



מדור מיוחד ביג דאטה בתכנון העירוני

הקדמה:

נתונים גדולים עבור אתגרים עירוניים גדולים

דניאל פלזנשטיין

הופעתו של הביג דאטה* מחולל מהפכה שקטה בכל הקשור לניתוח תהליכים עירוניים ולניהול המערכת העירונית. גם אם נשלול את כלל הסופרלטיביים המתבטאים סביב 'העיר הדיגיטלית' ו'העיר הימקושרת' (wired city), לא ניתן להתכחש לעובדה שביג דאטה מציע הזדמנויות רבות להבנה חדשה של הצמיחה הדינמית של ערים. בו בזמן, ערי העולם ניצבות בפני משברים כלכליים, חברתיים וסביבתיים המאיימים על המשך פיתוח וצמיחתן. בשלב זה, לא ברור עד כמה ביג דאטה מסוגל לתרום לתיאוריה ולשיטות ניתוח חדשות האמורות להתמודד עם איומים אלה. ברם, אין עוררין שלביג דאטה פוטנציאל רב שעדיין לא מומש.

אוסף מחקרים זה נמנע במכוון מלהכתיב הגדרה קשוחה למונח 'ביג דאטה'. במקום זה, מאומצת גישה כוללת הרואה בנתונים בהיקפים גדולים הבאים ממקורות בלתי קונבנציונליים והמסופקים באופן דינמי ולפעמים בזמן אמיתי, כמאפיינים בולטים של הביג דאטה. חומר הגלם של הביג דאטה מגיע לעיתים קרובות בצורה בלתי מבוקרת ובמבנה לא מתאים לניתוח ססטמטי. בנוסף הוא לא בהכרח מגיע כמידע מרחבי. מידע המגיע בצורת עסקאות אשראי, דווחי צריכת אנרגיה או הערות ברשתות החברתית אינו 'מיוצר' כמידע מרחבי. אולם ברוב המקרים ניתן לעגן אותו מרחבית. לכן הוא הופך להיות מידע בעל זיקה עירונית אינהרנטית.

מקורות רבים מספקים ביג דאטה וכולם מקבלים ביטוי בקובץ הנוכחי. ראשית גופים מוסדיים ציבוריים ופרטיים כגון רשויות ממשלתיות ועירוניות, ספקי שירותים ותשתית הולכות ומשחררות מידע לציבור כחלק ממגמה של 'ממשל פתוח' (Artibas-Bel 2014) (open government). ביג דאטה זה הינו בעל מבנה פנימי ואיכות מבוקרת (structured) המאפשר תשאול וניתוח. הוא

* בהעדר מונח מקביל בעברית המתקבל על הדעת, נשתמש בביטוי 'ביג דאטה' לאורך מדור זה

מיוצר מלכתחילה למטרה זו וכולל סדרות נתונים רבות העוסקות בתחבורה, איכות חיים, צריכה פרטית, חינוך, דיור, פשע וכד'. ניתן לקשר בין הסדרות ולאפשר יצירת קבצי מידע גדולים. למשל, הצלבת מידע מכרטיסי תחבורה חכמים (כרטיסי רב-קו למיניהם) עם מידע אודות צריכת מים ביתית עשויה לאפשר בחירת מועד מתאים להחלפת צנרת עירונית.

מקור שני הוא מידע שמקורו בציבור. מידע זה נוצר באקראי על ידי פעולות שוטפת וסטנדרטיות הנרשמות ברשתות מחשבים. מידע זה הוא ההיפך מהמידע המוסדי. הוא מגיע בהיקפים גדולים, במהירות רבה, בתצורות שונות ובמבנה שהוא לגמרי unstructured. המידע לא נוצר מלכתחילה למטרות אנאליטיות ודורש לא מעט ניפוי על מנת להתאימו לניתוח העירוני. בראש וראשונה מדובר במידע שמקורו ברשתות חברתיות כגון Twitter, Foursquare, Instagram ובמידע שמסופק על ידי הציבור באופן התנדבותי ובצורה בלתי מתוכננת (VGI - Volunteered Geographic Information). המידע הזה מאפשר למתכנן לתקשר באופן ישיר עם הציבור ולחולל שינויים בתכנון, שירות ורגולציה באמצעות אפיק שלא היה קיים בעבר (Schweitzer 2014). הקישור בינו לבין הביג דאטה המוסדי, למשל באמצעות ציורים מעוגנים מרחבית, מהווה הזדמנות להבנת התנהגות הפרט בסביבה העירונית. חשוב גם לציין כי שימוש צולב כזה מהווה גם מקור לדאגה בסוגיות אתיות כגון שאלות של צינעת הפרט ושימוש שרלטני במאגרי מידע שלא ייעודו לתשאל בעת הקמתם (Kitchin 2013).

בין שני קצוות-מידע אלה קיים מקור נוסף והוא הביג דאטה הנובע מפעילות ניטור ודיווח של מכשירי מעקב ניידים וניידים. מכשירים סטטיים כגון ציוד לחישה מרחוק (לוויינים) או לחישה מקרוב (חיישנים) מספקים כמויות רבות של מידע אמין בתדירות גבוהה וברזולוציות מרחביות שונות. לרוב מידע זה קשור לניהול ותפעול מערכות עירוניות בתחומי הסביבה, תחבורה, תחזוקת מבנים, בטחון הציבור ואספקת מים ואנרגיה. מכשירים ניידים כגון מערכות האיכון וה-GPS שבטלפונים החכמים מאפשרים כור נרחב של אפשרויות ליצירת ביג דאטה ברזולוציה עיתית ומרחבית גבוהה. בזמן האחרון ההתפתחויות בתחום ה-Internet of Things פועלות להפוך מכשירים ניידים למכשירים ניידים.

קיומו של ביג דאטה מעלה מספר סוגיות חשובות בלב ליבה של המחקר העירוני-תכנוני:

התפתחות התיאוריה העירונית: האם יש בכוחו של הביג דאטה להניע תיאוריה עירונית חדשה? האם הוא מסוגל להניב תובנות חדשות ביחס לתהליכי צמיחה עירוניים? מחד גיסא, ביג דאטה מספק כמויות גדולות של מידע פרטני לטווח הקצר ולכן מאלץ הסתכלות מצומצמת וקצרת-טווח על ההתפתחות העירונית (Batty 2013a). מאידך גיסא, תיאוריה עירונית עוסקת בשינוי לטווח הארוך. שינוי זה קשור בתהליכי עומק כגון תמורות דמוגרפיות ושינויים בדפוסי צריכה שמחוללים העדפות חדשות למגורים ולצורת עבודה. על מנת להתמודד עם פער זה מתפתח ענף ידע חדש בשם 'מדעי העיר' (Urban Science) או 'אינפורמטיקה עירונית' (Urban Informatics). תחום זה מחבר בין לימודי מערכות מורכבות, מדעי המחשב ותכנון עירוני (Batty 2013b).

מדידה ומתודולוגיה: מהי השפעתה של ביג דאטה על שיטות העבודה במחקר העירוני? אין זו הגזמה לומר שביג דאטה הופך אל ההסקה הסטטיסטית על

ראשה. בדיקות של סיבתיות ואנדוגניות בין משתנים ומניעת הטיות בבחירה הופכות לבלתי-רלוונטיות מול מאגרי מידע ענקיים כאשר כל מה שנדרש הוא קשר מתאמי ביו משתנים. בעיה נוספת היא שבמהותו ביג דאטה הינו 'דאטה שמן' (fat data): יש בו הרבה יותר משתנים מסבירים מאשר תצפיות. לעומת זאת, מידע סטנדרטי הינו 'דאטה גבוה' (tall data) שבו יש יותר תצפיות מאשר משתנים מסבירים. לכן בביג דאטה יש יותר מידי משתנים מנבאים מהנדרש דבר המביא ל'התאמת יתר' של מודלים סטטיסטיים. בעיה זו מעלה את הצורך בפיתוחם של כלי סלקציה מבוססי למידה חישובית (machine learning) וגם בשיטות סטטיסטיות המסוגלות לבצע אמידה בתנאים אלה, כגון שיטת ה-LASSO: least absolute shrinkage and selection operator (Varian 2014).

ניהול והנגשה של ביג דאטה: איך ניתן לארוז את כמות המידע שמספק ביג דאטה ולהנגיש אותו בצורה ידידותית לציבור? המורכבות והנפחים של המידע העומד לרשות המנתח דורשים כישורים מגוונים. מיומנויות אלו כוללות יכולות בתחומי כריית מידע, שימוש במודלים שונים (סימולטיביים, חישוביים) חיבורי מקורות מידע (data fusion) וכלים לשליפה ולויזואליזציה. אדם אחד או יישום בודד אינם מסוגלים לכך, בדרך כלל. דרך אחת בעלת פוטנציאל להתמודד עם אתגרים אלה היא באמצעות הנגשת המידע לציבור ולקהילה המקצועית באמצעות מיפוי אינטרנטי דינמי (Lichter, Grinberger and Felsenstein 2015).

הטמעת ביג דאטה בפרקטיקה התכנונית: איך ניתן להטמיע את השימוש בביג דאטה במערכות התכנון? לשאלה זו יש לכאורה שני כיווני פעולה המשרתים את התכנון. אחד הוא הטמעת השימוש בביג דאטה בכלים הייעודיים לקבלת החלטות (planning support systems). דרך זו עשויה לחבר את הביג דאטה לפרקטיקה התכנונית בצד של התפוקות התכנוניות (Russo et al 2015). אופציה אחרת היא הגברת ההשתתפות האזרחית בהליך התכנוני באמצעות עידוד ה-VGI המסופק על ידי הציבור (Evans-Cowley 2010 and Hollander). דרך זו מקשרת את הביג דאטה לתשומות התכנוניות. עם זאת, שני הכיוונים האלה אינם תלויי-ביג דאטה: מערכות קבלת החלטות והשתתפות הציבור קיימות גם ללא ביג דאטה. ייתכן שמסלול שלישי עובר דרך ההכשרה של המתכנן. איש המקצוע חשוף היום למערכות מיחשוב ורישות גדולות ויש לו נגישות למידע בכמויות שלא היו בעבר. ברם, הוא נזקק לכלים אופרטיביים על מנת לנצל הזדמנויות אלו. הכלים המשלימים קיימים בתחומי הויזואליזציה וכריית המידע. בלעדיהם, האפשרויות הנפתחות על ידי מהפכת הביג דאטה, עלולות להישאר כאבן שאין לה הופכים בפרקטיקה התכנונית.

מטרת מדור זה הינה להציג דוגמאות של מחקרים יישומיים שהמכנה המשותף שלהם הוא השימוש בביג דאטה בהקשר העירוני-תכנוני. מפאת חדשנות התחום, קיים חוסר בחומרים בעברית בנושא הביג דאטה בכלל ובתחום העירוני בפרט. מתוך כלל המאמרים, שניים מהם דנים בביג דאטה שמקורו בגופים מוסדיים, שניים עוסקים בביג דאטה הנובע מאיתותי איכון שבטלפונים חכמים ומאמר אחד מדגים את הפקת ביג דאטה מציוצי טוויטר. מטבע הדברים התעניינות החוקרים מתרכזת באותם תחומים ראשיים

שבהם הביג דאטה הוא הזמין ביותר: תחבורה, פיתוח עירוני, מוביליות הפרט וניהול מצבי חירום.

עבודתם של שובל וברנבוים מתמקדת בדרכים לרתום את שיטות איתור לתכנון העירוני. מאמרם דן בשיטות השונות לחלוץ נתוני מיקום וזמן מדויקים מטלפונים ניידים. הם מונים את היתרונות של שיטה זו על פני מקורות מידע אחרים מבוססי-מקום כגון מידע הנובע מרשתות חברתיות או VGI. ביניהם ניתן לציין את היכולת לאסוף נתונים רבים בזמן קצר, לבצע ניתוחים חוזרים בתדירות גבוהה ולהפיק מידע ברזולוציה עיתית ומרחבית גבוהה. בתחום התכנון העירוני יתרונות אלה מתבטאים במיפוי מדויק על בסיס מידע מעודכן יותר ממה שנהוג היה בעבר, חיזוק טכנולוגי של שיתוף הציבור בתכנון והערכה טובה יותר של תוצאות התכנון לאחר מעשה.

בן אליה, בנסון ורופא משתמשים בטכנולוגיות ביג דאטה ובנתוני מ"מ"ג עירוניים על מנת לבחון את השפעת הרפורמה בתחבורה ציבורית של שנת 2011, על הנגישות של קבוצות אוכלוסיה שונות בעיר תל-אביב-יפו. הם נעזרים ביישום מבוסס-גרף (CTgraph) המאפשר חישובי נגישות ברמה של הבניין הבודד ובכל שעות היום. זאת בניגוד לקנה מידה המקובל במחקרי נגישות הנשען על יחידות מרחביות גסות כגון שכונות או אזורי תנועה ועל יחידות הזמן אגרגטיביות של בוקר וערב. במצב כזה, ברור כי השונות הפנים-יחידתית 'מגוהצת' על ידי הניתוח המצרפי. הממצאים מראים כי הרפורמה תורמת לכלל הנגישות בעיר אבל יוצרת דיפרנציאציה בלתי-שוויונית בין חלקי אוכלוסיה. הם מוצאים שרוב התועלת מהרפורמה נמדדת באותם אזורים בעיר שנהנו מנגישות גבוהה גם קודם לכן.

יצירת ביג דאטה 'סינתטי' למטרות תכנוניות מהווה את המוקד של עבודתם של גרינברגר, לייכטר ופלזנשטיין. הם מדגימים כיצד ניתן ליצור מידע זה מנתונים מוסדיים רגילים (כגון סקרים של הלמ"ס) וכיצד ניתן להשתמש בו לצורך תכנון התאוששות בעקבות אסון עירוני. העבודה נעזרת באלגוריתם ייעודי שמקצה מידע מרחבי הנאסף ברמה של אזורים סטטיסטיים, למבנים. מידע זה משולב במודל מבוסס-סוכנים בכדי לייצר דינמיקות עירוניות בזמן ובמרחב. תהליך פירוק המידע כולל מעבר ממסד נתונים המתאר מאות או אלפי יחידות מרחביות בלבד, למאגר המכיל מידע עיתי ומרחבי על מליוני מבנים ופרטים. הנתונים שנוצרים מאפשרים לאכלס שכבות שונות של מידע מרחבי עם תכונות סינתטיות. לדוגמה, נערכת סימולציה של רעידת אדמה היפותטית במרכז ירושלים. הממצאים מתייחסים הן להשפעות ארוכות טווח על שימושי קרקע והמורפולוגיה העירונית והן להשפעות לטווח הקצר על הסגרציה בין אוכלוסיות ועל הפגיעות של ציבורים שונים. התוצאות מונגשות לציבור באמצעות פיתוח מערכת של מיפוי דינמי מבוסס-אינטרנט.

מאמרם של גל-צור, קופליק, מינקוב, שור, גרנט-מולר ונוצ'רה מתמודד עם השאלה, האם ניתן להפיק מידע תחבורתי משמעותי ממדיה חברתית? המחברים סוקרים את מגוון המקורות של המדיה החברתית ואת המאפיינים הרלוונטים הגלומים במידע זה, לתכנון תחבורה ולניהול תנועה. הם מזהים שלושה סוגי תוכן במדיה החברתית היכולים לשמש כתשומות למידע תחבורתי: מידע המתייחס לכוונה של הפרט לבצע נסיעה, זיהויים של אירועים חריגים המתרחשים ברשת התחבורה והבעת עמדה ע"י הפרט ביחס לאיכות שירותי תחבורה. העבודה מביאה דוגמה להפקת מידע תחבורתי מצימצים של

נוסעים למשחקי כדורגל של קבוצת ליורופול באנגליה. היא מתארת של שלבי התהליך של כריית המלל (text mining) ומחשבת את מדדי ההצלחה ביחס למידע התחבורתי שהופק. מסקנות המחקר מצביעות על היכולת לדלות מידע משמעותי ממקורות של ביג דאטה שאינם מיועדים לכך אבל גם מצביעות על המאמצים הכרוכים בכך.

המחקר של אשבל, צירני ורזין מתמקד במקור נוסף של ביג דאטה: השימוש בנתוני טלפונים סלולריים. הוא מנתח את כמות העובדים במע"ר של תל-אביב ומקומות מגוריהם. הביג דאטה מופק מאיתותים סלולריים לחודש ייצוגי (יוני 2013). מידע זה מותאם חזק עם נתוני מפקד האוכלוסין 2008. הנתונים הסלולריים גם מצביעים על הדומיננטיות התעסוקתית של המע"ר וקיומם של שוקי עבודה מרחביים מצומצמים. הרלוציה העתית של המידע גם מאפשרת ניתוח של כניסות ויציאות למע"ר לאורך כל שעות היממה, דבר שמקל על זיהויים של שעות עומס. מידע זה מאשש ידע אינטואיטיבי ביחס לתפקוד המע"ר בשעות הלחץ. העבודה מדגימה את הפוטנציאל הגלום בהצלבת ביג דאטה עם מידע עיתי נוסף בתחום התחבורה, כגון שעות יציאה מערי הלווין לכוון המע"ר וההשלכות של זרמי תנועה עתידיים על דפוס עומס המתהווים במע"ר.

מקורות

- Arribas-Bel D (2014) Accidental, Open and Everywhere: Emerging Data Sources for the Understanding of Cities, *Applied Geography* 49, 45-53.
- Batty M (2013a) Big data, Smart Cities and City Planning, *Dialogues in Human Geography*, 3 (3), 274-279.
- Batty M (2013b) *The New Science of Cities*, MIT Press. Cambridge MA
- Evans-Cowley, J., & Hollander, J. (2010). The New Generation of Public Participation: Internet-based Participation Tools. *Planning Practice and Research*, 25(3), 397-408.
- Kitchin R (2013) Big data and Human Geography: Opportunities, Challenges and Risks *Dialogues in Human Geography* 3 (3), 262-267
- Lichter M, Grinberger Y and Felsenstein D (2015) Simulating and Communicating Outcomes in Disaster Management Situations, *International Journal of Geographic Information*, 4, 1827-1847.
- Russo P, Costabile M.F., Lanzilotti R and Pettit C.J. (2015) Usability of Planning Support Systems: An Evaluation framework, pp 337-353 in S. Geertman, J. Stillwell, J. Ferreira, & R. Goodspeed (eds.) *Planning Support Systems and Smart Cities*, Springer International Publishing, Switzerland.
- Schweitzer L. (2014) Planning and Social Media: A Case Study of Public Transit and Stigma on Twitter, *Journal of the American Planning Association*, 80 (3), 218-238.
- Varian H. R (2014) Big Data: New Tricks for Econometrics, *Journal of Economic Perspectives*, 28 (2), 3-28.