיכה בניהול נתוני מחקר. מאמר זה סוקר את הפעילות של צוות התמיכה אליה נחשפתי במהלך השתלמות של חודש וממצאי פרויקט מחקרי בנושא שיתוף נתונים ומסמכים בין חוקרים.

מילות מפתח: ניהול נתוני מחקר; תמיכה במחקר; ספריות אקדמיות; תפקיד הספרייה; השתלמות

The Library's role in Research Data Management: The Case of the University of Edinburgh

Abstract:

Research Data Management (RDM) is a term that describes the organization, storage, and sharing of data collected during the lifetime of a research project as well as the practices that support long-term preservation, access, and use after the project has been completed. Today, an increasing number of funding agencies, journals, and other stakeholders are requiring data producers to share, archive, and plan for the management of their data. Both research funders and publishers increasingly expect that data resulting from research projects will be made available for scrutiny and re-use, whenever legal and ethical requirements allow it. Higher education institutions are required to support their academic staff in meeting these strict requirements but sometimes lack the expertise and knowledge as to how to provide suitable assistance for data storage and curation/preservation. Academic libraries are trying to fill these gaps by expanding the skills librarians in managing and organizing information and

taking the role of supporting research data management. One of the most active libraries in RDM support services is the University of Edinburgh Library with its team of skilled librarians. This article reviews the activities I was exposed to and findings from a research project on researchers' data and documents sharing practices.

Keywords: Research Data management; research support; Academic libraries; Library role; internship

מתחילת העשור השני של המאה ה-21 גוברת בעולם המערבי ההתעניינות בניהול נתוני מחקר. התעניינות זו מונעת בעיקר משינוי המדיניות של גופים מממנים הדורשים ממוסדות המחקר ומהחוקרים עצמם להנגיש את נתוני המחקר שלהם לציבור ולחוקרים אחרים. כדי להיענות לדרישה זו, החלו המוסדות האקדמיים לעסוק בפיתוח מדיניות, שירותים ותשתיות.

תמיכה בניהול נתוני מחקר היא אתגר לספריות האקדמיות שמעוניינות למלא תפקיד בתחום זה (Cox, Kennan, Lyon, & Pinfield, 2017). בארץ, תחום זה טרם פותח דיו, בין השאר בשל מיעוט קרנות המחקר העומדות לרשות החוקרים. בשלב זה לא הגיעה דרישה של קרנות המימון בארץ לעמוד בסטנדרטים לניהול נתונים תקין, ועל כן מוסדות המחקר והחוקרים אינם משקיעים את המאמצים והמשאבים הנחוצים לכך. עם זאת, סביר להניח שלאור העיסוק הגובר בעולם בנושא של מידע פתוח ונתונים פתוחים נראה דרישה כזאת בעתיד הלא רחוק. כבר היום חוקרים המעורבים במחקרים אירופאים ואמריקנים נתקלים בבעיות שיתוף נתונים. האיחוד האירופי אוסר על העברת נתונים למדינות מחוץ לאיחוד וביניהן ישראל וזאת בגלל שאינן עומדות בסטנדרטים של אבטחת המידע הנהוגים באירופה (Hoofnagle, van der Sloot, & Borgesius, 2019). נראה שבעיות אלה, והדילמות הנלוות להן, ילכו ויחמירו, והזמן נכון עכשיו להקים צוות שיוכל לתת פתרונות ולגבש מדיניות ותהליכי עבודה הקשורים לניהול נתוני המחקר בארץ. גם אם חוקרים בארץ אינם נדרשים לעמוד בדרישות גופים מממנים, ניהול נתוני מחקר והכנת תכנית לניהול הנתונים הן שיטות עבודה טובות שכדאי לאמץ לצורך שיתוף פעולה תקין בין חוקרים, העמדת הנתונים לשיפוט עמיתים בזמן סקירת מאמרים ומניעת אובדן נתונים או דליפת נתונים אישיים של נחקרים.

מה הם נתוני מחקר?

נתוני מחקר הם העדויות העומדות בבסיס התשובה לשאלת המחקר, וניתן להשתמש בהן כדי לאמת ממצאים ללא קשר לצורתם (למשל, הדפס, דיגיטלי או פיזי). נתוני מחקר עשויים לכלול סטטיסטיקות, אוספים של תמונות דיגיטליות, הקלטות קול, תמלילי ראיונות, נתוני סקר ותצפיות על עבודות שדה עם הערות מתאימות, פרשנות, יצירות אמנות, ארכיונים, חפצים שנמצאו ומאמרים שפורסמו. המטרה העיקרית של נתוני המחקר היא לספק את המידע הדרוש כדי לתמוך בתצפיות, בממצאים או בתפוקות של פרויקט מחקר ולאמת אותם. קבצי נתונים שונים מקבצים אחרים של המחקר – יש והם גדולים מאוד, ובמקרים רבים נדרשות תוכנות מיוחדות לצורך גישה לנתונים. חלק מהנתונים ייחודיים ולא ניתן לשחזרם אם אבדו או נהרסו, חלקם רגישים ונדרשים אמצעי אבטחה מיוחדים כדי להעבירם (Australian National Data Service, 2011).

ניהול נתוני מחקר

ניהול נתוני מחקר, הוא מונח כללי המכסה את האופן בו מארגנים, מאחסנים ומטפלים בנתונים שנאספו או נוצרו במהלך מחזור חייו של פרויקט המחקר – לפני התחלת הפרויקט, במהלכו ולאחר שהסתיים. הפעולות המבוצעות בניהול נתוני מחקר נועדו לאפשר גישה לנתונים ושימוש חוזר בהם והן כוללות: כתיבת תוכנית ניהול הנתונים (data management plan) לפני התחלת המחקר, ויצירה, תיעוד, שימוש, אחסון, שיתוף ושימור הנתונים לטווח הארוך במהלך ביצוע המחקר ואחריו (Ingram, 2016 ; Pinfield,) (Cox, & Smith, 2014).

שיתוף נתונים

בקהילה המחקרית גוברת ההכרה ביתרונות של שיתוף נתונים: השיתוף מאפשר אימות של תוצאות המחקר, מעמיד מאגר נתונים רחב יותר לרשות הקהילה המדעית ומאפשר שימוש חוזר בנתונים למחקרים עוקבים או שונים (Pinfield et al., 2014). הדחיפה לשיתוף נתונים מגיעה בעיקר ממוסדות המממנים מחקר והדורשים שקיפות בתהליך המחקרי. שיתוף הנתונים מאפשר גישה של חוקרים אחרים בתהליך שיפוט המאמר או התוצר האקדמי. מטרת שיתוף הנתונים היא לאפשר שימוש בנתונים לצורך ביצוע מחקרים נוספים באותו תחום או בתחומים אחרים, וכך לנצל טוב יותר את ההשקעה הכספית. גופי מימון בארצות הברית ובאירופה מחייבים כיום את החוקרים לעמוד בסטנדרטים נוקשים לשמירה על הנתונים ולאפשר גם לקהל הרחב גישה אליהם (Rice & Haywood, 2011).

ועדת המחקר הבריטית והנציבות האירופית הציגו סדרת עקרונות המבוססים על ההנחה שמחקר הממומן על ידי הציבור הוא משאב ציבורי. בתור שכזה, עליו להיות נגיש עד כמה שניתן ותוך זמן סביר בכפוף למגבלות הקיימות במתן גישה כזאת, כשהקו המנחה לגישה לנתוני מחקר ולתוצריו הוא "פתוח עד כמה שאפשר, סגור עד כמה שניתן" (UK Research and Innovation, 2018 ; European Union Open) (Data Portal, 2011).

פעילות מוסדות המחקר

בנוסף לדרישות של הגופים המממנים, למוסדות המחקר עצמם יש אינטרס לנהל נתוני מחקר: פרסום וציטוט הנתונים שנאספו במחקרים שנעשו במוסדות אלה עשוי להשפיע על שמו של המוסד ועל ההכרה בו. תיעוד מלא של תוצרי המחקר, של החוקרים ושל תלמידי המחקר יאפשר למצוא אותם ולהשתמש בתוצרים תוך ציון המקור (Bryant, Lavoie, & Malpas, 2017a). לשם כך ולצורך עמידה בדרישות הגופים המממנים, מוסדות להשכלה גבוהה משקיעים מאמצים לספק לחוקרים תשתיות וכלים מתאימים לניהול נתוני המחקר, להעמיד לרשותם כוח אדם מיומן שיוכל לעזור למלא את ההנחיות, לענות על שאלות ולפתור בעיות שמתעוררות בקשר לניהול הנתונים. שירותים אלה כוללים הדרכות, העלאת המודעות לחשיבות של ניהול תקין של נתוני מחקר, ייעוץ ומתן פתרונות לבעיות שצצות במהלך המחקר, כלים לאחסון נתונים רגישים ואחרים וכוננים מוסדיים לשמירת הנתונים לטווח ארוך והבטחת הגישה לנתונים (Bryant et al., 2017a). שירותי תמיכה במחקר אינם מערכת קבועה ונוקשה של שירותים אותם מעתיקים מאוניברסיטה אחת לשנייה אלא אסטרטגיה מותאמת לכל מוסד ומוסד. במחקר שערך OCLC (Online Computer Library Center), שהוא ארגון מחקר ללא מטרות רווח שמטרתו לעודד גישה לידע ולצמצם עלויות מידע, נבדקו שירותי ניהול נתוני מחקר של ארבע אוניברסיטאות. בשלוש יבשות שונות כמקרה בוחן. במחקר זהו שלוש קטגוריות של שירותי תמיכה בניהול נתוני מחקר:

הדרכות - מטרתן להעלות את המודעות לחשיבות של ניהול נתוני מחקר, לעודד פיתוח מיומנויות בתחום זה וליידע את החוקרים על הכלים והשירותים הקיימים לתמיכה בניהול נתוני המחקר

אוצרות נתונים – תשתית טכנולוגית ושירותי תמיכה בניהול נתוני המחקר לאורך תקופת ביצוע המחקר. אוצרות הנתונים מאפשרת גישה לנתונים בטווח הארוך לצורך שכפול המחקר או לטובת מחקרים נוספים.

מומחיות – תמיכה בקבלת החלטות ומציאת פתרונות בנוגע לבעיות ניהול נתוני המחקר.

כל אחת מהאוניברסיטאות שנבדקו מטמיעה שירותים אלה ברמה שונה ושמה דגש על מרכיבים שונים.

ספריות אקדמיות ותמיכתן בניהול נתוני מחקר

אופי המעורבות של הספריות בניהול נתוני מחקר, המידה שבה הספריות צריכות להוביל את הנושא, אלה שירותים צריך לספק ואלה תשתיות לפתח – כל אלה הם נושאים לדיון מקצועי. הספרנים יכולים לתרום לקידום הנושאים הללו באמצעות המיומנויות שיש להם בארגון מידע, ביצירת מטה-דאטה ובאחזור נתונים (Cox & Pinfield, 2014), ואכן, יש ספריות שמובילות יוזמות, במיוחד בכל הקשור לפיתוח מדיניות. יחד עם זאת הספריות מתמודדות עם אתגרים כמו רמת מעורבות נמוכה של בעלי העניין במוסד והחוקרים עצמם, פערי ידע בתחום של התשתית הטכנית הנחוצה, ידע משפטי והבנת סוגי הנתונים בדיסציפלינות השונות. מעבר לכך יש קושי לגייס מימון לפעילות התמיכה בניהול נתוני מחקר (Cox & Pinfield, 2014 ; Cox et al., 2017).

למרות האתגרים, חוקרים רבים טוענים שהספריות הן אלה שצריכות להוביל את סדר היום בניהול נתוני המחקר (Lewis, 2010; Cox, Verbaan, & Sen, 2012; Corral, 2012). ליון (Lyon, 2012) מנה 10 תחומים שבהם יכולה הספרייה לתמוך במחזור חיי המחקר:

1. ריכוז הדרישות של ניהול נתוני המחקר
2. עזרה ויעוץ לחוקרים בהכנת תוכנית לניהול נתוני המחקר
3. יעוץ טכני בנוגע לפורמט הנתונים והמטה-דאטה
4. ציטוט נתוני מחקר
5. הדרכות בנושא ניהול נתוני מחקר
6. מתן תעודות הסמכה בתחום
7. הערכת הנתונים והחלטה אילו כדאי לשמור
8. אחסון נתוני מחקר (יחד עם אנשי מערכות מידע)
9. גישה לנתוני מחקר
10. קידום ההשפעה של נתוני המחקר (יחד עם רשות המחקר)

השתלמות בניהול נתוני מחקר בספריית אוניברסיטת אדינבורו

אחת הספריות האקדמיות ששימשו כמקרה בוחן לדוח ה OCLC בנוגע לשירותי התמיכה בניהול נתוני מחקר היא ספריית אוניברסיטת אדינבורו בסקוטלנד. אוניברסיטת אדינבורו אימצה ב-2011 מדיניות ניהול נתונים כחלק ממחויבותה למצוינות במחקר, מדיניות שהאוניברסיטה מקצה לה משאבים, וכך תורמת לעמידה של החוקרים בסטנדרטים הגבוהים שמציבות מועצת המחקר הבריטית והיוזמה האנגלית למצוינות במחקר. אוניברסיטת אדינבורו מציעה חבילת שירותי מידע (Information Services Group) הכוללת את הספרייה, טכנולוגיות מידע, מרחבי לימוד ומרכז מומחיות דיגיטלית הממוקם באוניברסיטה.

בזכות המעורבות בפיתוח מקורות הדרכה ותרומתה לקהילה הרחבה העוסקת בניהול נתוני מחקר, הפכה ספרייה זו לאחת המובילות בתחום (Bryant, Lavoie, & Malpas, 2017b). בספטמבר 2019 נסעתי לאדינבורו להשתלמות בת ארבעה שבועות, השתלמות שכללה עבודה עם צוות התמיכה בניהול נתוני מחקר, השתתפות בישיבות, תצפית, שיחה עם בעלי עניין במחלקת התמיכה במחקר, הצטרפות להדרכות הצוות וביצוע פרויקט מחקרי לטובת המחלקה לתמיכה בניהול נתוני מחקר.

שירותי התמיכה בניהול נתוני מחקר בספרייה אוניברסיטת אדינבורו

באוניברסיטת אדינבורו הממוקמת במקום העשרים בדירוג האוניברסיטאות העולמי (QS World University Rankings, 2020) לומדים כ-40,000 סטודנטים בקמפוסים הממוקמים ברחבי העיר ויש בה כ-4000 חברי סגל אקדמי במשרה מלאה. צוות ספרייה האוניברסיטה מונה כ-700 עובדים. בצוות התמיכה בניהול נתוני מחקר יש שבעה עובדים במשרה מלאה והוא חלק ממחלקה גדולה יותר של תמיכת הספרייה במחקר. הצוות ממוקם יחד עם כ-100 אנשי שירותי מידע בבניין נפרד שאינו חלק מהספרייה המרכזית, זאת בשל חוסר מקום בספרייה המרכזית ועבודה בסמיכות לאנשי תמיכה בשירותים שונים, עבודה הדורשת שיתוף פעולה הדוק עם לבניית הכלים הנחוצים לניהול נתוני המחקר ולתחזוקתם. לאנשי צוות התמיכה התמחויות בסוגי נתונים שונים והם מעבירים הדרכות כלליות ונושאות על ניהול נתוני מחקר. אנשי הצוות גם חברים בצוותים ובאגודות לקידום מדע פתוח, אוצרות דיגיטליות ומאגרים מוסדיים.

צוות מחלקת התמיכה בניהול נתוני מחקר יושב במתחם משרדים פתוח. בקומת המשרד ישנם תאים קטנים עם מחשב ומסך המשמשים לישיבות קטנות של שניים-שלושה אנשים. ישנה קומה עם חדרים גדולים יותר המוזמנים מראש ומשמשים לישיבות עם מספר משתתפים גדול יותר. העבודה השוטפת כוללת פגישות פנימיות וחיצוניות, בניית אבני דרך לשנים הבאות, עדכון מדיניות המחלקה, הדרכות הנערכות ברחבי הקמפוסים השונים בעיר, התייעצות בנושאים משפטיים, בניית דפי מידע עבור תלמידים חדשים המגיעים לאוניברסיטה ועבור חוקרים ותלמידי תארים מתקדמים, ארגון כנסים, עדכון הבלוג ודף הטוויטר של המחלקה ועבודה עם משתלמים שמגיעים למחלקה. מדי שנה מגיעים תלמידי תארים מתקדמים לעבודה במחלקה בשכר ומבצעים פרויקטים שונים. אחת המשתלמות למשל, בנתה חוברת מידע לתלמידי תואר שני בנושא עבודה עם נתונים בזמן כתיבת התזה. בחוברת פורטו כל הנושאים שהמחלקה עוסקת בהם: מה הם נתונים, ארגון קבצים, היכן ניתן למצוא נתונים משניים והיכן ניתן לשמור אותם, מה הם נתונים רגישים ואיך מתמודדים עם פרטיות ודאגים לאנונימיות, מהם מטא-דאטה ואיך מתעדים אותם, כיצד מתמודדים עם ריבוי גרסאות ואיפה מאחסנים את הנתונים לאחר הגשת העבודה (Sanchez, 2019).

בשבוע הראשון של הסמסטר, סטודנטים חדשים עוברים אוריינטציה ומשתתפים בקבלות פנים. חלק מהצוות עוסק בהכנת תיקיות עם עלונים של שירותי המחלקה ומחלק אותם בקמפוסים השונים. כל תיקיה כזאת כוללת את הפריטים הבאים: עלון בצורת טבלה שבו שאלה כללית - איך אתה מנהל את הנתונים שלך כרגע? ושאלות נוספות לגבי תכנון המחקר הנוגעות לארגון, שמירה וגיבוי הנתונים למשל: איך אתה מארגן את נתוני המחקר שלך? איך ומתי אתה מגבה נתונים, איך אתה מתעד את ניתוח הנתונים, איך אתה מתכנן לשתף את נתוני המחקר לאחר פרסומו? לצד כל שאלה יש ארבעה טורים עם תשובות לשאלות וככל שמתקדמים שיטת העבודה נהיית מסודרת ומאורגנת יותר מבחינת ניהול הנתונים. לדוגמה, התשובה לשאלה הראשונה מתפתחת מ: "באשר לנתונים שלי, יש לי דרכים משלי לעשות דברים אבל אין לי תוכנית מתועדת" ועד: "אני מכין תוכניות לניהול הנתונים לצורך שימוש עתידי שלי או של

אחרים". עלון אחר מפנה לקורס מקוון שבנה צוות המחלקה לניהול נתוני מחקר. הקורס נקרא Research Data MANTRA (MANTRA= Management TRAIing) והוא נועד לשקף את הפרקטיקות הטובות ביותר לניהול נתונים. החומרים מיועדים לתלמידים מתקדמים ולחוקרים צעירים והקורס כולל הסברים, תרגילים וראיונות מצולמים עם חוקרים מובילים המספרים על האתגרים העומדים בפניהם בניהול נתוני מחקר. החומר מחולק למספר יחידות לימוד ומאפשר התקדמות בזמן ובקצב הנוחים לכל מתעניין. עלון נוסף שמקבלים הסטודנטים כולל 10 טיפים לאבטחת מידע ובאחר ניתנים פרטי יצירת קשר עם המחלקה והזמנה להשתתף בהדרכות. בנוסף, יש חוברת דקה ובה מדריך כללים שמספקת האוניברסיטה לחוקרים לאחסון ושיתוף הנתונים שלהם לפני המחקר, במהלכו ולאחר סיומו, וסימנייה עם הלוגו ופרטי המחלקה לתמיכה בניהול נתוני מחקר. התיקייה והתוכן שבה משמשים להעברת מידע לגבי שירותי המחלקה וגם לתזכורת חוזרת על קיומם של התמיכה והשירותים שהיא מספקת.

הדרכות לחוקרים ותלמידי מחקר

את הדרכות בנושא ניהול נתוני מחקר מעביר צוות התמיכה בניהול נתוני מחקר באוניברסיטת אדינבורו, והן מיועדות בעיקר לתלמידים מתקדמים. המחלקות הן המזמינות את ההדרכות ויש שהסטודנטים נרשמים אליהן מתוך עניין וצורך אישי. משך ההדרכות הוא כשלוש שעות (כולל הפסקה) והן נערכות בקמפוס בו לומדים הסטודנטים. צוות התמיכה פועל להעביר את ההדרכות בצורה אינטראקטיבית באמצעות יצירת דיון בין המשתתפים בשאלות הנוגעות לאותה הדרכה. העובדה שההדרכות נעשות בקמפוסים מחייבת את אנשי הצוות לטייל בין הקמפוסים השונים בעיר. נוסף להדרכות הפרונטליות אוניברסיטת אדינבורו מציעה שני קורסים מקוונים, אחד הוא קורס המנטרה שבנתה המחלקה לתמיכה בניהול נתוני מחקר <https://mantra.edina.ac.uk/> והשני הוא קורס MOOC של קורסרה (Coursera) שנבנה בשיתוף עם אוניברסיטת צפון קרולינה בארצות הברית <https://www.coursera.org/learn/data-management>. נושאי ההדרכות הם שיטות עבודה עם נתוני מחקר, הכנת תוכנית ניהול נתונים להצעת מחקר, שימוש בנתונים ב SPSS, ניקוי נתונים ועבודה עם נתונים רגישים.

כלים ושירותים שמעמידה אוניברסיטת אדינבורו לרשות החוקרים

כדי לעמוד בדרישות הגופים המממנים, האוניברסיטה מעמידה לרשות החוקרים כלים ושירותים שיסייעו להם לנהל את נתוניהם בצורה מסודרת ומאובטחת ולשתף פעולה עם עמיתיהם למחקר. עם הכלים והשירותים האלה נמנים:

טופס תוכנית ניהול נתונים: מוסדות רבים דורשים לצרף להצעת המחקר תוכנית לניהול הנתונים. האוניברסיטה מציעה לחוקרים טופס מקוון שנועד לעזור להם במילוי תוכנית כזאת. את הטופס הכין המרכז הבריטי לאוצרות דיגיטלית (DCC - The Digital Curation Centre) הממוקם באוניברסיטת אדינבורו.

אחסון נתונים באמצעות DataStore: כלי שנועד לשמירת נתוני מחקר אקטיביים (בזמן ביצוע המחקר) ועומד לרשות חוקרים ותלמידי מחקר. DataStore נותן מרחב אחסון נתונים אישיים וקבוצתיים וניתן להרחיבו במידת הצורך. אחד היתרונות של כלי זה הוא שהוא מגובה באופן מלא ולפיכך בטוח ויציב. ניתן לגשת לנתונים ממקומות שונים וממכשירים שונים.

Data Safe Haven: כלי המעניק סביבה בטוחה ומבוקרת לשימוש בנתוני מחקר רגישים (אישיים או חסויים). הכלי מאפשר העברת נתונים רגישים לסביבה מוגנת שבה ניתן לשמור אותם באופן בטוח, לעבוד

ולנתח את הנתונים. בסביבה זו יכולים חוקרים המורשים לכך לעבוד יחד תוך עמידה בדרישות של בעלי הנתונים, לדוגמה, שירותי רפואה לאומיים.

DataSync: כלי המאפשר לחוקרים לשתף נתונים עם עמיתיהם ולסנכרן את הנתונים במכשירים שונים. הכלי כולל ממשק ווב (Web) שמאפשר לשתף נתונים עם כל מי שיש לו כתובת דוא"ל תקפה בכל דפדפן בו הוא משתמש. יש בו אפליקציה המאפשרת סנכרון בין המחשב למכשירים ניידים.

GitLab: כלי לניהול גרסאות שמאפשר לעקוב אחרי שינויי קוד ומסמכים.

מחברות מעבדה אלקטרונית: חלופה למחברות נייר. אפשר להשתמש בהן כדי להכניס פרטים של חלק מהמחקר כמו למחברות המעבדה המסורתיות אך המחברת האלקטרונית מאפשרת גם הכנסת קישורים לקבצים דיגיטליים אחרים כמו תמונות, מודלים ופרוטוקולים. היתרון הגדול הוא שניתן לשתף מחברות אלה בין חברי הקבוצה ללא תלות במקום בו הם נמצאים וגם להפוך אותן לנגישות לציבור.

Pure: מערכת הכריס (CRIS – Current Research Information System) של האוניברסיטה. היא כוללת פרופילים אישיים של תוצרי המחקר והפרויקטים חוקרים ואנשי צוות. האוניברסיטה משתמשת בPure כדי לתעד מטה-דאטה על נתוני מחקר לצורך עמידה בדרישות הגופים המממנים. המטה דאטה מאפשרת לאחרים להבין איזה נתונים קיימים, איך נאספו וכיצד לגשת אליהם.

DataShare: מאגר נתונים חופשי ופתוח של אוניברסיטת אדינבורו המאפשר לחוקרים להעלות נתונים, לשתף בהם וליצור רישיון לשימוש בנתונים אלה לאחר סיום המחקר כך שניתן יהיה למצוא אותם ברשת ולעשות בהם שימוש חוזר. כלי זה מאפשר להגדיל את השפעתו של המחקר, הופך את הנתונים לברי גילוי וציטוט, מפחית את החשש לאובדן מידע ומבטיח שימור הנתונים לטווח ארוך.

DataVolt: שירות אחסון ארכיוני לטווח ארוך שבו הסגל האקדמי יכול לאחסן בביטחון נתונים שאינם פתוחים לציבור. נתונים אלה נשמרים באופן בטוח כך שלא ניתן למחוק אותם בטעות או לגשת אליהם בצורה לא נאותה.

נוסף לכלים אלה ישנם כמובן שירותי התמיכה של צוות התמיכה במחקר.

הפרויקט המחקרי

אף שהאוניברסיטה משקיעה משאבים כספיים רבים בפיתוח ובתחזוקה של הכלים הללו, ולמרות ההשקעה הכספית והעמדת צוות תמיכה מיומן, החוקרים אינם עושים שימוש רחב בהם. במסגרת ההשתלמות נערך מחקר שמטרתו הייתה לנסות להבין את פרקטיקות שיתוף הפעולה המחקרי של חוקרים ואת החסמים המונעים מהם להשתמש בכלי ניהול הנתונים שהאוניברסיטה מעמידה לרשותם.

שיטה

כלי המחקר היה ראיונות מובנים למחצה. המרואיינים היו שישה חוקרים מתחומי דעת שונים המשתפים פעולה עם חוקרים אחרים על בסיס קבוע ומעבירים ביניהם קבצי מסמכים ונתונים. המשתתפים גויסו באמצעות פנייה בדוא"ל לחוקרים ששלחו הצעות להרצאות בכנס שארגנה המחלקה, פנייה כללית באמצעות טוויטר ודרך קבוצות עבודה משותפות. הראיונות נערכו במשרדי החוקרים בקמפוסים השונים ברחבי העיר, ואחד הראיונות נערך בספרייה המרכזית. כל המרואיינים סימנו בטופס ההסכמה מדעת (ראו נספח 1) כי הם מוכנים שפרטי הראיון יופיעו תחת שמם המלא. הראיונות, שארכו בין חצי שעה לשעה, נפתחו בשאלה כללית על האופן בו הם עובדים עם שותפיהם. בהמשך, נשאלו המרואיינים על הכלים שבהם הם משתמשים להעברת מסמכים ונתונים, הסיבות לבחירת כלים אלה, באלה בעיות הם נתקלים, איך בונים תוכנית ניהול נתונים, איך הם מתמודדים עם ריבוי גרסאות, מה הם עושים כדי להגן על הפרטיות של הנבדקים ואלה אמצעים הם נוקטים כדי להימנע מאובדן מידע. לבסוף נשאלו על שירותי

התמיכה במחקר וכיצד ניתן לקדם. הראיונות הוקלטו אך לא תומללו. תקנות הגנת הפרטיות באירופה קובעות שלא ניתן להוציא נתונים למדינות שאינן חברות באיחוד האירופי ולכן ניתוח הנתונים נעשה במגבלת משך ההשתלמות. מתוך האזנה להקלטות הוצאו תמונות מרכזיות. לאחר כתיבת הדוח המחקרי בו הוצגו הנושאים שעלו בראיונות הועבר הדוח למרואיינים לצורך סקירה ואישור.

אישור לבצוע מחקר

לצורך עריכת המחקר נדרש תהליך (Data Protection Impact Assessment) DPIA שתפקידו לעזור לנתח, לזהות ולהקטין למינימום את הסיכונים בהגנת הנתונים של הפרויקט המחקרי. תהליך זה הינו חלק מהדרישה של (General Data Protection Regulation) GDPR – רגולציה של האיחוד האירופי לגבי הגנת המידע והפרטיות של תושבי האיחוד האירופי. התהליך כולל מילוי שאלון שבו כמה מודולים הבנויים בהתאם לשלבים שונים של המחקר ושאלות רבות ומפורטות לגבי האופן בו יבוצע, הנתונים שייאספו, למה ישמשו, איפה יאוחסנו וכמה זמן יישמרו לאחר סיום המחקר. לאחר מילוי השאלון צריך לחכות לאישור ועדה כדי שאפשר יהיה להתחיל את המחקר. לאחר קבלת האישור החל ביצוע הראיונות. המשתתפים קראו את דף המידע בנוגע למחקר (ראו נספח 2) וחתמו על טופס ההסכמה מדעת.

ממצאים

כלים לשמירה ושיתוף מסמכים ונתונים

החוקרים הזכירו כלים מוסדיים ומסחריים בהם הם משתמשים כדי לאחסן, לשתף ולערוך מסמכים ונתונים. כלים אלה היו דוא"ל (דואר אלקטרוני), גוגל דוקס, גוגל דרייב, דרופבוקס, וויציט (הגרסה הסינית לפייסבוק), אופיס 365, פלטפורמות וויקי וכוננים ארגוניים. ההחלטה במה להשתמש נבעה מגורמים שונים כגון: מי השותפים? האם הם מתוך המוסד או מבחוץ? גודל הקבצים, רמת רגישות הנתונים, מה מטרת השיתוף? (למשל כתיבת מאמר או ניתוח הנתונים). המשתתפים ציינו שהם משתמשים בדוא"ל או בכלים מסחריים כמו גוגל דרייב בעיקר כשהם עובדים עם חוקרים מחוץ לארגון ומחוץ למדינה. "דוא"ל תמיד עובד" אמר ד"ר תומס זלינסקי, מנהל משאבי מחשב בפקולטה לביולוגיה. השימוש בכלים המסחריים נעשה בדרך כלל בהעברת מסמכים ונתונים לא רגישים כי כלים אלה נוחים וידידותיים למשתמש ומוכרים לשותפים למחקר. ד"ר סיריל פרנט, חוקר מדעי המוח הקוגניטיביים, הדגיש שהוא וחבריו משלמים מכספם הפרטי עבור מקום אחסון נוסף בגוגל דרייב, על אף שהאוניברסיטה מספקת מקום אחסון רב על הכונן המוסדי. יש לציין כי הכלים אותם מספקת האוניברסיטה מאפשרים עבודה עם שותפים חיצוניים אך אלה צריכים לפתוח חשבון חיצוני. כאשר הם עובדים עם שותפים שונים במדינות שונות, נוח להם יותר לעבוד עם הכלים המסחריים שנגישים לכולם. ד"ר פרנט הוסיף שכאשר הוא עובד מהבית או ממדינה אחרת, קל לו יותר להשתמש בכלים המסחריים כי הכניסה לחשבונות של האוניברסיטה דורשת התחברות באמצעות VPN, כלומר מחייבת עוד כמה צעדים.

פלטפורמות וויקי הוזכרו ככלי יעיל לעבודה עם מסמכים, רשימות, פרוטוקולים ואחסון נתונים המשמשים בעיקר לכתיבת מאמר. המשתתפים ציינו פלטפורמות אלה לעיתים קרובות ככלי השיתוף בתוכנית ניהול הנתונים המצורפת להצעת המחקר שכן פלטפורמות אלה הן כלי אינטואיטיבי ופשוט. לצורך פרויקטים הדורשים ניתוח נתונים מיוחד נבנים כלים מתוחכמים אך הצוותים משתמשים בכלים הפשוטים יותר כשניתן.

לצורך עבודה והעברת קבצים גדולים הוזכרו כלים כמו WeTransfer, כונן השיתוף של האוניברסיטה שנקרא DataSync, דרופבוקס, ופלטפורמות מדע פתוח (Open Science) שמאפשרות לשלוח קישור מכל כלי.

בעיות וחסימים לשימוש בכלים שמציעה האוניברסיטה

מתוך הראיונות עלו כמה בעיות שמונעות מהחוקרים להשתמש בכלים של האוניברסיטה

הכלים לא מתאימים לצרכים המיוחדים של קבוצת המחקר

הבעיה בחלק מהכלים שמציעה האוניברסיטה היא שהם מספקים את התכונות הבסיסיות לקהל הרחב. החברות שמספקות את הכלים אינן מפתחות אותם לפי דרישה מאחר שלא יהיה לפיתוחים אלה מספיק ביקוש. לא פעם זקוקות קבוצות מחקר לכלים מיוחדים לצורך המחקר. כלי אחסון הנתונים של האוניברסיטה בנויים כמיכל של קבצים והם אינם מתאימים למבנה הנחוץ לארגון ועבודה עם נתוני המחקר. פרופסור מילר, ראש המחלקה למערכות ביולוגיות, וד"ר פרנט ציינו שבתחומים שלהם (ביולוגיה ומדעי המוח) קבוצות המחקר מסתמכות על סטנדרטים בינלאומיים יותר מאשר על המערכות הכוללות של האוניברסיטה. ד"ר זלינסקי משוכנע ששיתוף הוא רק מטרה משנית עבור החוקרים ושהם ישתמשו בכלי מסוים רק אם יהיה לו ערך נוסף עבור המחקר שלהם. חוקרים לא ישקיעו זמן בהעלאה מסודרת ומאורגנת של קבצי הנתונים אלא אם וכאשר יידרשו מהם זאת ואז סביר שלא יזכרו את כל הפרטים הנחוצים לתיאור הקובץ. גם פרופסור מילר חשב שנחוץ ערך מוסף לשיתוף הקבצים, למשל דרך נוחה לניתוח הנתונים.

הכלים אינם מספיק ידידותיים למשתמש ואינם פועלים במיזוג עם מערכות אחרות

כל המשתתפים הסכימו שעל הכלים להיות יותר ידידותיים למשתמש אם המוסדות רוצים שהחוקרים יעבדו איתם. כפי שהדברים עומדים היום, נראה שהחוקרים ישתמשו בכלי הפשוט ביותר אלא אם יכפו עליהם שיטה אחרת. הולי טיבל, דוקטורנטית העוסקת בסטטיסטיקה, ציינה שאחת הבעיות עם הכלים של האוניברסיטה היא שחלק מהמערכות לא מאפשר אינטגרציה עם כלים אחרים הנחוצים לחוקר.

כשדובר על ניהול גרסאות הייתה הסכמה בין המשתתפים שאנשים שאינם מתכנתים לא ישתמשו במערכות ניהול גרסאות כמו GitHub ו Subversion. לדעתם מערכות אלה מסובכות מדי, גם אם ישתתפו בקורסים שמציעה האוניברסיטה לחוקרים שאינם אנשי תוכנה. בנוסף לכך, מערכות אלה אינן מיועדות לניהול גרסאות של מסמכים שכן הן פועלות בצורה בינארית. כלים בהם משתמשים החוקרים כדי להתמודד עם בעיית ריבוי הגרסאות של מסמכים הם כלים השומרים את הגרסה האחרונה בלבד ומציעים היסטוריית גרסאות כמו גוגל-דוקס, וויקי ושר-פוינט.

בטיחות הנתונים ופרטיותם

לעבודה עם נתונים רגישים מספקת האוניברסיטה כלים מיוחדים. לעיתים נדרשים החוקרים לעבור הכשרה מיוחדת ומבחן כדי שיוכלו להשתמש במערכות אלה. הולי טיבל, שעובדת עם נתונים רגישים, סיפרה שמנחי העבודה שלה לא יכולים לגשת לנתונים מאחר שלא עברו את המבחן. היא עצמה מודעת לחשיבות השמירה על הנתונים ומפעילה אמצעי בטיחות וזהירות גם עבור קבצים פחות רגישים אך טוענת שלעיתים לוקח לה חצי שעה להגיע לנתונים בגלל תהליך הזהירות ארוך ומייגע.

חשש לאובדן מידע

החוקרים לא חששו מאובדן מידע. הם שומרים את הנתונים בענן או בכוננים מוסדיים שבהם נעשה גיבוי באופן קבוע ודואגים בעצמם ליצור עותקים נוספים. יחד עם זאת הם מודעים לאפשרות של אובדן מידע של נתונים שנשמרו על גבי דיסקים או כוננים חיצוניים ולאפשרות שחברות המספקות שירותי ענן יפשוטו רגל. בעיה נוספת עם החברות המסחריות היא בטיחות הנתונים. גם כשדואגים לאנונימיות יש אפשרות להצליב את המידע ממקורות שונים. למרות מודעות זאת, השימושיות של הכלים המסחריים גוברת על החשש לבטיחות ופרטיות.

קידום שירותי התמיכה במחקר

המשתתפים הסכימו שצוות התמיכה בניהול נתוני המחקר עושה עבודה טובה. וונלונג לי, דוקטורנט בתחום המשפטים, חשב שצריך לפרסם את השירותים לעיתים קרובות יותר כי כשהסטודנטים מגיעים הם מוצפים במידע. ד"ר בריאן קאהיל, מנהל פרויקטים במחלקה להנדסה, ציין שמניסיונו, כשהחוקר הראשי או ראש הקבוצה מכתיבים דרך מסוימת, שאר החוקרים יפעלו בדרך זו ולכן כדאי לצוות התמיכה להפנות את מאמציהם להכשרה, הדרכה ותמיכה בחוקרים הראשיים.

דיון

המחקר נועד לבדוק את הכלים שבהם משתמשים חוקרים לצורך שיתוף נתונים ומסמכים, ואת החסמים המונעים מחוקרים אלה להשתמש בכלים שמעמידה לרשותם האוניברסיטה. הממצאים מצביעים על כך שמשתתפי המחקר עושים שימוש גם בכלי האוניברסיטה וגם בכלים של חברות מסחריות בהתאם לצורך ולנחות. למרות המודעות ליתרונות של כלי האוניברסיטה בשמירת הנתונים ובאבטחת המידע לטווח ארוך, החוקרים מעדיפים במקרים רבים את הכלים המסחריים שהם לדבריהם ידידותיים יותר למשתמש. המשתתפים דיברו על הצורך בכלים פשוטים, כאלה שמוכרים לכל שותפי המחקר. סיבה נוספת לבחירה בכלים המסחריים הייתה עבודה עם שותפים מחוץ לארגון. שיתוף באמצעות כלי האוניברסיטה מחייב את השותפים לפתוח חשבון חיצוני מיוחד בזמן שחוקרים ושותפיהם עובדים עם קבוצות מחקר שונות ומתאימים עצמם לדרך העבודה שנוחה לשותפים. שיקול נוסף בבחירת כלי השיתוף היה העדפות החוקר הראשי. ממצאים אלה מתקשרים למחקרים אחרים המסבירים את הבחירה בכלי השיתוף על פי ההעדפות של משתמשים אחרים או על פי הניסיון וההרגלים הקיימים ולא על בסיס התכונות הספציפיות של הכלי (Rader 2009; Capra, Vardell, & Brennan, 2014). המשתתפים הזכירו פעמים רבות את הדוא"ל ככלי לשיתוף מסמכים או קבצים שאינם גדולים. העדפת הדוא"ל ככלי פשוט ויעיל שמאפשר לכל אחד שימוש בתוכנה הנוחה לו, מתועדת בספרות המחקרית (פלך, 2013; ישראלי, 2019; Bergman & Whittaker, 2016). השימוש בדוא"ל לשיתוף קבצים נמשך גם כיום, למרות קיומם של כלים חדשים ונוחים לניהול ושיתוף מידע קבוצתי.

המשתתפים דיברו על יתרונותיהם של כלי השיתוף הפשוטים אך באותה עת תיארו כלים מתוחכמים אותם הם בונים במיוחד לצורך המחקר ותוכנות וכלים שבהם הם לומדים להשתמש מאחר שהם דרושים לצרכי המחקר. בין המשתתפים היו שהדגישו ששיתוף הנתונים הוא רק מטרה משנית ושחוקרים ישתמשו בכלי רק אם יהיה לו ערך למחקר שלהם. ניתן להבין מכך שחוקרים מוכנים להשקיע מאמץ בלמידה ובהטמעת כלים חדשים, אך ערך השיתוף והפעולות הנדרשות לצורך שיתוף כפי שמתבטאות בדרישותיהם הברורות של גופי המימון ומוסדות המחקר אינו עומד בראש סדר העדיפויות שלהם. ממצא זה תואם את מחקרם של מרקאווקייט ושותפיו (Markauskaite, Kennan, Richardson, Aditomo, & Hellmers, 2012) שבדקו את הנכונות לשיתוף בקרב חוקרים באוסטרליה. סיבה נוספת שעשויה להסביר את

ההתעלמות החוקרים מהדרישות היא שאין אכיפה של הכללים. לאחר מילוי טופס התוכנית לניהול נתונים אף אחד לא בודק מה נעשה עם הנתונים לאורך הליך המחקר, כך עלה משיחות עם אנשי צוות התמיכה בניהול נתוני המחקר. כאשר מדובר בנתונים המוגדרים רגישים, המוסדות או הארגונים שמספקים את הנתונים דואגים בעצמם לשמירה על הנתונים על ידי הגבלת הגישה אליהם במערכות ובדרכים שפורטו למעלה.

משתתפי המחקר משתמשים בכלים המסחריים למרות מודעותם לנושאים של פרטיות, בטיחות מידע ועמידה בדרישות הגופים המממנים ואפילו משלמים מכיסם הפרטי כדי להרחיב את יכולת אחסון הקבצים בכלים אלה. ממצא זה חשוב לשיקולים שמפעילים מוסדות המחקר בפיתוח כלים ותשתיות עבור החוקרים. ייתכן שההשקעה הגדולה בתקצוב ופיתוח כלי האוניברסיטה לא מניבה את התוצאות המקוות. אחד הפתרונות שטבעי לשקול לאור הממצאים ואף מיושם במידה מסוימת במוסדות שונים הוא הטמעת הכלים המסחריים כמו הכלים של גוגל ודרופבוקס בתוך שירותי האוניברסיטה. כך, האוניברסיטה העברית בירושלים מעמידה לרשות החוקרים חשבונות דוא"ל במערכת Gmail. באירופה, בה נושא הפרטיות נמצא במקום גבוה בסדרי העדיפויות ובא לידי ביטוי בתקנות ה-GDPR, הסבירות לבחירה בפתרון זה היא נמוכה.

סיכום ומסקנות

נושא ניהול נתוני המחקר באוניברסיטת אדינבורו בפרט ובבריטניה בכלל מפותח מאוד. ספריית אוניברסיטת אדינבורו לוקחת חלק פעיל ונרחב בתמיכה בתחום זה. האוניברסיטה משקיעה משאבים כספיים וכוח אדם מיומן כדי לספק לחוקרים את הכלים והשירותים הנחוצים להם לצורך ניהול נתוני המחקר. אחד הגורמים המרכזיים לפיתוח כלים ושירותים אלה הוא דרישות הגופים המממנים את המחקר. גופים אלה דורשים שהנתונים יהיו נגישים לשימוש חוזר לצורך אישוש מסקנות המחקר וכן לשימוש במחקרים נוספים באותו תחום דעת או באחרים. הם גם דורשים שהנתונים ינוהלו כראוי בצורה שמבטיחה הן מניעת אובדן והן שמירה מפני דליפת נתונים חסויים או אישיים. צוות התמיכה בניהול נתוני מחקר זוכה לשיתוף פעולה צמוד מאנשי מערכות המידע. הצוות דואג להפיץ את שירותיו באמצעות נוכחות בימי האוריינטציה של הסטודנטים, הדרכות יזומות ומוזמנות לתלמידי מחקר, פגישות עם חוקרים ובעלי עניין, ארגון כנסים, השתתפות בכנסים ואירועים, חברות בוועדות ותמיכה אישית לפונים בשאלות. הצוות מיומן בתחומים שונים הנוגעים לניהול נתונים, כללי האתיקה וחוקים הנוגעים לשמירת נתונים בבריטניה ובאיחוד האירופי ונעזר בוועדות משותפות כדי לקבוע כללים בנושאים בהם חסרים קיום מנחים.

שאלתי את רובין רייס, ראשת הצוות, מה הייתה ממליצה לי לעשות לו רציתי ליישם בארץ דבר אחד מכל תחומי העבודה אליהם נחשפתי. אחרי שחשבה מעט על השאלה יעצה לי להקים צוות של אנשים שיוכלו לחשוב יחד איך לקדם את הנושא באופן שיתאים לעבודת החוקרים בארץ.

ועדת התמיכה במחקר של האוניברסיטאות עוסקת בנושאים שונים הנוגעים לצרכים של החוקרים. ועדה זו בוודאי מתאימה לקידום הנושא של ניהול נתוני מחקר. פעילות נוספת אותה ניתן ליישם גם ללא תשתיות או תמיכה כספית של המוסד היא העלאת מודעות החוקרים לניהול טוב יותר של נתוני המחקר שלהם ומתן המלצות ליישום של כמה כללים בסיסיים הנוגעים לשמירה, גיבוי ועבודה משותפת.

תודות

ההשתלמות באוניברסיטת אדינבורו מומנה על ידי הקרן לקידום מדעי הרוח והחברה מטעם האקדמיה הלאומית הישראלית למדעים. אני מודה לקרן על תמיכתה הנדיבה.

מקורות

ישראלי, ת' (2019). *שיתוף קבצים: שאלת הסקייבליות*. עבודה לשם קבלת תואר דוקטור לפילוסופיה, אוניברסיטת בר-אילן, רמת-גן.

פלך, נ' (2013). *שיתוף מסמכים בפרספקטיבת האחזור*. עבודה לשם קבלת תואר מוסמך, אוניברסיטת בר-אילן, רמת-גן.

Australian National Data Service. (August 2011). What is research data? Retrieved from: <https://www.ands.org.au/guides/what-is-research-data>.

Bergman, O., Whittaker, S. (2016). *The science of managing our digital stuff*. Cambridge MA: MIT Press.

Bryant, R., Lavoie, B., & Malpas, C. (2017a). *The realities of research data management: Part one: A tour of the Research Data Management (RDM) service space*. Dublin, OH: OCLC Research. doi: 10.25333/C3PG8J

Bryant, R., Lavoie, B., & Malpas, C. (2017b). *The realities of research data management: Part two: Scoping the university RDM service bundle*. Dublin, OH: OCLC Research. doi: 10.25333/C3Z039

Capra, R., Vardell, E., & Brennan, K. (2014). File synchronization and sharing: User practices and challenges. *Proceedings of the American Society for Information Science and Technology*, 51(1), 1-10.

Corrall, S (2012) Roles and responsibilities: Libraries, librarians and data. In: G. Pryor (Ed.), *Managing research data* (pp. 105–133). London: Facet.

Cox, A. M., Kennan, M. A., Lyon, L., & Pinfield, S. (2017). Developments in research data management in academic libraries: Towards an understanding of research data service maturity. *Journal of the Association for Information Science and Technology*, 68(9), 2182-2200.

Cox, A. M., & Pinfield, S. (2014). Research data management and libraries: Current activities and future priorities. *Journal of Librarianship and Information Science*, 46(4), 299-316.

Cox, A. M., Verbaan, E. & Sen B. (2012) Upskilling liaison librarians for research data management. *Ariadne 70*. Retrieved from: <http://www.ariadne.ac.uk/issue70/cox-et-al>

European Union Open Data Portal: <http://data.europa.eu/euodp/en/data/>

- European Commission. 2011. *Open data: An engine for innovation, growth and transparent governance*. COM/2011/0882 final. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF>
- Hoofnagle, C. J., van der Sloot, B., & Borgesius, F. Z. (2019). The European Union general data protection regulation: What it is and what it means. *Information & Communications Technology Law*, 28(1), 65-98.
- Ingram, C. (2016). *How and why you should manage your research data: A guide for researchers* [web page]. Updated 25 February 2019. Retrieved 23 August 2019 from <https://www.jisc.ac.uk/guides/how-and-why-you-should-manage-your-research-data>.
- Lewis, M. (2010). Libraries and the management of research data. In S. McKnight (Ed.), *Envisioning future academic library services: Initiatives, ideas and challenges* (pp. 145-168). London: Facet.
- Lyon, L. (2012). The informatics transform: Re-engineering libraries for the data decade. *International Journal of Digital Curation*, 7(1), 126-138. doi: 10.2218/ijdc.v7i1.220.
- Markauskaite, L., Kennan, M. A., Richardson, J., Aditomo, A., & Hellmers, L. (2012). Investigating eResearch: Collaboration practices and future challenges. In A. Juan, T. Daradoumis, M. Roca, S. Grasman, & J. Fauli (Eds.), *Collaborative and distributed e-research: Innovations in technologies, strategies and applications* (pp. 1-33). Hershey, PA: IGI. doi: 10.4018/978-1-4666-0125-3.ch001
- Pinfield, S., Cox, A. M., & Smith, J. (2014). Research data management and libraries: relationships, activities, drivers and influences. *PLoS One*, 9(12), e114734.
- Rader, E. (2009). Yours, mine and (not) ours: Social influences on group information repositories. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2095-2098). ACM.
- Rice, R., & Haywood, J. (2011). Research data management initiatives at University of Edinburgh. *International Journal of Digital Curation*, 4(2), 232-244.
- Sanchez, C. (2019). *Data mindfulness = Making the most of your dissertation*. Retrieved November 24, 2019, from <http://datablog.is.ed.ac.uk/files/2019/08/Data-Mindfulness-Making-the-Most-of-your-Dissertation-handbook-1.pdf>
- UK Research and Innovation. (2018). *UKRI open access principles and high level policy*. <http://www.rcuk.ac.uk/research/datapolicy/>.

Title: Collaborating with Research Data**Consent Form for Interviews: a Qualitative study**

Thank you for reading the information sheet about the interview sub-study. If you are happy to participate then please complete and sign the form below. Please initial the boxes below to confirm that you agree with each statement:

***Please
Initial
box:***

I confirm that I have read and understood the information sheet dated 16/09/2019 and have had the opportunity to ask questions.

I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without there being any negative consequences. In addition, should I not wish to answer any particular question or questions, I am free to decline.

Please choose one of the following two options:

I would like my **real name** used in the above

I would like a **pseudonym** to be used in the above.

I would like my comments to be attributed **anonymously**

I agree for this interview to be tape-recorded. I understand that the audio recording made of this interview will be used only for analysis and that extracts from the interview, from which I may or may not be personally identified (according to the choice above), may be used in any conference presentation, report or journal article developed as a result of the research. I understand that no other use will be made of the recording without my written permission, and that no one outside the research team will be allowed access to the original recording.

I agree that my anonymised data will be kept for future research purposes such as publications related to this study after the completion of the study.

I agree to take part in this interview.

Name of participant

Date

Signature

Dr Tamar Israeli_____
Principal Investigator

Date

Signature

Collaborating with Research Data: a study for the Research Data Support team

I would like to invite you to take part in a project study. Before you decide you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. Ask questions if anything you read is not clear or if you would like more information. Take time to decide whether or not to take part.

My name is Dr Tamar Israeli. I am an intern with the Research Data Service, and I am doing this study to help understanding researchers' data collaboration needs and practices, with a view to improving the service.

During the interview I will ask you to describe the ways in which you and your colleagues share files (particularly data files) throughout the research collaboration process: what makes you choose the tools you are using, any difficulties in data sharing and how you overcome them. I would also like to hear about the creation of a research data management plan (DMP), and especially about the issues that are relevant for collaboration. This interview should take between 30 minutes and one hour and will take place in a location of your convenience. The interview will be audio recorded for later analysis.

Since we are mainly interested in the collaboration side of the things, the criteria for choosing participants is that they are University-based researchers at any career stage who collaborate with other colleagues (internal and external) on a regular basis.

Participation is completely voluntary. You have the right to refuse participation, refuse any question and withdraw at any time without any consequence whatsoever.

Data can be anonymized, pseudonymized or attributed according to your wishes.

The information will be stored on an institutional shared drive. Only the project team will have access to the data, and it will be erased after the findings are published.

The final findings will be used for the research data support department for service evaluation. It is also possible that they will be presented as a poster or talk at an upcoming conference.

For further information please contact:

Investigator: Dr Tamar Israeli y1tisrae@exseed.ed.ac.uk

Western Galilee College (home institution) / University of Edinburgh (visitor)

Supervisors: Martin Donnelly martin.donnelly@ed.ac.uk

Research Data Support Manager, Library & University Collections, University of Edinburgh

Robin Rice r.rice@ed.ac.uk Data Librarian and Head, Research Data Support

Library & University Collections, University of Edinburgh

Thank you very much for your cooperation.