

## חלק ב – תיאוריה

מטרת חלק ב' של הספר, היא להציג בפני הקורא אוסף של מודלים ותיאוריות, המהווים תשתית להבנה ומימוש המסע הדיגיטלי. הפרקים בחלק זה הם:

### פרק 5 – חדשנות עסקית מבוססת דיגיטל

בפרק זה נסקור את נושא החדשנות (Innovation), כנושא בעל חשיבות אסטרטגית בכל ארגון עידן המודרני. נסביר את מושגי היסוד של היתרון התחרותי ומקורותיו – הובלת עלות ובידול – ואת החשיבות של החדשנות ביצירת היתרון התחרותי. נגדיר את עקומת ה S, עקומה המציגה את מחזור החיים של כל טכנולוגיה, מתחילתה ועד להופעת הטכנולוגיה או הרעיון החדש שמחליפים אותה. נציג את החדשנות המערערת (Disruptive Innovation) מאת פרופ' Clayton Christensen. נציג בקצרה שתי מתודולוגיות לפיתוח החדשנות: חשיבה עיצובית (Design Thinking) ועקומת הערך (Value Curve), הלקוחה מתוך אסטרטגיית האוקיינוס הכחול.

### פרק 6 – מודלים עסקיים מבוססי דיגיטל

פרק זה סוקר את מושג המודל עסקי (Model Business), ומתייחס לחשיבותו בעידן הדיגיטלי. הטרנספורמציה העסקית הדיגיטלית הוא אחד הכוחות רבי העוצמה שמניעים מודלים עסקיים חדשים, המשנים את כללי המשחק בכל הענפים ובכל הזירות העסקיות. יש חשיבות לכך שכל בעלי העניין בארגון – הדירקטוריון, ההנהלה הבכירה, המובילים הדיגיטליים יבינו את המודל העסקי הקיים, וכיצד לבנות מודלים עסקיים חדשים המתאימים לעידן הדיגיטלי.

### פרק 7 – נתונים – הנפט של העידן הדיגיטלי

פרק זה מוקדש לנתונים (Data). הנתונים היוו תמיד משאב ארגוני חשוב ותשתית הכרחית עבור כל המערכות הדיגיטליות. קשה לחשוב על יישומים דיגיטליים ללא מרכיב הנתונים. כל היישומים הדיגיטליים, בלי יוצא מן הכלל, משתמשים, מעבדים ומנהלים נתונים. כמובן שלנתונים תפקיד חדש בעידן הדיגיטל.

### פרק 8 – המערבולת הדיגיטלית – מבט ענפי

בפרק זה נמקד את הדיון ברמת הענף העסקי. הפרק מציג את מודל "המערבולת הדיגיטלית", המשנה באופן משמעותי את כל הענפים העסקיים הקיימים, ומושכת אותם אל עין המערבולת שבה הם הופכים לענפים דיגיטליים.

### פרק 9 – חמש פלטפורמות תומכות חדשנות

בפרק זה נרחיב את הדיון על התרומה הייחודית של הטכנולוגיות הדיגיטליות, בכל הקשור לחדשנות. הפרק מציג חמש קטגוריות של פלטפורמות דיגיטליות שונות, בעלות פוטנציאל רב להאצת חדשנות. הבנת פלטפורמות החדשנות האלה ואת זיקתן לטכנולוגיות הדיגיטליות, יאפשרו הבנה טובה יותר כיצד ניתן להביא ערך לארגון ולחולל חדשנות, נושא שהוא בעל חשיבות אסטרטגית לכל ארגון בעידן הדיגיטלי.

## פרק 7. נתונים – הנפט של העידן הדיגיטלי

*The Information about the package is as important as the package itself*

*Fred Smith, CEO, FedEx*

### 7.1 מבוא

מטרת פרק זה היא למקד את הזרקור באחד הנושאים החשובים של הטכנולוגיה הדיגיטלית, **הנתונים (Data)**, ולהצביע על אוסף של נושאים בהם הנתונים הם לב העניין. חברת McKinsey פרסמה בדצמבר 2016 ראיון עם Don Callahan שכותרתו *Rewiring Citi for the Digital Age*.<sup>76</sup> Don הוא ראש חטיבת התפעול והטכנולוגיה של Citi, אחד הבנקים הגדולים בעולם. הוא מצוטט כאומר "האינפורמציה חשובה בדיוק כמו הכסף". הוא התייחס כמובן לענף הבנקאות, אבל ניתן לומר שבעידן הדיגיטלי האינפורמציה אכן הפכה לחשובה כמעט כמו המוצר או השירות ברוב הארגונים. באחד הפרסומים של חברת FedEx, מנכ"ל החברה Fred Smith מצוטט כאומר *The Information about the package is as important as the package itself*.<sup>77</sup>

נתונים היוו תמיד משאב ארגוני חשוב ותשתית הכרחית עבור מערכת דיגיטלית. קשה לחשוב על יישומים דיגיטליים ללא מרכיב הנתונים: מערכות מידע ארגוניות כגון CRM, ERP, Billing או BI; יישומי מסחר אלקטרוני באינטרנט כגון Amazon, eBay, Alibaba ואחרים; יישומי רשתות חברתיות כגון Facebook, Twitter, YouTube, WhatsApp ואחרים; יישומי מובייל לניווט כמו Waze, יישומים להזמנות חדרים בבתי מלון כמו Booking.com או TripAdvisor, להשכרת חדרים כמו Airbnb או להזמנת הסעות כמו Uber ואחרים. כל היישומים הנ"ל, בלי יוצא מן הכלל, משתמשים, מעבדים ומנהלים נתונים. למרות חשיבותם הרבה לכל היישומים הנ"ל, ההתייחסות אל הנתונים הייתה דומה להתייחסות שלנו למנוע של מכונית – אנחנו לא ממש מבינים כיצד הוא עובד אולם ברור לנו שבלעדיו המכונית חסרת ערך. כך גם הנתונים – בלעדיהם רוב היישומים שאנו מכירים הם חסרי ערך.

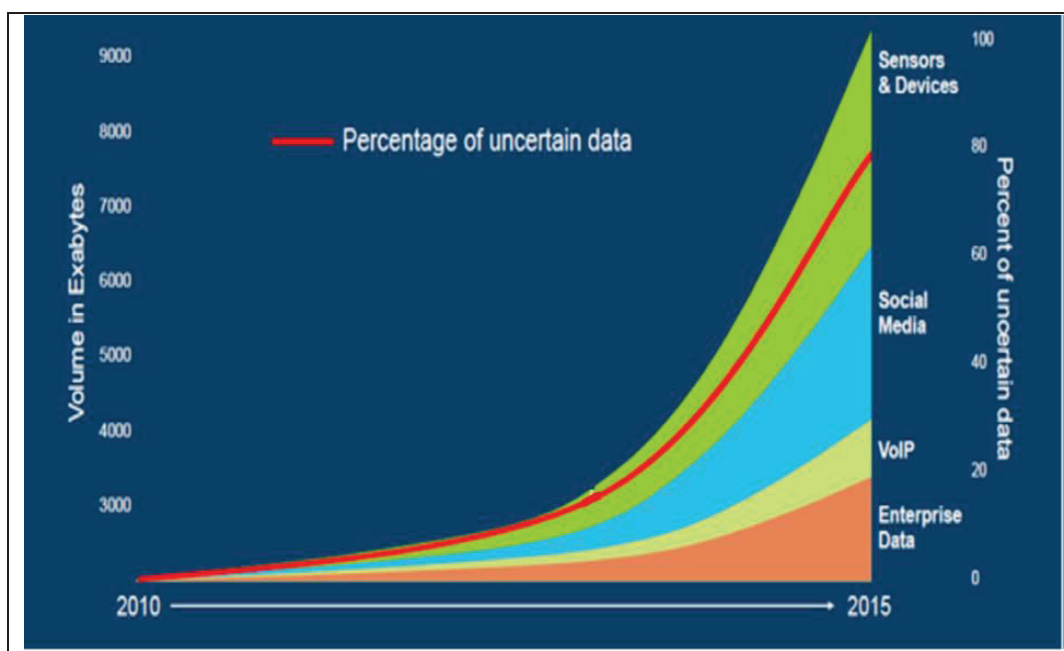
בעידן הדיגיטלי הנתונים עוברים לחזית העסקית, זוכים לכבוד הראוי להם ובמידה מסוימת ל"חיים משל עצמם". הם הופכים למשאב משמעותי ביכולת הארגון לייצר יתרון תחרותי, לקבל החלטות חכמות, לייצר מוצרים חדשניים ובחלק מהמקרים הנתונים עצמם הופכים למקור חדש של הכנסות. יש כאלה הקוראים לתופעה זו של העלייה בחשיבות הנתונים בשם כלכלת הנתונים (Data Economy).

בפרק זה נציג מספר נקודות מבט שונות על העלייה בחשיבות הנתונים והפיכתם למשאב קריטי, או כפי שיש כאלה הקוראים להם **הנפט של העידן הדיגיטלי**. אנו נציע לארגון מספר נקודות מבט וחיפוש הזדמנויות ממוקדות נתונים.

## 7.2 התפוצצות הנתונים – Data Explosion

אחד המאפיינים הבולטים של העידן הדיגיטלי הוא התפוצצות הנתונים – הגידול הבלתי נתפס בכמות הנתונים שנוצרים בעידן הדיגיטלי. מספר המערכות, המוצרים והסנסורים המייצרים ומשתמשים בנתונים גדל בצורה דרמטית והגידול בנפח הנתונים הפך לאקספוננציאלי.

מערכות המידע בארגונים אוגרות כמויות הולכות וגדלות של נתונים; כמות עצומה של מיילים נשלחים בעולם בכל דקה; מיליארדים של מכשירי סמארטפונים וטאבלטים, שהם מחשבים לכל עניין ונושא, מייצרים ומשתמשים בכמויות עצומות של נתונים; כמויות עצומות של נתונים נוצרים כתוצאה מפעילות המשתמשים ברשתות החברתיות (כמו YouTube, Twitter, LinkedIn, Facebook, e-Snapchat, WhatsApp, Instagram ועוד); אתרי המסחר האלקטרוני באינטרנט כגון Amazon, Alibaba, Bay ו-Aliexpress משתמשים ומייצרים כמויות עצומות של נתונים; מיליארדים של אנשים הגולשים באינטרנט ומפעילים יישומים שונים, כגון מנועי חיפוש של Google ואחרים מייצרים כמויות עצומות של נתונים; מיליארדים של סנסורים המותקנים כמעט בכל מוצר (כמו למשל במכוניות, במטוסים, במנועי הסילון, בטורבינות רוח, ברכבות, במעליות, במצלמות המותקנות ברחבי העיר, בעמודי תאורה ורמזורים, בבתים ובבניינים שהופכים ליותר ויותר חכמים, בחינוכים הפועלים ללא מגע יד אדם, במכוונות כביסה, במקררים, תגי RFID המותקנים בחבילות ועל גבי פריטים, מכשירי ניווט מבוססי GPS ועוד. כל אלה מייצרים ומשתמשים בכמויות עצומות של נתונים. האיור הבא מציג כיצד נפחי הנתונים הלכו וגדלו במשך השנים וכיצד המערכות שמייצרות אותם השתנו – ממערכות מידע ארגוניות לסנסורים והתקנים אחרים.



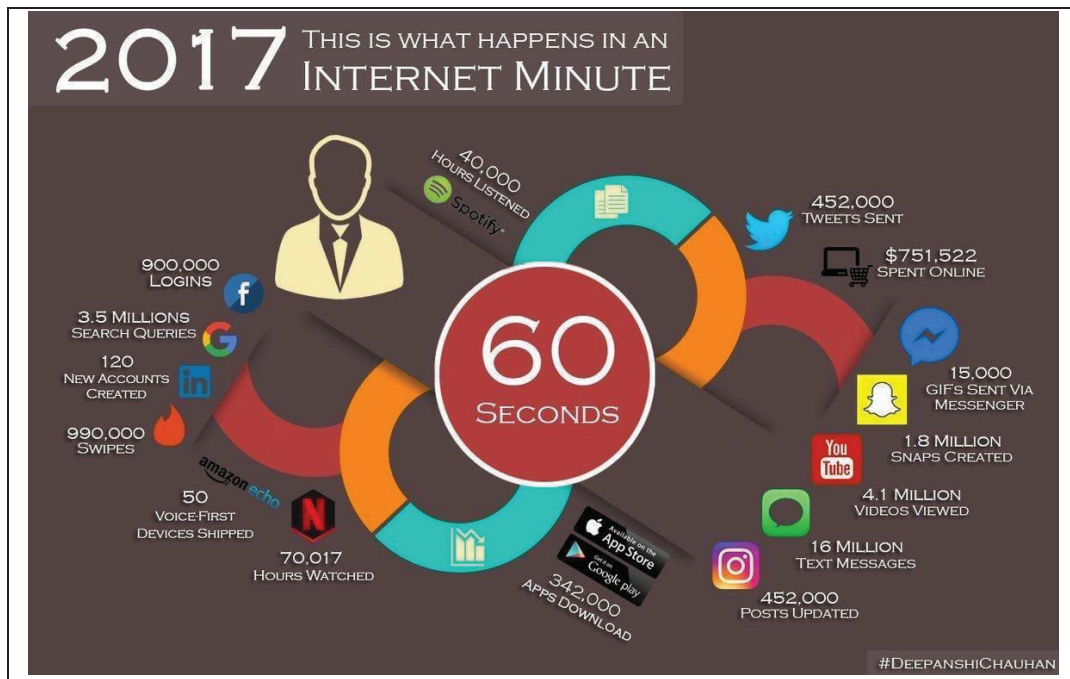
איור 7-a – הגידול בנפח הנתונים על פי המקורות השונים

לפי תחזיות שונות, אנו נהיה עדים בעידן החפצים החכמים (IoT) לעשרות מיליארדים של חפצים חכמים המחוברים לרשת האינטרנט, משדרים נתונים שניתן לנתח ולהפוך לתובנות ולשגר חזרה

הנחיות בזמן אמת לתפעול ותחזוקה חכמים יותר של החפצים. המציאות היא שארגונים מוקפים בכמויות עצומות של נתונים, אולם לא תמיד יודעים ומצליחים לנצל ולהפיק מהם את התובנות ולהשתמש בהם בתהליכי העסקיים ובתהליכי קבלת ההחלטות. כבר בשנת 1991 כתב John Naisbitt, סופר ומרצה פופולרי, בספרו <sup>78</sup> Megatrends 2000: **We are drowning in data but starving for knowledge** - **אנחנו טובעים בנתונים אבל רעבים לידע**. האתגר של כל ארגון בעידן הדיגיטלי הוא ללמוד כיצד להפוך את נפחי המידע העצומים למקור של תועלת עסקית ויתרון תחרותי.

### 7.3 נתונים – משאב בגדילה מתמדת

האיור הבא מציג את הפעילות ביישומים אינטרנטיים שונים במהלך דקה אחת בלבד! כ 21 מיליון מסרים של WhatsApp, כ 350 אלף ציוצים חדשים ב Twitter, כ 150 מיליון מיילים נשלחים, 2.4 מיליון שאילתות נשאלות ב Google, מורידים כ 51 אלף יישומונים מחנות ה AppStore של Apple, 701 אלף משתמשים מתחברים ל Facebook וכך הלאה. פשוט בלתי נתפס!



איור 7-7 – התפלגות הפעילות באינטרנט במהלך דקה אחת - 2017

ב-2017 יש כ 2 מיליארד אנשים שמחוברים לאינטרנט, כ 4 מיליארד מנויים ברשתות הסלולריות, מיליארדים של מכשירים וסנסורים מחוברים לאינטרנט של החפצים (IoT). ההערכה היא כי עד שנת 2020 יהיו בין 40 ל 60 מיליארד חפצים חכמים שיהיו מחוברים לאינטרנט. חברת Walmart מאחסנת כ 2.6 Petabytes של נתונים מדי שעה בעקבות הפעילות המסחרית של לקוחותיה. 90% מכל הנתונים בעולם נוצרו בשנתיים האחרונות. ההערכה היא שמדי חודש נוצרים כ 2.5 Exabytes של נתונים חדשים (Exabyte אחד מכיל מיליארד Gigabytes).



איור 7-7 – פעילות באינטרנט במהלך דקה אחת בלבד

מנוע במטוס של חברת Airbus מייצר 10 Terabytes של נתונים על כל חצי שעת טיסה. ארבעת המנועים המותקנים במטוס הענק Airbus 380 מייצרים במהלך טיסה אחת מלונדון לניו יורק 640 Terabytes של נתונים. למוח האנושי קשה לתפוס את כל המספרים שהמכשירים והשימושים הדיגיטליים מייצרים ואת קצב הגידול המדהים שלהם.

## 7.4 התפתחות הבגרות האנליטית

מקובל להתייחס אל Clive Humby, מתמטיקאי ואחד מהוגי הרעיון של מועדון הלקוחות המצליח של ענקית הקמעונאות Tesco, כאל מי שהשתמש ראשון באנלוגיה בין הנפט והנתונים. בשנת 2006 הוא צוטט כמי שאמר: "הנתונים הם הנפט החדש. הנפט הוא בעל ערך וניתן להשתמש בו רק לאחר שהוא עובר תהליכי עיבוד וזיקוק. חייבים להפוך את הנפט לגז, בנזין, פלסטיק, כימיקלים ועוד כדי להפוך אותו לבעל ערך שיכול להביא ולהניע פעילויות ורווחיות. בדומה לנפט, גם את הנתונים יש לעבד, לסנן, להעשיר ולנתח כדי שהם יהיו בעלי ערך לארגונים".

מאז רבים המשיכו להשתמש באנלוגיה זו בין נפט ונתונים, שני משאבים משמעותיים ליצירת עושר ותועלת לאנושות בכלל ולארגונים בפרט. למשל Peter Sondengaard, סמנכ"ל בכיר בחברת המחקר Gartner, כתב במאמר שהתפרסם ב 2015 ב Forbes<sup>79</sup>: "ביג דאטה הוא הנפט של המאה ה 21. למרות כל הערך שלהם, הנתונים הם טיפשים. הם אינם עושים כלום עד שאתה לומד כיצד להשתמש בהם. גם הנפט הגולמי הוא חסר ערך עד שמזקקים אותו והופכים אותו לדלק. גרסת הביג דאטה של נפט

מזוקק – אלגוריתמים ייחודיים שנועדו לפתור בעיות מסוימות וניתן לתרגם אותם להחלטות ופעולות. זהו המתכון הסודי של הארגונים המצליחים בעתיד. הבהלה לזהב של העידן הדיגיטלי תהיה ממוקדת בכיצד ניתן לעשות דברים בעלי ערך עם הנתונים". מנכ"לית IBM, Ginni Rometty, אמרה ב 2013: "אני מבקשת שתחשבו על נתונים כעל המשאב הטבעי הבא".

הנפט כמעט חסר ערך במצבו הגולמי. רק לאחר עיבודו וזיקוקו הוא הופך לאוסף של מוצרים יקרי ערך שקשה לדמיין את האנושות בלעדיהם – בנזין, פלסטיק, ניילון, שמני סיכה וכד'. בדומה לנפט, גם הנתונים במצבם הגולמי כמעט חסרי ערך. רק לאחר עיבודם הם הופכים למוצרים יקרי ערך המסוגלים לתמוך בפעילות השוטפת של הארגונים בתהליכי קבלת החלטות, בניתוחי מגמות ואיתור חריגים, בסימולציות What If, בהצגת תמונה גרפית למנהלים (Dashboards) ועוד. לשם כך, הנתונים עוברים תהליך של עיבוד - הם מופקים ממקורות המידע השונים (בסיסי נתונים תפעוליים, רשתות חברתיות, מערכות חיצוניות, סנסורים וכד'), עוברים תהליך של ניקוי והעשרה ונטענים למאגרי נתונים מיוחדים (מחסני נתונים, אגמי נתונים – Data Lakes ועוד). אגמי הנתונים (Data Lakes) הם תפיסה חדשה יחסית המאפשרת אחסון מהיר של נתונים בפורמטים שונים מבלי לכפות עליהם מבנה (סכמה) כלשהו ומבלי לבנות להם אינדקסים עם כניסתם למאגר (פעולה שמאיטה מאד את תהליך הכנסת הנתונים למאגר). להבדיל מהנפט, שהוא משאב מתכלה ככל שממשיכים בהפקתו מבטן האדמה, הנתונים רק הולכים ומתרבים וכל יום שעובר מוסיף כמויות עצומות של נתונים.

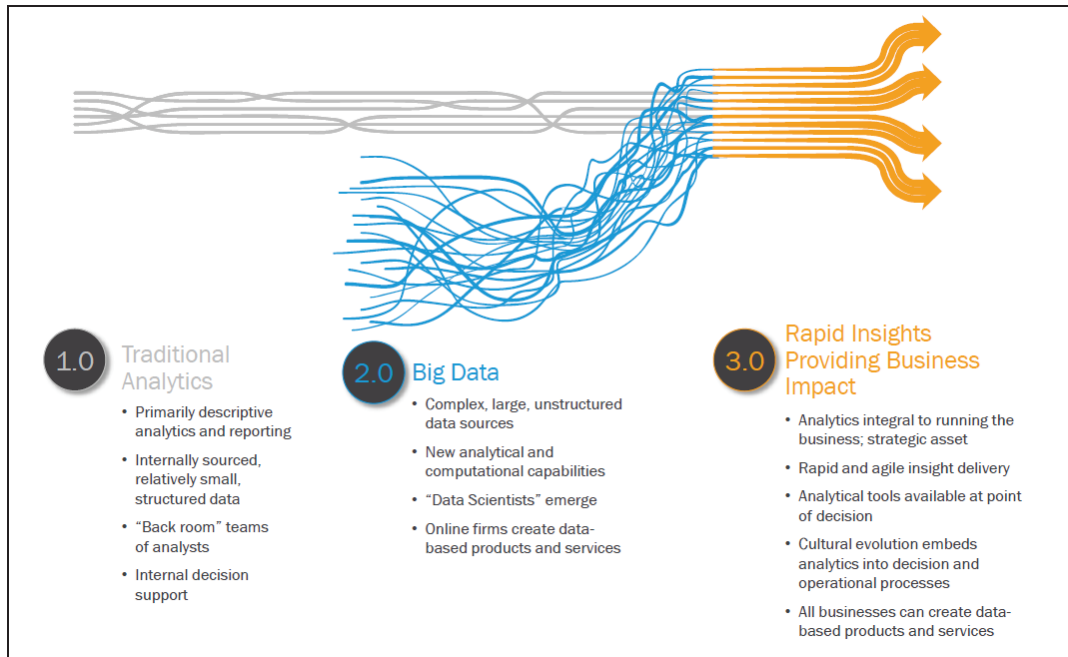
אם הנתונים הם הנפט של העידן הדיגיטלי, אזי האנליטיקה העסקית היא בתי הזיקוק של העידן הדיגיטלי. בית הזיקוק הוא המתקן המשמש לעיבוד הנפט הגולמי והפיכתו לאוסף מוצרים בעלי ערך רב. בדומה, מערכות האנליטיקה העסקית הן אלה המעבירות את הנתונים הגולמיים דרך אוסף של תהליכים מורכבים (גזירה, העשרה, בדיקה, טעינה, עיבוד וניתוח) עד להפיכתם למידע, ידע ותובנות עבור המשתמשים והארגונים.

אנליטיקה איננה נושא חדש. היא החלה להתפתח קצת אחרי הופעת מערכות המידע בשנות ה 60 של המאה הקודמת. במהלך השנים המערכות האנליטיות זכו לשמות רבים וביניהם – Business Intelligence, Data Warehousing, Decision Support Systems ובשנים האחרונות Data Analytics והפכו ממערכות פשוטות יחסית שעסקו בעיקר בדיווח (Reporting) למערכות העוסקות בניתוחים מורכבים ועוסקות בחיזוי, ניתוח מגמות, סימולציות ועוד.

מאמרו של פרופ' Tom Davenport ימשש אותנו כבסיס להבנת וסקירת ההתפתחות של המערכות האנליטיות. פרופ' Davenport, אחד החוקרים המובילים בתחום האנליטיקה העסקית ומי שחיבר מחקרים וספרים רבים בתחום וכיום חוקר ב MIT, פרסם מאמר במגזין Harvard Business Review בדצמבר 2013 שכותרתו הייתה Analytics 3.0 – In the New Era Data Will Power Consumer Products and Services<sup>80</sup>.

פרופ' Davenport חילק את התפתחות המערכות האנליטיות לשלוש תקופות – אנליטיקה 1.0 עידן הבינה העסקית שהחל להתפתח אי שם בשנות ה 60 של המאה הקודמת, אנליטיקה 2.0 – עידן הביג דאטה שהחל בסביבות שנת 2000 ואנליטיקה 3.0 העידן המודרני של השנים האחרונות המשלב את

הבינה העסקית עם הביג דאטה כתשתית לכלכלת נתונים (Data Economy). בעידן הזה לא ניתן יותר להבדיל בין העסקים לבין הנתונים – הם שלובים זה בזה. האיור הבא מציג את שלושת השלבים האלו ואת המאפיינים העיקריים של כל שלב.



איור d-7 – שלושת השלבים של אנליטיקה עסקית עפ"י Davenport

נסקור בקצרה את שלושת השלבים:

### 7.4.1 אנליטיקה עסקית 1.0 – עידן הבינה העסקית Business Intelligence

זהו העידן הראשון שבו החלו להתפתח תפיסת מחסן הנתונים והבינה העסקית. תחילה מחוללי דוחות ושאליות ואחר כך החלו להופיע כלים מתוחכמים לגזירת הנתונים ממערכות המקור (כלי ETL) ובהמשך כלים מתקדמים לניתוח, וויזואליזציה ותחקור. עם השנים כלים אלה הפכו למתוחכמים וגמישים. כיום רוב הארגונים מיישמים את תפיסת מחסן הנתונים ומשתמשים בכלים אנליטיים כחלק מההליכי קבלת החלטות ואף לפיתוח היתרון התחרותי. ניהול מחסן הנתונים הפך לחלק בלתי נפרד של המערכות המנוהלות ע"י אגפי טכנולוגיות המידע. המאפיינים העיקריים של עידן אנליטיקה 1.0 הם:

- א. מקורות הנתונים מצומצמים יחסית ועיקרם משמש למטרות פנים ארגוניות. הנתונים ברמת מבניות גבוהה.
- ב. הפעילות האנליטית מוגבלת יחסית וממוקדת בעיקר בהפקת דוחות וניתוח של מה שכבר קרה (Descriptive Analytics and Reporting). מי שמבצע את הניתוח הם אנליסטים בסיסיים של BI.

- ג. תהליך הניתוח מורכב יחסית ומחייב גזירת הנתונים ממערכות המקור, הכנתם לטעינה למחסן הנתונים וביצוע של ניתוחים באמצעות כלים אנליטיים ייעודיים.
- ד. האנליסטים שעסקו בניתוח הנתונים פעלו מאחורי הקלעים והכינו דוחות וניתוחים עבור המנהלים ובדרך כלל לא היו חלק מהתהליכים העסקיים וקבלת ההחלטות.
- ה. רמת הזיקה בין האנליטיקה לבין האסטרטגיה העסקית הייתה מוגבלת והיא בדרך כלל ללא השפעה ישירה על היתרון התחרותי של הארגון.
- ו. בחלק ניכר מהארגונים, מנהלים המשיכו לבסס את ההחלטות שלהם על אינטואיציה וניסיון ופחות על ניתוח של נתונים.

### 7.4.2 אנליטיקה עסקית 2.0 – עידן ה Big Data

ההתפתחות העצומה בכוח המחשוב, הירידה המתמדת בעלויות החומר ועלויות יחידות האחסון ובמקביל התפוצצות הנתונים בעקבות יישומי אינטרנט ומובייל, גרמו לאנליטיקה לעלות שלב, שלב הביג דאטה, שלב שהחל אי שם בשנות ה 2000. אמנם המונח Big Data הופיע בשלב מאוחר יותר, בסביבות 2010, אבל חלק ממאפייניו החלו להופיע מוקדם יותר, מיד עם תחילת עידן האינטרנט. אתרי המסחר האלקטרוני והרשתות החברתיות שצמחו במהירות עצומה, הביאו איתם אתגרים חדשים בניהול הנתונים ולצורך לטפל בכמויות עצומות של נתונים ובמגוון רחב של פורמטים (טקסט, תמונות, וידאו, קול וכד').

חברות האינטרנט החדשות והגדולות וביניהן Google, Amazon, YouTube, e-Bay, Facebook, ואחרות, החלו לפתח שיטות וטכנולוגיות חדשות להתמודדות עם היקפי הנתונים הגדולים. הנתונים הפכו לאחד הנכסים העיקריים של חברות אלה והן החלו בפיתוח מוצרים ושירותים חדשים מבוססי נתונים. לדוגמה חברת LinkedIn פיתחה מוצרים כגון People You May Know ו Jobs You May Be Interested In, מוצרי פרמיום עליהם היא גובה תשלום ועוד.

הצורך בעיבוד מקבילי מסיבי הביא להתפתחות טכנולוגיית Hadoop שהיוותה את אחד הפתרונות לכך. דור חדש של מערכות לניהול בסיסי נתונים נולד – NoSQL (Not Only SQL), בסיסי נתונים המסוגלים לטפל במגוון עשיר של פורמטים שונים של נתונים ובזרימה מהירה של נתונים הנכנסים לבסיס הנתונים. טכנולוגיית מחשוב ענן גם כן סיפקה פלטפורמה לאחסון ועיבוד כמויות עצומות של נתונים. החלו להופיע גם טכנולוגיות לעיבוד נתונים בזיכרון (למשל Hana של SAP) וכן טכנולוגיות שהביאו את כוח העיבוד אל בסיס הנתונים, במקום להביא את הנתונים אל השרתים. כפי שניתן לראות, הדרישות האנליטיות של העידן החדש היו שונות מאד מהעידן הראשון ואף הביאו להופעת מקצוע חדש – **מדען הנתונים** (Data Scientists). נסכם את מאפייני עידן האנליטיקה 2.0:

- א. מקורות נתונים חדשים, עצומים בנפחם ובגיוון גדול של פורמטים לא מובנים. ההגדרה המקובלת של Big Data נקרא 3V ומצביע על נפח הנתונים הענק – **Volume**, על מגוון גדול של סוגי הנתונים – **Variety** ועל המהירות העצומה בכניסת הנתונים למאגר – **Velocity**<sup>81</sup>.



- ב. הופעת יכולות אנליטיות חדשות הנדרשות להתמודד עם האתגרים החדשים.
- ג. הופעת ההתמחות החדשה של מדעני נתונים (Data Scientist) שנועדה להתמודד עם האתגרים של עיבודים אנליטיים מורכבים וחקר נתונים מתקדם להפקת תובנות (Insights) חדשים מתוך הנתונים.
- ד. חברות חדשות, בעיקר ענקיות האינטרנט, החלו לפתח זן חדש של מוצרים, מוצרי נתונים (Data Products), שהיוו עבורם מקור חדש של הכנסות.
- ה. חלק משמעותי ממקורות הנתונים המשמשים לאנליטיקה עבר ממקורות נתונים פנימיים למקורות חיצוניים כמו סנסורים, רשתות חברתיות וכד'.
- ו. הופעת למידת מכונות (Machine Learning) האיצה את הצורך ביכולות מתקדמות של עיבוד נתונים וגם סיפקה פלטפורמה ליכולות ניתוח מהירות ביותר.

### 7.4.3 אנליטיקה עסקית 3.0 – עידן כלכלת הנתונים

רבים יחשבו שאנו עדיין בעידן השני, אולם כבר ניתן לראות סימנים לעידן חדש של אנליטיקה. זהו העידן המשלב בצורה מיטבית את שני העידנים הקודמים. זוהי סביבה הלוקחת את הטוב מהבינה העסקית ומהביג דאטה כדי לקבל תובנות מהירות ולפתח מוצרים חדשים מבוססי נתונים בעלי השפעה משמעותית על הארגון. זהו עידן כלכלת הנתונים.

אמנם אנו עדיין בשלב ראשוני של עידן אנליטיקה 3.0, אולם כבר ניתן לראות ארגונים מתחומי התעשייה, שירותי הבריאות, קמעונאות, פיננסים ועוד המשתמשים בנתונים כדי לייצר אוסף חדש של מוצרים מבוססי נתונים המביאים ערך חדש ללקוחותיהם ופלטפורמות מתקדמות להפקת תובנות מהירות למנהלי הארגון. המאפיינים העיקריים של עידן אנליטיקה 3.0 הם:

- א. האנליטיקה הפכה לנכס אסטרטגי, היא הכרחית לפעילות הארגון.
- ב. יכולות לייצור מהיר של תובנות עסקיות.
- ג. כלים אנליטיים מתקדמים זמינים למקבלי החלטות המסופקים ביעילות בנושאים בהם הם נדרשים.
- ד. ההתפתחויות בתרבות הארגונית מביאות לשילוב של האנליטיקה בתהליכי קבלת החלטות ובתהליכים העסקיים (Data Driven Decision Making).
- ה. כל הארגונים יכולים לייצר מוצרים מבוססי נתונים (ולא רק ארגוני האינטרנט). אנו עדים להופעה של כלים חדשים המשלבים יכולות בתחום אחסון/שינוע/אנליזה/עיבוד והגשה בו זמנית לעומת מספר פתרונות נפרדים.

להלן מספר דוגמאות לחברות העושות שימוש באנליטיקה כנכס אסטרטגי:

- א. **חברת התובלה הענקית Schneider National** עושה שימוש הולך וגובר בנתונים המגיעים ממקורות חדשים (כמו רמות הדלק במשאיות, מיקום המכולות, התנהגות הנהגים ואינדיקטורים נוספים) כדי להשתמש באלגוריתמים מתוחכמים לאופטימיזציה של הלוגיסטיקה. המטרה שלהם היא לחשב מסלולים אופטימליים, להוריד את עלויות הדלק ולהקטין את מספר התאונות.
- ב. **ענקית התעשייה General Electric** הגדירה את נושא האינטרנט התעשייתי, ה Industrial IoT, כנושא אסטרטגי וכמקור חשוב ליצירת היתרון התחרותי שלה בעידן הדיגיטלי. כוונתה היא לחבר את כל ההתקנים התעשייתיים שלהם ולשלב בכל מוצריה סנסורים המספקים נתונים על הביצועים שלהם. היא הקימה חטיבת תוכנה ענקית, פיתחה אוסף של פרוטוקולים וכלים תחת המותג Predix ומשקיעה רבות בקידום תפיסת ה IoT התעשייתי שלה. החברה אוספת את הנתונים, מנתחת אותם באמצעות כלי ביג דאטה ומספקת ללקוחותיה תובנות רבות בנושא תחזוקת המוצרים וניצול אופטימלי שלהם. החברה מקדמת את הרעיון של Digital Twin, אובייקט תוכנה חכם המדמה את פעולת המכשיר, כלומר אובייקט דיגיטלי המדמה בזמן אמת את האובייקט הפיסי, ומאפשר הפקת תובנות בזמן אמת וגם הנחיות חוזרות להתקן הפיסי עצמו. למשל לכל טורבינת רוח שהם מייצרים ומתקינים יהיה תאום דיגיטלי, המקבל נתונים בזמן אמת מהטורבינה עצמה. התאום הדיגיטלי מאפשר ניתוח בזמן אמת של מצב הטורבינה ולהפיק תובנות בדבר תפעול הטורבינה להפקת חשמל בצורה היעילה ביותר ולשדר אותם מידית לטורבינה כדי לשנות את זווית הפעולה, מהירות הסיבוב וכד'. תדמינו מצב שבו לכל מנוע סילון, טורבינה, קטר, מכונה לדימות רפואי ש GE מייצרת ומתקינה יהיה גם תאום דיגיטלי המחובר לתאום הפיסי בזמן אמת ומדמה את פעולתו.
- ג. **ענקית המוצרים Procter & Gamble** עברה משימוש בבינה עסקית בנושאים נקודתיים לשימוש בליבת עשיית העסקים שלה. החברה שילבה את נושא האנליטיקה העסקית בכל תהליכי ובנתה 500 חדרי דיונים מיוחדים, Business Sphere. חדרים אלה עתירים במסכי תצוגה בהם המנהלים עוסקים בקבלת החלטות ומסתייעים בכלים ובתצוגות האנליטיות.
- ד. **חברת UPS** משתמשת במערכת ORION כדי לבצע אופטימיזציה בזמן אמת של מסלולי הנסיעה עבור 55,000 נהגיה.

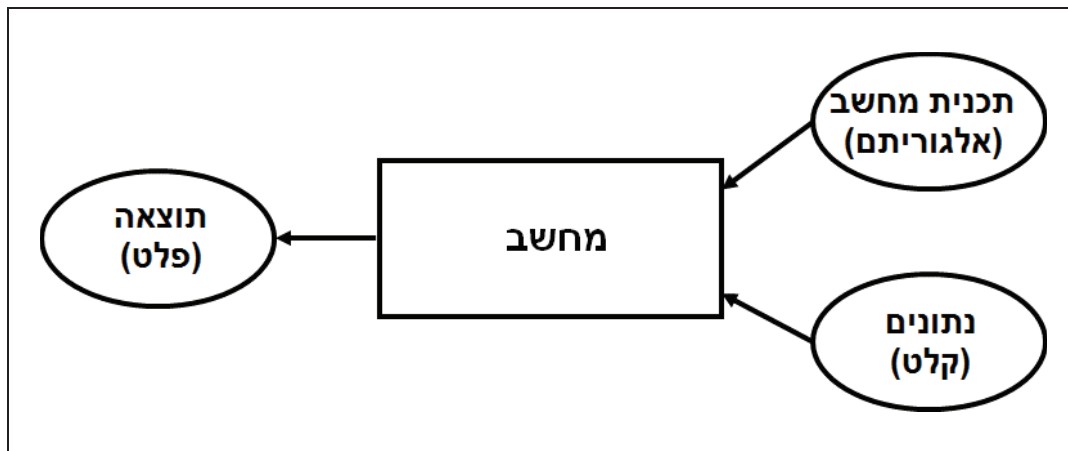
כל הדוגמאות הנ"ל של יישומים מתקדמים של אנליטיקה עסקית המשלבים את הבינה העסקית ואת הביג דאטה כדי לייצר תובנות מהירות ויעילות, ובחלק מהמקרים לייצר מוצרים מבוססי נתונים ללקוחותיה.

## 7.5 נתונים – הדלק של למידת מכונה – Machine Learning

השנים האחרונות הביאו לשינוי דרמטי בתחום מדעי המחשב, מהפכה שקטה שהביאה לעליה חדה בשימוש בנושא למידת המכונות, ענף של תחום האינטליגנציה המלאכותית. האינטליגנציה

המלאכותית היא ענף מחקרי וותיק בתחום מדעי המחשב, נושא שעורר את דמיונם של רבים אולם לא הצליח לפרוץ למרכז הבמה.

משך עשרות שנים שלטה בעולם המחשוב פרדיגמה אחת, Rule Based Programming – האדם (מותכנת) מפתח את האלגוריתם ותוכנה מיוחדת (הקומפיילר) מתרגם אותו לשפת מחשב כלשהי. המחשב מבצע במהירות ויעילות את מה שהאלגוריתם קובע. הוא מקבל קלט מסוים והופך אותו לפלט מסוים בהתאם להנחיות האלגוריתם. המשמעות היא שעל מפתח האלגוריתם לחשוב ולהתייחס מראש לכל המצבים האפשריים ולהגדיר מה על המחשב לעשות בכל מצב כזה. האיור הבא מתאר פרדיגמה זו.



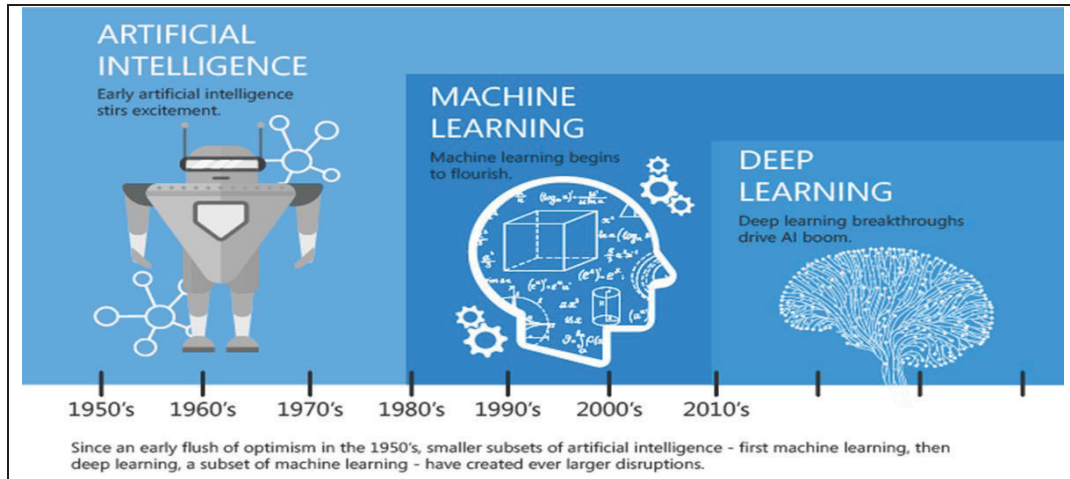
איור e-7 – הפרדיגמה המקובלת של תוכנות מחשב

בסוג מסוים של יישומים וביניהם יישומים קוגניטיביים כגון ראייה, זיהוי דיבור, תרגום, רובוטיקה ועוד, פרדיגמה זו לא פעלה היטב. למרות שנות מחקר רבות, המדענים לא הצליחו לפתח אלגוריתמים טובים. שני שינויים בולטים שהתרחשו בשנים האחרונות, הביאו לשינוי הפרדיגמה:

א. **ההתפוצצות בכמות הנתונים הדיגיטליים** – תופעה אותה תיארו כבר והביאה לבניית יכולות של ניהול וניתוח נתונים בעלי פורמטים שונים.

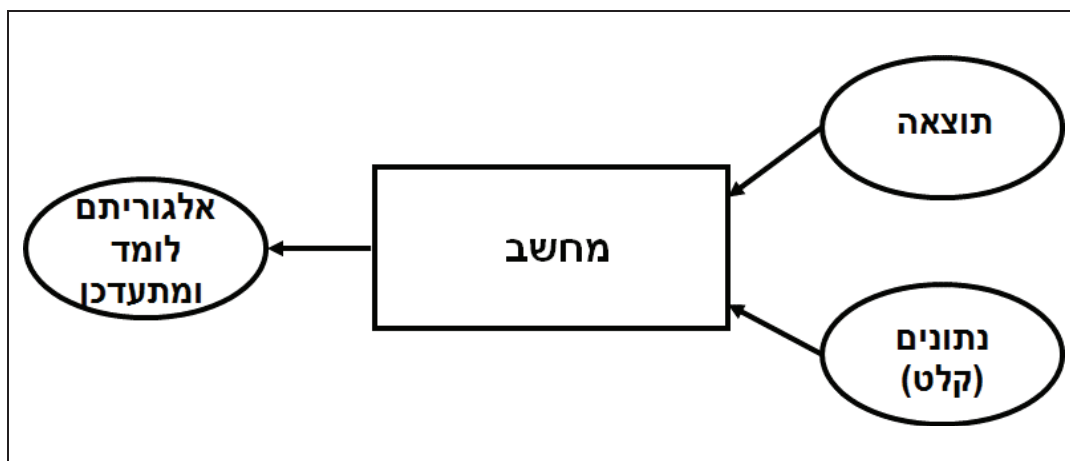
ב. **התפתחות כוח המחשוב** – עצמת המעבדים, גודל הזיכרון, מהירות התקשורת, יכולות עיבוד מקביליות – כל אלה הביאו לעוצמות חדשות של מחשב.

שתי התפתחויות אלה, שהשלימו זו את זו, במקביל לפיתוח והתקדמות בחקר הרשתות הנוירוניות המלאכותיות (Artificial Neural Networks), הביאו לפריצת הדרך בלמידת מכונות ובעיקר ביכולת הלמידה מתוך הנתונים. הרעיון, הלא חדש חייבים לומר, של ללמוד ולהסיק את האלגוריתם מתוך הנתונים ואת החלפת התכניתנים במערכות תוכנה לומדות, הביא להופעת הפרדיגמה החדשה – **למידת מכונות** ובשנים האחרונות **לרשתות למידה עמוקה (Deep Learning)**. אלו מערכות תוכנה, המשנות את עצמן ככל שהן נחשפות לנתונים רבים יותר, כלומר לומדות מתוך הנתונים. האיור הבא מציג את ההתפתחות של האינטליגנציה המלאכותית ולמידת המכונות והוא לקוח מתוך המאמר של Milan Manwar<sup>82</sup>.



איור f-7 - ההתפתחות של האינטליגנציה המלאכותית

כיום מערכות תוכנה אלו מסוגלות לזהות בצורה טובה דיבור (Speech Recognition), לזהות פרצופים של אנשים (Recognition Face) או חפצים, להמליץ באמצעות מנוע המלצות (Recommendation Engine) על הספרים שכדאי לקרוא או הסרטים שמומלץ לו לראות בעקבות למידה של הרגלי הקריאה או הצפייה בסרטים של לקוחות. אם בפרדיגמה הקודמת, התכניתן נדרש להגדיר מראש את האלגוריתם ומה על המחשב לעשות בכל מצב, בפרדיגמה החדשה ולמשימות מסוימות, אין צורך בכתיבת תוכנה אלא באימון המערכת (שלב ה Supervised Learning או Training) ע"י חשיפתה לכמויות גדולות של נתונים (כלומר לדוגמאות רבות). הנתונים כוללים את המופעים ומה על המערכת לעשות. למשל מציגים למערכת דוגמאות רבות של תמונות בהן יש ציון (Label) אם בתמונה מופיע חתול או כלב - והמערכת מעדכנת את המשקלות בצמתים של הרשת הניורנית שלה עד שתוכל לזהות באופן אוטומטי את מה שמופע בתמונות חדשות (שלב הפעלת המערכת). למשל אם נראה למערכת תמונה כלשהי, היא תוכל לזהות אם מדובר בחתול או בכלב. כאמור, היישום של למידת מכונות מתבסס במידה רבה על התיאוריה של רשתות עצביות מלאכותיות. בתהליך הלמידה מאמנים את הרשת ע"י הצגת דוגמאות והמשקלות של הצמתים ברשת משתנות בהתאם. האיור הבא מתאר את הפרדיגמה החדשה.



איור g-7 – הפרדיגמה החדשה של לימוד מכונות

דיון מעמיק יותר בנושא למידת מכונות והתיאוריה המתמטית והסטטיסטית העומדות בבסיסה הוא מעבר למטרות פרק זה. אולם דבר אחד ברור ובוולט – **הנתונים וה Big Data, במקביל להתפתחות המדהימה בעצמת המחשוב** (המחשוב המקבילי והיכולת לאחסן נפחים ענקיים של נתונים), מניעים ומקדמים את טכנולוגית למידת מכונות והאינטליגנציה המלאכותית. מערכות אלה מהוות תשתית משמעותית ליישומים רבים ומתקדמים בעידן הדיגיטלי, עידן המתאפיין בצורך ביצירת תובנות עסקיות מתקדמות המתעדכנות תוך כדי זרימת נתונים חדשים ולעיתים אף בזמן אמת. מערכות כאלה כבר משולבות במכשירי הסמארטפון ואחרים במערכות כגון Siri, Alexa, Cortona ומערכות מחשוב קוגניטיביות כגון Watson.

## 7.6 מודלים עסקיים מבוססי נתונים – Data Driven Business Models

נתונים היו מאז ומעולם רכיב חשוב ביכולת הארגון לפעול ולממש את האסטרטגיה שלו. הנטייה הייתה לחשוב על היישומים כעל הרכיב הקריטי ואל הנתונים כעל מרכיב משני. המגמה בשנים האחרונות היא הפיכת הנתונים עצמם לרכיב אסטרטגי, כלומר למרכיב המשפיע ישירות על היתרון התחרותי של הארגון ומאפשר לארגון לחולל זרם חדש של הכנסות. חברות המחקר השונות (IDC, Forrester, Gartner) צופות שארגונים רבים יבינו את העוצמה הטמונה בנכס החדש, הנתונים (Data as an Asset), ויפעלו לנצל אותו ליצירת תועלות עסקיות חדשות. במילים אחרות – ארגונים רבים יגבשו מודלים עסקיים מבוססי נתונים. ניתן לזהות מספר סוגי מודלים עסקיים הממוקדים בנתונים:

א. **מכירת נתונים לעסקים אחרים** – זהו המודל העסקי הידוע והוותיק ביותר המבוסס על נתונים. מזה שנים רבות ארגונים נוהגים למכור נתונים שונים תמורת כסף – למשל ארגון יכול למכור את המאפיינים הדמוגרפיים של לקוחותיו (כמובן במסגרת מה שהרגולציה והחקיקה בנושא צנעת הפרט מאפשרת להם). מה שהשתנה לעומת השנים הקודמות הוא הגידול העצום בכמויות הנתונים ובסוגיו. להלן מספר דוגמאות:

- **חברת Waze** – מציעה לרשויות מקומיות וגורמי תחבורה, נתונים על נפח התעבורה בתחומם (לאחר הורדת הפרטים המזהים). הרשויות המקומיות יכולות לעשות שימוש בנתונים אלה כדי לתכנן ולהרחיב את נתיבי התחבורה, לבצע אופטימיזציה של מחזורי הרמזורים וכד'.
- **חברת General Motors** – מוכרת לחברות ביטוח נתונים (לאחר סילוק פרטים מזהים) על הנסיעות המתבצעות ברכבים שלהם לפי הנתונים הדמוגרפיים של לקוחותיה, נתונים הנוצרים ע"י מערכת On Star המותקנת ברכביה. נתונים אלה יכולים לשמש את חברות הביטוח ללימוד היקפי נסיעות לפי חתכים שונים ולשפר את ניתוח הסיכונים שלהן.
- **חברת Strava** – מפתחת היישומון הפופולרי המשמש כרשת חברתית של עשרות מיליוני אתלטים העוסקים בריצה וברכיבה על אופניים למטרות ספורט.

המשתמשים מעלים ליישום את הפעילות הספורטיבית שלהם והיישום מאפשר למשתמשים לשתף מידע, לבחור מסלולים וכד'. המידע שהמשתמשים יוצרים כמו מסלולי ריצה או רכיבה, שעות הפעילות, משך הפעילות ועוד, הוא מידע רב ערך למתכנני ערים ורשויות מקומיות. החברה פיתחה יישום ייעודי, Strava Metro, וחתמה על הסכמים עם מספר עשרות עיריות גדולות ומעבירה להם את המידע (ללא פרטי המשתמשים כמובן) לשיפור תכנון העיר.

- **חברות הסלולר** – יכולות לספק נתונים לגבי מספר האנשים הנמצאים בזמן נתון באזור מסוים (ללא הפרטים המזהים כמובן). נתונים כאלה יכולים לשמש מסעדות או בעלי עסקים אחרים לתכנן את הביקוש לשירותיהם לפי שעות היממה ולתכנן מיקומים לעסקים חדשים. חברת פלאפון השיקה לא מזמן שירות חדש "פלאפון סמארט-דאטה". השירות יאפשר לארגונים לקבל מידע על מיקומם של אנשים, של מכוניות, על שעות בהן קיימת פעילות רבה באזור מסוים, על עומסי תנועה באזורים מסוימים ועוד. ניתוח של מידע כזה ע"י ארגונים יאפשר להם לקבל תובנות ולקבל החלטות חכמות.

ב. **מכירת נתונים מנותחים ללקוחות** – מודל עסקי נוסף שבמרכזו נתונים, מתבסס על מידע שהארגון מנתח ומוכר כשירות ללקוחותיו. לדוגמה חברת טלקום יכולה להציע כלי ללקוחותיה שיאפשר להם לנתח את השימושים במכשיר הנייד – כמה שיחות, מיהם עשרת היעדים הפופולריים לשיחות, ניתוח לפי שעות וכד'. חברות כרטיסי האשראי מספקות שירות ללקוחותיהן המנתח את הרגלי הצריכה שלהם תוך ניתוח סל הרכישות לפי סוגים שונים (מזון, תרבות, פנאי, ביגוד וכד'). חברה המוכרת מחשוב לביש (כמו שעונים או צמידים) המודדים פרמטרים כגון מרחק, מספר הצעדים, מסלולי הריצה, קצב איבוד הקלוריות ועוד, יכולה לספק חלק מהניתוחים ללקוחותיה בחינם ולהציע ניתוחים מתקדמים יותר תמורת עלות (מודל Freemium). ניתוחי מידע אלה מהווים דרך לספק ערך נוסף ללקוחות הארגון ולשמר אותם לאורך זמן, גם אם חלק מהניתוחים נמכרים ללא תוספת עלות. עם ההתקדמות בטרנספורמציה הדיגיטלית והפיכת מוצרים לשירותים, מגמה זו תתגבר. למשל חברת Rolls Royce המוכרת כיום מנועים כשירות במסגרת התכנית Power by the Hour, יכולה לבצע ניתוחים מתקדמים של המידע המשודר מכל הסנסורים במנועים ולהציע לחברות התעופה שירותי מידע מתקדמים תמורת עלות.

ג. **תיווך נתונים (Data Brokerage)** – קיימות חברות שאחד העסקים העיקריים שלהם הוא תיווך נתונים. למשל חברות כגון Bloomberg L.P או BDI אוספות נתונים כלכליים ממקורות שונים ומספקות ללקוחותיהן מידע מנותח על תיק ההשקעות שלהם או בהתאם לצרכים שלהם.

קיימים מודלים רבים נוספים המבוססים על נתונים. העידן הדיגיטלי מאופיין בהתפוצצות הנתונים, מאפשר לארגונים להתייחס לנתונים שהם צוברים כאל נכס מניב. הם יכולים לחשוב ולפתח מוצרים חדשניים מבוססי נתונים ולחולל זרם הכנסות חדש. בהמשך נתייחס לתפקיד חדש

- מנהל נתונים ראשי. אחד מתפקידיו החשובים הוא לאתר צרכים ורעיונות מבוססי נתונים ולסייע בפיתוח מוצרים חדשים אלה, כלומר לחולל הכנסות מנתונים (Monetizing the Data).

## 7.7 גישה לנתונים של ארגונים אחרים

באופן טבעי הנתונים שייכים לארגון שאוסף ומנהל אותם. אלה גם הנתונים הנמצאים ברמת איכות הגבוהה ביותר מאחר והארגון אחראי עליהם ומכיר אותם. משך שנים רבות התפיסה הרווחת הייתה לא לחלוק נתונים אלה עם ארגונים או אנשים אחרים מסיבות די ברורות – נכס ארגוני שיש להגן עליו מפני חשיפה, בגלל שיקולי פרטיות וצנעת הפרט, בגלל הרגישות העסקית או הביטחונית וכד'. ארגונים נטו לשמור על נתוניהם מאחורי חומות גבוהות. בשנים האחרונות אנו עדים לשינוי משמעותי בתחום זה. נציין שתי תופעות חדשות יחסית:

א. פתיחת וחשיפת מאגרי נתונים של ארגונים ממשלתיים וציבוריים, תופעה הנקראת Open Data.

ב. פתיחת וחשיפת נתונים של ארגונים עסקיים באמצעות ממשקי תכנות (API - Application Programming Interfaces) המאפשרים לארגונים לשלוף ולשתף ביניהם יישומים ונתונים.

בהקשר זה, נדגיש שאיכות הנתונים המגיעים מצד ג' משתנה בהתאם למקור הנתונים – ישנם מקורות ברורים ומזוהים בעלי רמת איכות גבוהה וישנם מקורות ברמת איכות לא ברורה ולכן יש לנקוט לגביהם זהירות רבה יותר. נפרט שתי מגמות חשובות אלה.

### 7.7.1 פתיחת מאגרי נתונים ממשלתיים – Data Government Open

הרעיון של פתיחת וחשיפת מאגרי מידע של ארגונים לשימוש ע"י גורמים מחוץ לארגון למטרות שונות וללא מגבלות על זכויות יוצרים, הוא רעיון שקיים כבר שנים רבות. ניתן לראות רעיון זה כחלק מהמגמה שצברה פופולריות ומקדמת את אי הגבלת השימוש (Open) בתוכנה (Open Source). מגמה זו זכתה לחיזוק משמעותי בשנת 2013 כאשר הממשל של הנשיא אובמה פרסם את המדיניות ואוסף של הנחיות לכל הרשויות הפדרליות בארה"ב בדבר פרסום ופתיחת מאגרי מידע ממשלתיים וציבוריים. מגמה זו יצאה מתוך ההנחה שהמידע המנוהל ע"י הרשויות הממשלתיות והמקומיות שייך בסופו של דבר לאזרחים ולמשלמי המיסים. הממשלות מחזיקות מאגרים אלה בשם האזרחים שלהן ולכן עליהן לחלוק אותם ולחשוף את מה שניתן לחשוף כדי לעודד פתיחות ושקיפות וכן כדי לעודד יצירת חדשנות, צמיחה כלכלית וערך חדש. הרעיון של חשיפת מאגרי נתונים ממשלתיים התפשט במהירות למדינות אחרות וכיום קיימים מיזמי Open Data במספר רב של מדינות, כולל בישראל. המדינות המובילות בתחום זה הן ארה"ב עם אתר ה Data.gov והבריטים עם Data.gov.uk. גם בישראל פותח אתר Data.gov.il המכיל עשרות מאגרים ממשלתיים הפתוחים וזמינים לציבור ולארגונים.

ההגדרה של Open Data היא שאלה נתונים שניתן להשתמש בהם באופן חופשי ומותר לכל אחד להפיץ אותם בהגבלות מינימליות (לעיתים רק של אזכור המקור). הרעיון העומד מאחורי מגמה זו הוא שמאגרי הנתונים הם נכס שניתן ורצוי למנף למטרות שונות – לתועלת הציבור וגם לתועלות

עסקיות. רשימת המאגרים הממשלתיים שניתן לפתוח היא עצומה – מאגרי נתונים של NASA המתקבלים מחלליות המחקר, מאגרי מפות, מאגרי עם מידע רפואי (כמו למשל של הגנום האנושי), מאגרי נתונים על איכות הסביבה, נתונים מטאורולוגיים ועוד. למשל פתיחת מאגר הנתונים של השירות המטאורולוגי מאפשר לכל אחד לשלוח את נתוני מזג האוויר שהצטברו משך שנים רבות ולעשות בהם שימושים שונים ומגוונים. ארגונים העוסקים בחקלאות ובתיירות יכולים לעשות שימוש בנתוני המשקעים בחודשי השנה השונים באזורים בארץ ולפתח יישומים הממליצים לחקלאים מתי לזרוע ולקצור יבולים מסוימים או יישומים תיירותיים המציגים לתיירים את מזג האוויר באתרים בהם הם מבקשים לבקר.

במהלך השנים אף התפתחה פלטפורמת תוכנה (המבוססת על קוד פתוח כמובן) בשם CKAN המאפשרת לארגונים וממשלות לפתוח מאגרים במהירות וביעילות ומאפשרת איתור נוח של המאגרים ע"י הגורמים המעוניינים. מערכת זו אף מנהלת את ה Meta Data כדי לאפשר את הבנת מבנה הנתונים ואת הגישה אליהם. גם ממשלת ישראל עושה שימוש ב CKAN כבסיס ל Data.gov.il.

פתיחת מאגרים ממשלתיים אפשרה לחברות שונות לפתח יישומים חדשניים. להלן מספר דוגמאות:

א. **חברת Moovit** שהחלה את פעילותה בארץ עם פתיחת המאגרים של משרד התחבורה על

קווי האוטובוסים והתחבורה הציבורית, כולל כל התחנות להורדת והעלאת נוסעים. על בסיס מידע זה החברה פיתחה יישום המאפשר לכל אחד לתכנן את מסלול הנסיעה שלו באמצעות התחבורה הציבורית ממקום כלשהו ליעד כלשהו. לשימוש בסיסי זה, החברה הוסיפה מגוון רחב וגדול של אפשרויות ופונקציות תוך הצגת מפות של העיר ומסלולי הנסיעה, רשימת התחנות הקרובות, יכולת לשלב בין אמצעי תחבורה שונים. החברה מספקת כיום יישום זה ביותר מ 1,200 ערים ברחבי העולם.

ב. **חברת מדלן** הישראלית פיתחה יישום השואב חלק מהמידע ממאגרי מידע ממשלתיים כדי

להציג למשתמש נתונים על עסקאות המכירה/קניה ומחירי הדירות בסביבה (הנשאבים ממאגרים של רשות המיסים), על בתי ספר וגני ילדים בשכונה (מתוך מאגרי משרד החינוך) והמרחק שלהם מהדירה שהמשתמש מתעניין בה, על מיקום האנטנות הסלולריות בשכונה (מתוך מאגרים של המשרד להגנת הסביבה), על רמת הפשיעה (מתוך מאגרי המשטרה), על מיקום חנויות הסופרמרקט באזור ועוד.

משרד הבריאות בארץ יצא לדרך, בשנת 2017, עם פרויקט "שיתוף מידע רפואי" להקמת מאגר לאומי של נתונים רפואיים שיאוחסן בענן של Amazon, כמובן לאחר שכל פרטי הזיהוי מותממים (עוברים תהליך חכם של אנונימיזציה וסילוק מידע מזהה, תהליך שאינו פוגם בקשרים הלוגיים והתוכן הסטטיסטי הטמון בנתונים). בכוונת משרד הבריאות לאפשר לגופי מחקר של קופות החולים, אקדמיה, חברות לייצור תרופות ועוד, לבצע מחקרים על בסיס הנתונים הרפואי הלאומי וכל זאת מתוך מטרה לקדם את איכות הרפואה בארץ ושיפור הבריאות של הציבור.

קיימות דוגמאות רבות נוספות לחברות מסחריות, בעולם ובארץ, שידעו לנצל את חשיפת המאגרים הממשלתיים ולפתח על בסיסם יישומים חדשניים שהביאו ערך רב לציבור.



## 7.7.2 פתיחת מאגרי נתונים של ארגונים עסקיים

התופעה של חשיפת מאגרי נתונים התפתחה גם בקרב ארגונים עסקיים, כמובן מסיבות אחרות. ארגונים עסקיים עושים זאת על מנת לפתח את עסקיהם ולכן ארגונים אלה בדרך כלל עושים זאת תמורת תשלום. יש גם ארגונים הרואים בפתיחת מאגרי הנתונים שלהם הזדמנות לקבוע עובדות וסטנדרטים ולייצר הכנסות באופן ישיר או עקיף. הדרך לגשת לנתונים היא ע"י ממשק תכנות (API), השיטה הרווחת, או ע"י העברות קבצים, שיטה פחות פופולרית. להבדיל ממאגרי הנתונים הממשלתיים שמסופקים לגורמים המעוניינים ללא תמורה, הארגונים העסקיים מפתחים מודלים עסקיים שונים ובחלק מהמקרים גובים תשלום בגין הגישה לנתונים שלהם.

מודל עסקי מעניין שארגונים מפתחים נקרא **החלפת נתונים** (Data Swapping). במקום לגבות כסף בגין השימוש בנתונים, הארגונים מחליפים ביניהן נתונים, נתונים בעלי ערך לשני הארגונים. להלן מספר דוגמאות:

א. **חברת Waze** – החברה מספקת ערכת תכנות (SDK) המאפשרת גישה למאגרי הנתונים שלה ומאפשרת למפתחי יישומי צד ג' לשלב מידע של Waze בתוך היישומים שלהם, כולל נתונים, מפות וחישובים. באמצעות ערכת התכנות ניתן לקבל מידע על זמן הגעה צפוי ליעד כלשהו (ETA – Expected Time of Arrival) בהתבסס על נתוני עומס התנועה כפי שידועים ומנוהלים ע"י Waze. החברה מאפשרת גם גישה לנתוני ניווט וחישוב המסלול המהיר ביותר בין שתי נקודות ולשלב את התוצאות בתוך יישום צד ג'. זו למעשה גישה לשכבה אפליקטיבית ולא ישירות לנתונים.

ב. **חברת Waze** – החברה מאפשרת לרשויות מקומיות (לדוגמא עם עיריית Rio de Janeiro), לקבל נתוני מצב תנועה, תאונות, מהירות התנועה של מכוניות ועוד מתוך מאגרי Waze ובסיועם מסוגלות לנהל בצורה טובה יותר את התנועה בעיר (Traffic Management). בתמורה לנתונים אלה, העירייה מספקת ל Waze נתונים כגון צילומים מהמצלמות הפזורות בעיר, נתונים מרמזורים וסנסורים שונים המותקנים על הכבישים ומופעלים ע"י העירייה.

ג. **חברת Moovit** – לחברה יש מודל עסקי דומה המספק לרשות מקומית מידע על היקפי ויעדי הנסיעות ובתמורה הרשות המקומית מספקת ל Moovit מידע בזמן אמת על אוטובוסים ורכבות והיא יכולה גם לספק התראות למשתמשי Moovit על עומסי תנועה, עיכובים בתנועת האוטובוסים והרכבות וכד'.

ד. **חברת Strava** – החברה בנתה יישום מובייל של רשת חברתית של ספורטאים במקצועות שונים (למשל רוכבי אופניים, ריצה וכד') וממנפת את מאגר הנתונים שלה להכנסות חדשות. החברה גובה מרשויות מקומיות 80 סנט עבור כל רוכב אופניים רשום ומאפשרת לעיריות ללמוד את הרגלי הרכיבה בעיר. כמובן שהנתונים המועברים לרשויות מקומיות הם ללא נתוני זיהוי.

ה. **חברת Google** – החברה מאפשרת גישה באמצעות ממשק תכנות (API) לנתוני Google Maps ולשלב את המפות שלה ביישומי צד ג'.

ו. **חברת IBM** – החברה מספקת ממשק תכנות למערכת הקוגניטיבית שלה, Watson הפועלת בענן, ומאפשרת למפתחי יישומים לנצל את העוצמה והחוכמה של Watson. דוגמא יפה לשימוש בממשק תכנות זה עושה חברת Cognitoy שפיתחה צעצוע בצורת דינוזאור קטן וחמוד ובאמצעות ממשק התכנות ל Watson מאפשרת לילדים ללחוץ על כפתור ולשאול שאלות בנושאים רבים ומגוונים. צעצוע שעלותו כ 150 דולר מאפשר גישה ומשתמש באחת המערכות הקוגניטיביות העוצמתיות ביותר הקיימות כיום!!

כפי שניתן לראות מהדוגמאות הנ"ל, הנתונים הנאגרים ע"י הארגון הם לא רק נכס שנועד לשימוש פנימי אלא נכס שיכול להביא תועלות רבות לארגונים אחרים. ארגונים יכולים לפתח ולמנף את הנתונים שהם אוגרים ומנהלים ולחולל זרם חדש של הכנסות, כלומר להפוך את הנתונים לכסף (Monetizing the Data) ובמקרה של ארגונים ציבוריים להביא תועלת לציבור.

## 7.8 סיכום: ידע זה כוח

בפרק זה סקרנו את הפוטנציאל העצום הטמון בנתונים בעידן הדיגיטלי. האנלוגיה של משאב הנתונים לנפט מסייעת בהבנת היכולות להפיק ערך עצום מנתונים שעוברים עיבוד, בדומה לערך הנובע מהמוצרים המופקים מהנפט לאחר זיקוקו ועיבודו. העידן הדיגיטלי מאופיין, בין היתר, בהתפוצצות הנתונים והגידול האקספוננציאלי בכמות הנתונים, ומאפשר לארגונים לפתח תובנות חדשות על בסיס אנליטיקה מתקדמת ושילוב הנתונים כחלק אינטגרלי מהמודל העסקי שלהם. לתופעה זו קרא פרופ' Davenport בשם אנליטיקה 3.0.

במקביל לגידול בחשיבות האנליטיקה, אנו עדים להתפתחות המהירה של למידת המכונות, מערכות המתבססות על היכולת ללמוד מתוך כמויות גדולות של נתונים. יותר ויותר ארגונים מבינים את הפוטנציאל הטמון בנתונים ומפתחים מודלים עסקיים חדשניים מבוססי נתונים. גופים ממשלתיים פותחים את מאגרי הנתונים שלהם כדי לעודד חדשנות ויצירת ערך חדש. אנו עדים לארגונים הממנים Chief Data Officer שמטרתו לנהל ולקדם את הנכס החדש והחשוב הזה. ניתן להצביע על ארגונים מתחום התעופה, הבנקאות והביטוח, כרטיסי אשראי, ארגוני קמעונאות ועוד, שהבינו את החשיבות של משאב הנתונים ומינו CDO כבעל תפקיד בכיר לזיהוי ומיצוי פוטנציאל חדש זה.

ארגונים שילמדו לרתום את הנתונים הנוצרים בתוך הארגון ולשלב אותם עם נתונים חיצוניים, עם מערכות של למידת מכונות, אינטליגנציה מלאכותית ואנליטיקה מתקדמת, ויצליחו לפתח מודלים עסקיים חדשניים המבוססים על נתונים, הם הארגונים שיצליחו לייצר יתרון תחרותי בעידן הדיגיטלי.

לסיכום: ידע שמנוהל נכון (איסוף, אחזור, ניצול, עדכון והפצה וכו') זה כוח.