

פרטיות – מזכות יסוד למושג נלמד

ארז ויסברד*

המקוונים שלנו. גם פושעים מסורתיים יותר נהנים מהמידע שאנו חולקים ברשת. פרסום המידע שלפיו אנו נמצאים כעת באתר נופש אקזוטי מסייע מאוד בידי פושעים. כעת הם יודעים שהבית שלנו פנוי. האתר pleaserobme.com מציג את המידע שאנשים חולקים בטוויטר בנוגע למיקומם כדי להעלות את המודעות לסכנה.

אחת הבעיות המובילות לשיתוף יתר היא השימוש במושג "חבר" ברשתות חברתיות. ייתכן שיש מקום לתת אמון במעגל החברים הקרוב ביותר, ואולי אף בחבריהם, אך אין מקום לעשות כן ברשתות החברתיות. הדבר נכון בפרט לאנשים שמספר ה"חברים" שלהם ברשת החברתית גבוה מאלף. גם אם מבחינה טכנית אפשר להגביל את השיתוף ברשתות חברתיות לקבוצות מצומצמות, רובנו משתפים אוטומטית עם כל ה"חברים".

פרטיות בראי הזמן והזכות להישכח

כולם מחפשים את המידע שאנו חולקים. לפני ריאיון, מעסיק פוטנציאלי יחפש על המרואיין מידע גם ברשתות החברתיות, והוא עלול לראות בעין לא יפה תמונות או אמירות מסוימות. תמונות ומחשבות ששיתפנו לפני שנים, לעתים ברגע של חוסר מחשבה, עלולים לצוץ שנים אחרי כן ולהשפיע על חיינו. יש לזכור גם שהנורמות החברתיות משתנות - מה שהיה מקובל לפני עשר או עשרים שנה לא בהכרח מקובל היום. אך לא רק המעסיקים מחפשים אותנו. לפני כל דייט מחפשים שני בני הזוג מידע באינטרנט בכלל וברשתות החברתיות בפרט. היום אנו עדים ליוזמות רגולציה, בייחוד באיחוד האירופי, שנועדו לעגן את הזכות להישכח - את הזכות שהעבר לא ירדוף אותנו לנצח (Mantelero, 2013). כפועל יוצא של הזכות הזאת אנשים פרטיים יכולים לבקש ממנועי חיפוש להסיר תוצאות חיפוש לא מחמיאות (גם אם אמיתיות) מהעבר הרחוק. זכות זו אינה עומדת לפוליטיקאים ולדמויות ציבוריות אחרות כיוון שזכות הציבור לדעת על עברם עולה על זכותם לפרטיות.

ניתוח מתקדם של מידע

אחד הדברים שהטכנולוגיה הביאה בשנים האחרונות הוא היכולת להסיק תובנות מתוך מידע רב שנאסף. התחום הצומח של נתוני עתק (Big data) נותן תובנות רבות, בכל תחומי החיים, על ידי זיהוי של קשרים בין נתונים לאירועים. בתחום הרפואי יודעים היום לזהות קשרים בין גנים מסוימים לסיכוי לחלות במחלת הסרטן (גם אם לא יודעים להסביר את הקשר). דוגמה אחרת: חברת גוגל הצליחה לזהות התפרצות של שפעת מניתוח שאילתות החיפוש של משתמשים מאזור גאוגרפי אחד עוד לפני שהמידע היה זמין למרכז הארצי למניעת מחלות.

בצד היישומים הללו, ניתוח נתוני עתק מאפשר לזהות הרגלי קנייה ולמקד פרסומות בקהל היעד. כאשר אדם רוכש ספר מסוים באמזון, מיד מוצעים לו ספרים אחרים לצד הכיתוב "לקוחות שרכשו את הספר שבחרת הביעו עניין גם בספר הבא". ההיגיון שמאחורי ההצעה פשוט - לאנשים שקנו את אותו ספר טעם דומה, ולכן יש סיכוי טוב שאותו אדם יגלה עניין גם בספרים האחרים שהם קנו. כמובן שקניה של ספר אחד אינה מספיקה כדי לקבוע ששני אנשים דומים. אלגוריתמים מתוחכמים מבססים את ההמלצות על מספר רב ככל האפשר של תבחינים.

חברת נטפליקס (Netflix), המשכירה סרטים, רצתה לשפר את יכולתה לחזות את טעמם של המשתמשים כדי שתוכל לקלוע לטעמם כאשר היא מציעה להם לשכור סרטים. זהו אתגר לא פשוט משום שאין לאדם טעם קבוע אחד - אותו אדם ירצה לראות סרט אחד כאשר מצב רוחו טוב וסרט אחר בשבוע המעבודה אחרי יום לא מוצלח. בשנת 2009 הכריזו בחברה על תחרות למציאת

תקציר: בעולם של ימינו, המתאפיין בשיתוף מידע, המושג פרטיות הולך ומתפוגג. אחת הסיבות לכך היא שקשה להבחין במידת הנזק שגורם פריט מידע יחיד שאדם חולק. חינוך לפרטיות ומודעות בשילוב עם טכנולוגיות משמרות פרטיות יאפשרו גם לדור הבא לשמור על זכות יסוד זו.

השינויים הטכנולוגיים המתרחשים בשנים האחרונות מאפשרים לנו לתקשר באופן שלא הכרנו בעבר. המארג החברתי שלנו אינו מוגבל עוד לסביבה הפיזית. אנו יכולים לשתף בזמן אמת חוויות עם קרובי משפחה הגרים ביבשת אחרת וכן להיות חלק מקהילה וירטואלית עולמית של אנשים החולקים תחום עניין משותף. אולם אותה תקשורת הופכת את המידע שלנו לנגיש גם לגורמים אחרים, שהוא לא היה נגיש להם בעבר. למשל, בעבר היה צורך לעקוב אחר אדם באופן פיזי כדי לדעת היכן הוא נמצא. היום די להביט במידע שאותו אדם חולק ברשתות החברתיות כדי לדעת היכן היה, עם מי היה ומה עשה.

האם יש בכך בעיה? יש שיטענו שהשליטה על מידת החשיפה נמצאת בידינו ושכל אדם בוגר יכול לבחור איזה מידע לחלוק ועם מי לחלוק אותו. במאמר זה נראה מדוע טענה זו אינה נכונה. נראה שרובנו איננו מודעים להיקף המידע שאנו חולקים ולהשלכותיו על פרטיותנו. לכן עלינו ללמוד מחדש את המושג פרטיות ואת השלכות העידן הדיגיטלי עליו, ולהקנות אותו לדור אשר נולד לתוך עולם שהמושג הזה כמעט זר בו.

בעיית השקיפות

הטכנולוגיה המשתכללת והפיכת הממשקים לפשוטים יותר ויותר מקשים מאוד על המשתמשים לדעת כיצד בדיוק הטכנולוגיה עובדת ואיזה מידע היא חולקת. כולנו מכירים את חלון ההרשאות שעולה בזמן התקנה של אפליקציה על הטלפון שלנו. רובנו מאשרים מיד את ההרשאות, אבל אם נעצור ונביט לרגע ברשימת ההרשאות המתבקשות נראה לא פעם שהאפליקציה מבקשת גישה גם למידע פרטי שלא נדרש כלל לצורך פעולתה. ייתכן שגם אתם בין המיליונים אשר התקינו אפליקציית פנס פופולרית שביקשה הרשאות ליומן, לרשימת אנשי הקשר ולמיקום (אם אינכם בטוחים אילו הרשאות התבקשתם לתת בזמן ההתקנה, תוכלו לבדוק זאת בפרטי היישומים המותקנים). את המידע הזה המפתחים אוספים ומוכרים לגורמים בעלי עניין בלא שנשים לב לכך. עקב ממדי התופעה, ה-FTC (ארגון המסחר הפדרלי בארצות הברית) החל לתבוע מפתחי אפליקציות שפגעו בפרטיות המשתמשים וחרגו ממה שהצהירו (FTC).

גם כאשר אנו יודעים מה אנו משתפים, לא תמיד אנו יודעים עד כמה אנו נחשפים ומה השלכות החשיפה. דוגמה משעשעת לכך סיפר לי מנכ"ל של חברה שאחד מעובדיו ביקש ממנו לצרף אותו לרשימת חבריו בפייסבוק. העובד לא חשב על כך שכאשר ישבור שיא במשחק רשת בשעות העבודה, גם חברו החדש יחשף לתחביביו.

עד כה עסקנו במודעות להיקף המידע המשותף. כעת נעבור לדון בהשלכות האפשריות של המידע שאנו חולקים. למעשה, קשה מאוד להבין עד כמה המידע שאנו משתפים יכול להיות רגיש. גם מידע תמים לכאורה, כגון שם בית הספר שלמדנו בו, מקום המגורים שלנו ואפילו אילן היוחסין שלנו נראים כולם פרטים תמימים, שממילא ידועים לסובבים אותנו. אך יש לזכור ששם בית הספר, שם האם ושמו של בעל החיים הם בדיוק הפרטים המשמשים בשאלות האבטחה שנשאלו כשפתחנו חשבון דוא"ל. המידע התמים הופך לכלי בידי תוקפים המעוניינים לקבל גישה לחשבונות

* ד"ר ארז ויסברד - חוקר במעבדות בל (Bell Labs) ומרצה באוניברסיטה הפתוחה.

לרגולציה המתפתחת בתחום (Commission).

דוגמה לטכנולוגיה כזאת אפשר לראות במאמר המתאר שיטה להתאמת פרסומות לצופי טלוויזיה בלא לפגוע בפרטיותם (Tzachy, 2012). לכאורה שתי הדרישות סותרות זו את זו - ביסוס הפרסומות על הרגלי צפייה אישיים ואף שיתוף המפרסמים בנתונים בנוגע לחשיפת הצופים לפרסומות, לצד שמירה על פרטיות. אולם החוקרים הראו שאפשר להתאים את הפרסומות המוצגת בממיר הטלוויזיה של הצופה בלא שהשדר ידע איזו התאמה בוצעה, ובכל זאת לדווח על נתוני הצפייה בפרסומות. הפתרון שהם מצאו מתבסס על שני עקרונות עיקריים. האחד, ביצוע ההתאמה בצד המשתמש, כלומר התאמת הפרסומות נעשית בממיר של המשתמש ולא בשרתים של השדר (וכך יורד מכתפי השדר עול הגנת המידע, כולל לפני עובדים). העיקרון השני הוא דיווח מידע על צריכת הפרסומות (לשם קבלת מידע על הטעם של המשתמש) באמצעות הרעשה סטטיסטית. כלומר, כל משתמש מדווח נתוני צפייה שאינם מדויקים באופן ששך כל הדיווחים ממליוני הצופים המגיעים למפרסם נתונים קירוב טוב לנתוני האמת.

לסיכום, הפרטיות, אשר עד לא מזמן הייתה זכות בסיסית, הולכת ונעלמת. הטכנולוגיה מאפשרת לשתף מידע ולנתח בהיקפים שלא הכרנו, ופרטיותנו צפויה להישחק עוד בעתיד. דמיינו מה יקרה כשהמצלמות הרבות הפזרות סביבנו יחוברו לתוכנה לזיהוי פנים, וזו תחובר לאינטרנט ולרשתות החברתיות. כל אדם מולנו וכל בית עסק שניכנס אליו ידעו מיד מי אנחנו ומה העדפותינו. בשל השינויים המהירים אנו נדרשים ללמוד מחדש את מושג הפרטיות, ובעיקר ללמד עליו את הדור שנולד לעידן שיתוף המידע. באמצעות חינוך, התערבות רגולטורית ושימוש בטכנולוגיות משמרות פרטיות, אולי נצליח גם בעתיד לשמר מעט מהזכות הבסיסית הזאת.

מקורות

- AOL search data leak (n.d.). In Wikipedia. Retrieved September 19, 2016, from https://en.wikipedia.org/wiki/AOL_search_data_leak
- Commission, E. (n.d.). General data protection regulation. Retrieved September 19, 2016 from <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679>
- FTC. (n.d.). FTC approves final order settling charges against flashlight app creator. Retrieved September 19, 2016 from <https://www.ftc.gov/news-events/press-releases/2014/04/ftc-approves-final-order-settling-charges-against-flashlight-app>
- Hill, K. (2012, February 16). How Target figured out a teen girl was pregnant before her father did. Forbes. Retrieved from <http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#b5a-87de34c62>
- Mantelero, A. (2013). The EU proposal for a general data protection regulation and the roots of the "right to be forgotten". Computer Law & Security Review, 29(3), 229–235.
- Netflix. (n.d.). Netflix prize. Retrieved September 19, 2016 from <http://www.netflixprize.com/>
- Reinman, T., & Waisbard, E. (2012). PPI-TTA – Preserving Privacy in TV Targeted Advertising. Secrypt, 327–332. doi:10.5220/0004076103270332

אלגוריתם החיזוי הטוב ביותר (ראו Netflix). לשם כך הפיצו בסיס נתונים גדול אשר הכיל את הדירוג של מיליוני דירוגים לאלפי סרטים. כמובן, לפני שפורסם המידע, החברה הסירה את הפרטים המזהים של המשתמשים והחליפה אותם במספרים. כמו כן, מקצת הדירוגים לא פורסמו אלא שימשו קבוצת ביקורת. האלגוריתם שחזה את הדירוגים שלא פורסמו באופן הטוב ביותר (ולמעשה שיפר את יכולת החיזוי של נטפליקס ביותר מעשרה אחוזים) זכה במיליון דולר.

לכאורה הכול היה מושלם, אלא שבמהרה התברר שהחלפת שמות המשתמשים במספרים לא הספיקה כדי לשמור על פרטיותם. באתר האינטרנט IMDB משתמשים מזהים מדרגים סרטים. בהצלבה בין שני מאגרי המידע נמצאה התאמה בכמה מן הסרטים, וההתאמה אפשרה לזהות את מקצת המדרגים במאגר האנונימי. כך, משתמש שדירג סרט מסוים באתר IMDB והזהדה, מצא שגם סרטים אחרים שהוא צפה בהם, ושבנוגע להם רצה להישאר אנונימי, זוהו עמו כעת.

לא רק נטפליקס כשלה בהגנה על פרטיות משתמשיה כשהסתפקה בהסרת הפרטים המזהים. בשנת 2006 הפכו שאילתות החיפוש של לקוחות ספקית האינטרנט AOL לנחלת הכלל. החברה שחררה את שאילתות החיפוש שהקישו יותר מחצי מיליון מלקוחותיה במשך שלושה חודשים, וזאת לצורכי מחקר. המידע המזהה האישי הוסר, ושמות המשתמשים הוחלפו במזהה אקראי. אף על פי כן, תחומי העניין שנחשפו בחיפושם אפשרו לחשוף מידע רב על המשתמשים - גיל, אזור גאוגרפי, תחומי עניין, מצב רפואי ועוד - וכך אפשר היה לזהות את מקצת המשתמשים. בעקבות החשיפה מיהרה חברת AOL להסיר את המידע מהאתר שלה, אך, כצפוי בעידן האינטרנט, עותקים של המידע עדיין קיימים ברשת ("AOL search data leak", n.d.).

עד כמה רב כוחם של אלגוריתמים לניתוח נתוני עתק ועד כמה יכולתנו להתחמק מהם קטנה אפשר ללמוד מסיפור שקרה לנערה מתבגרת. הנערה קיבלה לביתה, מהחנות טרגט (Target), קופונים לנשים בהיריון. אביה של הנערה הגיע זועם לחנות ודרש מהמנהל התנצלות על שליחת הקופונים לנערה תמימה. אולם כמה ימים לאחר מכן התקשר האב להתנצל לפני המנהל. הוא גילה שביתו המתבגרת אכן בהיריון. כיצד ידעו המוכרים בחנות שנערה מסוימת בהיריון לפני הוריה? מתברר ששינויים מסוימים בהרגלי הקנייה, דוגמת רכישת קרמים מסוימים וסוגים מסוימים של תוספי מזון, הם סימנים טובים עוד בשלב מוקדם בהיריון. מתאם בין קניית תוספי מזון וקרמים בחודש מסוים להתחלה של רכישת חיתולים חצי שנה מאוחר מבוסס על אלגוריתמים המנתחים ללא הרף מאגר נתונים גדול בחיפוש אחר קשרים נסתרים בין הנתונים. לעתים, אותם קשרים אינם גלויים, אלא מתבססים על תבחינים רבים בו בזמן. כמובן, ככל שכמות המידע שנאסף גדלה, אפשר לזהות קשרים רבים יותר (Hill, 2012, February 16).

הדוגמאות הללו מראות שהמידע שאוספים גופים גדולים עלינו גדל, ואילו יכולתנו להגן על פרטיותנו פוחתת. האנשים שבחרו להזדהות בשםם כשדירגו סרטים ב-IMDB יכלו אולי להגן יותר על פרטיותם אילו ויתרו על ההכרה הציבורית בתרומתם, אך אותה נערה לא יכלה להגן על פרטיותה מחברת טרגט. לשם כך נדרשת התערבות רגולטורית. בדיוק כמו הזכות להישכח, נדרשת אפשרות לבקש להיות מוחרגים מניתוח התנהגותי. יתרה מכך, בעבור אוכלוסיות מסוימות, דוגמת קטינים, רצוי שההחרגה תהיה אוטומטית.

טכנולוגיה משמרת פרטיות

במבט ראשון נוצר רושם שהטכנולוגיה הולכת בכיוון אחד בלבד - הכחדת הפרטיות. אולם אנו עדים למספר הולך וגדל של מחקרים ופיתוחים טכנולוגיים שעניינם סיפוק שירותים מבוססי מידע אישי בלא לפגוע בפרטיות המשתמשים. שיטות אלו מצמצמות את המידע שספק השירות אוסף, וכך גם מקטינות את עול אבטחת המידע הרובץ על ספק השירות ומתאימות את איסוף המידע