

חיים דיגיטליים // המחשב שלי/התיקיות שלי/החיים שלי

המידע הדיגיטלי הולך ומצטבר, הבלגן במחשב ובנייד חוגג, והגיע הזמן לסדר. ד"ר עופר ברגמן מהמחלקה ללימודי מידע באוניברסיטת בר-אילן, מסביר מאיפה להתחיל

11:52 13.07.2017 מאת: [ענן דינר](#)

לפני שנים לא רבות כמעט שלא היה דבר כזה – "מידע אישי דיגיטלי". מכתבים ומסמכים נשמרו בקלסרים ובמגירות; כתובות ומספרי טלפון נכתבו בספר מרופט שהונח בסמוך למכשיר הטלפון הביתי; ותמונות? אלה מהן שלא נחשבו טובות מספיק כדי להיכלל באלבומים הייצוגיים, אוחסנו בקופסאות נעליים. עידן המידע היה אמור לפטור אותנו מכל זה ולשים את כל המידע האישי במקום אחד – בקצות אצבעותינו. אלא שיחד עם הנגישות והזמינות הגיעו גם אתגרים חדשים: איך לארגן את המידע? איך לסדר אותו כך שניתן יהיה למצוא פריט מסוים בקלות במועד מאוחר יותר? איך לשמור על הפרדה בין עיקר וטפל בלי להקדיש זמן רב למיון ולמחיקה?

השאלות האלו מעסיקות חוקרים בתחום המחקר הצעיר של ניהול מידע אישי (Personal Information Management), ועומדות במרכזו של ספר חדש, The Science of Managing Our Digital Stuff ("המדע של ניהול הדברים הדיגיטליים שלנו"), שכתב ד"ר עופר ברגמן מהמחלקה ללימודי מידע באוניברסיטת בר-אילן, ופרופ' סטיב וויטאקר מאוניברסיטת סנטה קרוז בקליפורניה. הספר, שראה אור לאחרונה בהוצאת MIT Press, מסכם שני עשורים של מחקר משותף על הרגלי הניהול של מידע דיגיטלי אישי, והוא מציג גישה חדשה לפיתוח מאפיינים של מערכות הפעלה, המבוססת על הידע המצטבר ומתמקדת בצורכי המשתמשים.

הספר כולל גם לא מעט ביקורת כלפי חברות התוכנה הגדולות שמפתחות את מערכות ההפעלה שמשמשות את כולנו מדי יום. לדברי ברגמן וויטאקר, המערכות האלה שבזכותן תפישתה שגויה הרואה במשתמש צרכן של מידע, בזמן שהרבה מהשימוש שאנשים עושים במערכות ההפעלה שלהם הוא דווקא אצירה (Curation) של אוספי ענק של מידע

אישי (כגון קבצים ודוא"ל). "ראייה זו מתעלמת מכמה בעיות דוחקות הנוגעות לניהול מידע", הם כותבים, והתוצאה היא שלעתים קרובות משתמשים רבים טובעים במידע לא נחוץ שמצטבר במחשביהם, ומתקשים למצוא מידע חשוב. "אחד הדברים הבולטים שגילינו במחקרים שלנו הוא שנלווים לעניין הזה הרבה רגשי אשמה", אומר ד"ר ברגמן למגזין TheMarker. "רבים מהאנשים שפניתי אליהם בהצעה להשתתף במחקר הזהירו אותי שאצלם 'יש הרבה בלגן' במידע הדיגיטלי, או טענו שהם 'לא דוגמה', כי הם לא יודעים לנהל את המידע היטב".

מהי הבעיה העיקרית עם ניהול מידע אישי?

"חוקרים רבים סבורים שהבעיה נעוצה בשיטת התיקיות. זו השיטה הכי נפוצה ולכן קל והגיזני להאשים אותה. זו גם שיטה ותיקה מאוד שראשיתה בשנות ה-60, עם המחיצות (Directories), שהיו נהוגות במערכות הפעלה טקסטואליות כמו DOS וקודמותיה. כשהגיעו מערכות ההפעלה הוויזואליות כמו ווינדוס, בשנות ה-80, הצטרף גם הייצוג הוויזואלי של התיקיות, ואז יכולת לבצע עוד פעולות כמו לשים קובץ בתוך תיקייה, לגרור קבצים מתיקייה לתיקייה וכדומה. איש מחשבים יגיד שזה לא באמת בתוך תיקייה, זו רק מטאפורה, אבל אנחנו מרגישים נוח בתוך המטאפורה הזאת. למרות זאת, חוקרים רבים, לצד אנשים בתעשיית התוכנה וגם משתמשים, מרגישים שמהו במטאפורה של התיקיות לא עובד כשורה, ושצריך להחליף אותן במשהו חדש".

מה הבעיה עם התיקיות?

"קודם כל, תיקיות מחביאות מידע. אם אני שם משהו על הדסקטופ, או על השולחן הפיזי, אני רואה אותו והוא מזכיר לי את עצמו. אם אניח אותו כאן ביום חמישי, ביום ראשון כשאגיע לעבודה אני אזכר בו. לא כך בתיקיות. פריט מידע מסוים שנמצא בתוך תיקייה, לא גלוי לעיני המשתמש אלא אם כך הוא זוכר בדיוק היכן שמר אותו. יש גם בעיות של סיווג – הדברים בעולם האמיתי לא מתחלקים כל כך טוב לשמות של תיקיות. וגם כשאתה מוצא לתיקיה שם, השימוש של התיקייה יכול להשתנות לאורך זמן והשם כבר לא יהיה רלוונטי. והדבר הכי מהותי וחשוב הוא שפשוט קשה לזכור. אם שמת משהו בתיקייה לפני חודשיים או לפני שנתיים, ואתה רוצה לאחזר אותו עכשיו, אם לא תזכור בדיוק את הדרך לאותו קובץ, אתה לא תגיע אליו".

מה האלטרנטיבה שהתעשייה מציעה?

"התעשייה מקדמת שני פתרונות: חיפוש ותגיות. חיפוש הוא הרבה יותר גמיש מתיקיות. זורקים הכל בבלגן אחד גדול, ומנוע חיפוש מביא את המידע ללא צורך לזכור את המיקום שלו. חוקרים רבים האמינו שהחיפוש יחליף את התיקיות כי זה בעצם מה שקרה ברשת. בתחילת שנות ה-90 האתרים הגדולים היו בנויים כאינדקסים שהיו מחולקים על פי נושאים ותתי-נושאים. במקביל התפתחו גם מנועי החיפוש, וכיום אני לא צריך לומר לך מי מבין שתי השיטות האלה ניצחה. אז ההנחה היא שגם במידע האישי יתרחש אותו תהליך, ומנועי החיפוש יחליפו את התיקיות".

בספר אתם טוענים שזו הנחה שגויה, ושאנשים מעדיפים להשתמש בתיקיות למרות הבעיות.

"נכון. במחקר שנערך לפני כשנה ובדק רישומים (לוגים) של שימוש, השתתפו צעירים שהשתמשו בגרסה החדשה ביותר של המק. נמצא ששיעור החיפושים היה 4.5%. כלומר, מתוך כלל הפעמים שבהן המשתמשים היו צריכים להגיע לקבצים, רק ב-4.5% מהפעמים הם השתמשו בחיפוש. גם מחקרים אחרים, ביניהם כאלו שערכנו, הראו זאת. כששאלנו אנשים שהשתתפו בניסוי מתי בכל זאת הם משתמשים בחיפוש, הם אמרו, 'רק כשאנחנו לא זוכרים בדיוק איפה שמנו את הפריט שאנחנו רוצים לאחזר'. כלומר, כאשר יש לאנשים ברירה, הם יעדיפו לנווט כמעט ב-100% מהמקרים. המחקר שלנו הראה שהשיפור הגדול שהיה במנועי חיפוש לפני כעשור לא הביא ליותר חיפושים. אפשר וצריך לשפר עוד ועוד את מנועי החיפוש, אבל זה לא מה שיפתור את הבעיות הגדולות של ניהול מידע אישי".

מה לגבי תגיות?

"תגיות באות לפתור בעיה אחרת שיש עם תיקיות, והיא שיש פרטי מידע שיכולים להיות קשורים לכמה נושאים. למשל, דוקטורנט שלי שולח לי מצגת שמיועדת להצגה בכנס. תחת איזו תיקייה אני אשמור אותה? תחת שמו של הדוקטורנט? תחת שם הכנס? תחת הנושא של הכנס? יכולות להיות כמה אפשרויות ואם אבחר באחת, זה יוציא את האפשרויות האחרות. ואם בעוד שנה ארצה למצוא את הקובץ הזה, לא בטוח שאצליח להתחקות אחרי קו המחשבה שלי עכשיו, שמוביל אותי לשים את הקובץ הזה כאן ולא שם.

"תגיות מאפשרות סיווג מרובה – אפשר לשים כמה תגיות על אותו פריט מידע – וגם כאן יש הסכמה מקיר לקיר לגבי זה שהתגיות הן הפתרון הבא, בין השאר בעקבות ההצלחה המרשימה של השימוש בתגיות ב-Web 2.0. בתעשייה מאמינים שזה העתיד. אנחנו חקרנו את זה, בין השאר אצל משתמשי ג'ימייל וגם בווינדוס 7, ושוב התברר שזה ממש לא המצב. מצאנו העדפה חזקה לשימוש בתיקיות על פני שימוש בתגיות. כשביקשנו ממשתמשי ווינדוס 7 להשתמש בתגיות, הם עשו זאת במהלך הניסוי, אך כאשר חזרנו אליהם אחרי שהניסוי הסתיים, הם הפסיקו להשתמש בהן כמעט לחלוטין. "משתתפי הניסוי בג'ימייל השתמשו בתגיות בכשליש מהאחזורים שלהם, אולם ב-92% מהמקרים הם השתמשו רק בתגית אחת. כלומר, הם בכלל לא ניצלו את האפשרות של הסיווג המרובה, ובמקרים המועטים שהשתמשו בה, הם לא ניצלו זאת בזמן אחזור המידע. תגיות הן רעיון טוב שמתאים לזמן 2, אבל לא כל כך מתאים לניהול מידע אישי".

אם הבעיה העיקרית בניהול מידע אישי אינה התיקיות. מהי הבעיה וכיצד אתם

מציעים לפתור אותה?

"הבעיה היא שבעיצוב מערכות לניהול מידע אישי לא מביאים בחשבון את המאפיינים הייחודיים שלהן. יש הבדל עמוק בין מערכות לניהול מידע אישי לבין מערכות להצגת מידע

ציבורי. במערכות ציבוריות, כגון אתרי אינטרנט, יש מישהו שעורך את המידע עבור האנשים הרבים שצורכים אותו. זו מעין תקשורת בינו לבינם, שבה הוא חייב להשתמש במאפיינים אובייקטיביים של המידע, כדי לאפשר למגוון רחב מאוד של אנשים להשתמש במידע הזה. ניהול מידע אישי הוא בעל מאפיינים אחרים. כאן אותו אדם ששומר את המידע הוא גם זה שמאחזר אותו מאוחר יותר, ולכן זוהי תקשורת בין האדם לבין עצמו – בשני זמנים שונים. המערכת צריכה לאפשר לאדם לגשר על פער הזמנים הזה, וכדי לעשות זאת היא יכולה להשתמש לא רק במאפיינים האובייקטיביים של המידע אלא גם במאפיינים סובייקטיביים, ייחודיים לאותו משתמש. זאת הגישה הסובייקטיבית-למשתמש, שאת העקרונות והיישומים שלהן אנחנו מפרטים בספר.

"יש שלושה עקרונות בסיסיים, והראשון שבהם קשור לנושא החשיבות. חשיבות היא כמובן דבר סובייקטיבי ותלוי משתמש – מה שחשוב לי עשוי להיות חסר ערך עבורך, ולהפך. יש כל מיני דרכים שבהן משתמש יכול לסמן פריטי מידע מסוימים כחשובים. הוא יכול למשל לשים אותם על הדסקטופ שלו. אבל מה קורה במצב הפוך, כשיש פריטים שאינם כה חשובים? במקרה כזה ישנן שתי אפשרויות גרועות. תוכל למחוק את אותו פריט, ואז אולי תצטער על כך מאוחר יותר, אם תזדקק לו. לעומת זאת אם לא תזרוק שום דבר, כל הפריטים הפחות חשובים יתחרו על הקשב שלך. מחקרים רבים הראו שזו הבחירה של מרבית האנשים. הם ממעטים למחוק, ונאלצים להתמודד עם העומס. "הגישה הסובייקטיבית-למשתמש מציעה כפתרון את **עיקרון ההצנעה**. המערכת צריכה לאפשר למשתמש להצניע פריטי מידע ולהפוך אותם לפחות בולטים, אבל להשאיר אותם בתוך ההקשר המקורי שלהם. ככה הם לא יתחרו על הקשב, אבל יהיו זמינים בעת הצורך".

איך תיראה מערכת כזאת?

"התיאור שלנו הוא מופשט בכוונה, כדי לאפשר למעצבים את החופש להשתמש בעיקרון הזה בדרכים שונות, אבל פיתחנו כמה יישומים קונקרטיים שאותם בדקנו בניסויים אמפיריים וקיבלנו תוצאות חיוביות ביותר. למשל, 'אזור אפור' בתוך החלון של תיקיית קבצים, שאליו אפשר לגרור את כל הקבצים שרוצים להצניע. "יישמנו את עיקרון ההצנעה גם ברשימות אנשי הקשר בטלפונים ניידים. יש המון פריטי קשר שאינם בשימוש – מספרי טלפון של אנשים שמצטברים עם הזמן, ואף פעם לא נמחקים. ראינו במחקרים שלנו ששיעור הפריטים הללו הוא בערך מחצית מכל פריטי הקשר, והם מתחרים גם על המקום המוגבל בתצוגה של המכשיר. מערכת הפעלה שתיישם את עיקרון ההצנעה תוכל להצניע את כל הפריטים שלא השתמשת בהם, נניח, יותר מחצי שנה, על ידי הצגתם אחרי יתר הפריטים. תוכל עדיין להגיע אליהם, אבל הם לא יפריעו לך בשימוש השוטף ברשימת אנשי הקשר.

"יישום נוסף הוא באוספי תמונות. כולנו מצלמים המון תמונות דיגיטליות ושומרים הכל, ואז יש בלגן בתיקיות של התמונות והרבה תמונות דומות זו לזו. היישום מציע שהמערכת

תזהה תמונות דומות ותציג רק אחת מהן, כשהאחרות מסודרות מאחוריה, כמו חבילה של קלפים".

מהם שני העקרונות הנוספים של הגישה הסובייקטיבית-למשתמש?

"העיקרון השני נקרא 'עיקרון הפרויקט', והוא אומר שהממשק צריך לאפשר למשתמש לשמור ביחד פריטי מידע ששייכים לאותה קטגוריה, ללא קשר לפורמט שלהם. זה נשמע טריוויאלי, אבל זה רחוק ממה שקורה עכשיו. כיום, פריטים שקשורים לאותו פרויקט יכולים להיות מפוזרים על פני כמה מקומות – תיבת המייל, ספריית המסמכים, הסימניות בדפדפן וכן הלאה. עבורי כמשתמש, כל פריטי המידע הללו קשורים זה לזה, אבל מבחינת המחשב אין כל קשר בין הפריטים הללו, ולכן כשאתה עובד על פרויקט מסוים אתה נדרש לדלג בין כמה מקומות במחשב.

"עיקרון הפרויקט מציע שינוי די רדיקלי באופן שבו מערכת ההפעלה מציגה את פריטי המידע האלו. ניתן יהיה לשים בתיקות פרויקטים פריטים מסוגים שונים – לא רק קבצים, אלא גם אימיילים, סימניות אנשי קשר או כל פריט אחר, וליצור עץ תיקיות אחד שבנוי על פי ההקשרים הסובייקטיביים שרלוונטיים עבורי כמשתמש. לצורך נוחות התצוגה, אפשר להפריד בין פריטים מסוגים שונים שנמצאים באותה תיקייה באמצעות לשוניות (טאבים), באופן דומה לשימוש בלשוניות בדפדפן".

אפשר להניח שיש משתמשים מתוחכמים שכבר נוהגים כך כיום, ושמים בתיקות שלהם אימיילים רלוונטיים, סימניות וכיוב'.

"בדקתי את זה בדוקטורט שלי, וזה נכון חלקית. אנשים באמת שומרים במקום אחד קבצים מסוגים שונים, בתיקות משנה על פי פרויקטים, והם עושים זאת כי העיצוב של מערכת ההפעלה מאפשר זאת. אולם הם לא נוהגים כך היכן שהעיצוב אינו מעודד זאת: אפשר למשל לשמור אימיילים בתיקיית המסמכים, אבל לא ניתן להשיב עליהם. אפשר גם לשמור סימניות בתיקיית מסמכים, אבל המחקר שלי הראה שרוב המשתמשים אינם יודעים שזה אפשרי ולא משתמשים באפשרות הזאת.

"העיקרון השלישי הוא עיקרון ההקשר הסובייקטיבי, והוא מתייחס למצב שבו משתמש צריך לאחזר פריט מידע כלשהו שהוא שמר בעבר. אחרי שהמשתמש מסיים לעבוד עם פריט מידע מסוים ושומר אותו, ההקשר שבו אותו פריט מידע נשמר עשוי לעזור למשתמש לאחזר אותו בעתיד. לדוגמה, תחשוב על סטודנט שמקשיב להרצאה וכותב במכשיר הנייד שלו נקודות מתוכה. תוך כדי שהוא כותב, המכשיר מקליט את מה שאומר המרצה, מתוך הנחה שמה שחשוב לסטודנט הוא כותב, ומה שלא כל כך חשוב, לא נכתב. אחרי זמן מה, אותו סטודנט לומד לבחינה ונעזר ברשימות שלו. לעתים קרובות רשימות כאלה הן לא כל כך מסודרות, וייתכן שיהיה קשה לשחזר את מה שנאמר בזמן שהן נכתבו. אבל יישום שיתבסס על עיקרון ההקשר הסובייקטיבי יוכל למשל לאפשר לו לשמוע מה שאמר המרצה בכל מקום שבו הרשימות אינן ברורות – הוא יוכל פשוט להקיש על משפט או קטע ברשימות, ולשמוע מה המכשיר שלו הקליט בזמן שכתב אותו".

לרעיונות שאתם מפרטים בספר יש פוטנציאל כלכלי לא מבוטל. למה בעצם לעסוק

בזה באקדמיה ולא בתעשייה?

"זו שאלה שהתמודדתי איתה לא מעט בראשית הדרך. בתחילת העשור הקודם, כשאמרתי למנחה שלי בדוקטורט שאני רוצה לחקור ניהול מידע אישי, הוא לא היה בטוח שיש בכלל תחום מחקר כזה. הוא אמר לי 'תביא לי שלושה מאמרים בנושא, ואז אני אדע שהתחום הזה קיים'. הצלחתי למצוא שלושה מאמרים כאלה, אבל אם הוא היה מבקש חמישה או עשרה, לא בטוח שהייתי מצליח ואולי לא היינו מדברים עכשיו. מאז המצב מאוד השתנה וכיום יש קהילה מדעית של יותר מ-100 חוקרים שעוסקים בזה ומתפרסמים עשרות מאמרים בשנה. מאמרים אלו מגדילים את הידע בתחום, זהו ידע שהתעשייה יכולה וצריכה להשתמש בו בעיצוב מערכות הפעלה חדשות. "לגבי יישום מסחרי, הגישה שלנו היא שאלה קודם כל רעיונות גנרטיביים, ואנחנו מקווים שאנשים אחרים ייקחו אותם הלאה ויישמו אותם. אנחנו קודם כל מחויבים לקדם את הידע, ולכן אנחנו מפרסמים מחקר מדעי ובכוונה לא רושמים פטנטים על יישומים שאנחנו מפתחים. אנחנו לא רוצים להגביל את היישום למשתמשים של חברה אחת שתרכוש את הפטנט, אלא לפתוח אותו לכל החברות. מעבר ליישומים הספציפיים, אנחנו מעדיפים שכל חברה תיקח את העקרונות שלנו לכיוון אחר. מה שהכי הייתי רוצה הוא שאנשים בתעשייה ייקחו את העקרונות האלה ויחשבו על עיצובים משל עצמם, ואני מקווה שלספר יהיה אפקט כזה".