

סקירה ותובנות ממפגש Data as an Asset

במפגש תקפנו את נושא הערך שניתן לייצר מנתונים וההסתכלות על נתונים כ"נכס". התחושה היא שזה אכן הדיבור בקרב ארגונים, סממאות כמו: data is the new oil, נתונים מהווים נכס חשוב לצמיחה ואף הישרדות הארגון בשנים הקרובות וכו'. אבל במקרים רבים יש תחושה שאלה אכן רק סיסמאות, ושקשה מאוד להוכיח אותן ולהצמיד להן מספרים. כשמנהלי דאטה מבקשים יותר תקציבים, הם נדרשים להראות ערך מדיד ולהצדיק את העלויות הגבוהות שנלוות לתשתיות דאטה, ניהול נתונים, וכ"א שדורש משכורות גבוהות בעולם ה data science.

איך עוברים מסימאות לערך מדיד? איך מסתכלים על נתונים כנכס שאינו שונה משאר נכסי הארגון? איך מכמתים את השווי של הנתונים עבור הארגון? איך מגדילים באופן שיטתי את הערך? אלה הנושאים שעמדו בלב המפגש Data as an Asset, מובאים כאן עיקרי התובנות שעלו מהמפגש.

המניע העיקרי כיום: פחד

בדומה לטרנספורמציה הדיגיטלית, טרנספורמציה הדאטה בארגונים מונעת כיום מפחד, הרבה יותר מכל סיבה אחרת. ארגונים פיננסיים כנראה הראשונים להרגיש זאת (חשש מאותם פינטקים מונעי-AI שעשויים לחולל תהפוכות של ממש בתעשייה) אך גם Retailers, גופי בריאות ועוד. מה יקרה אם לא נהיה ערוכים, ויגיע ארגון AI-driven לתעשייה שלנו ויעשה את מה שאנחנו עושים, רק יותר מהר, יותר ממוקד והרבה יותר יעיל? מה יקרה אם לא נתחיל להיערך כיום?

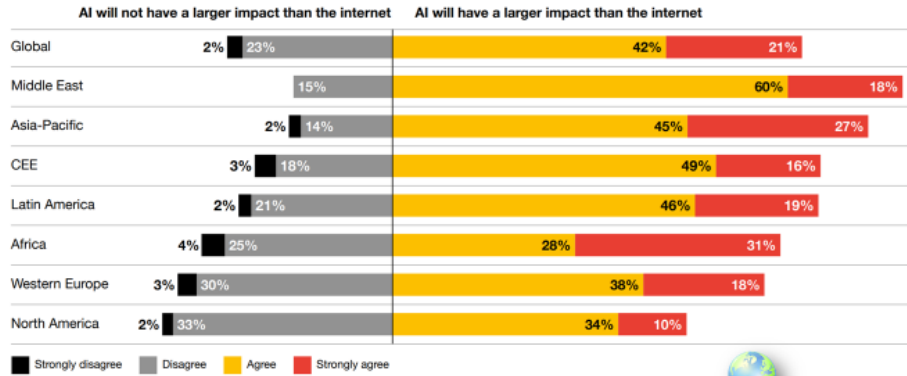
יש כמובן גם תועלות לעליית חשיבות הנושא בקרה ההנהלה הבכירה. ישנם ארגונים ש"נמנמו" בנושא הנתונים כבר שנים וכעת חשים שהם חייבים להיכנס לנושא ולמצות יותר פוטנציאל ממנו. כתוצאה מאותה תחושת דחיפות, מוקמות יחידות דאטה, מאוישים תפקידי Data Officers, מוקמות מעבדות נתונים, קמות יוזמות חדשניות סביב אנליטיקה מתקדמת, מגייסים ומכשירים כ"א בעולם ה Data Science. גם רגולציות סביב עולם הנתונים מסייעות בהנעת התחום קדימה (ולעתים גם עצירתו, באופן אירוני).

על פי התחזית של פורסטר, סטארטאפים מונעי-AI יצליחו לחטוף 1.2 טריליון \$ ממתחריהם עד 2020. כתוצאה מהעננה המרחפת הזו (פחד?), קרוב ל 70% מארגונים כבר החלו במרוץ ההצטיידות ל-AI. במפגש אף דובר על כך שחייבים להנמיך ציפיות מנושא ה AI... או לכל הפחות להפוך אותו לפחות מיסטי.

במחקר מצוין של PWC, נשאלו CEOs כיצד הם משווים את האפקט של תחום ה AI על התעשייה שלהם בהשוואה שלאינטרנט בזמנו הייתה. 63% סברו של AI השפעה גדולה יותר! זה נתון די מדהים בהתחשב בזה שלרוב, בכלל לא ברור מה זה AI והכל מאוד מעורפל, אבל יש תחושה של פחד ודחיפות.

The race for AI

63% of CEOs think AI will have a greater impact than the internet



Source: PwC, 22nd Annual Global CEO Survey
Base: AI respondents (2019-1,178)



Source: PwC

Copyright@STKI_2019 Do not remove source or attribution from any slide or graph

עוד כמה נתונים מעניינים של חברת המחקר NewVantage מצביעים בדיוק על אותה אווירה:

54% מהארגונים שנכללו במחקר ציינו כי חוסר ה"אגיליות" שלהם בעולם הנתונים מהווה את האיום התחרותי הגדול ביותר שלהם, ו 80% מהם חוששים מ disruption של חברות מונעות-נתונים.

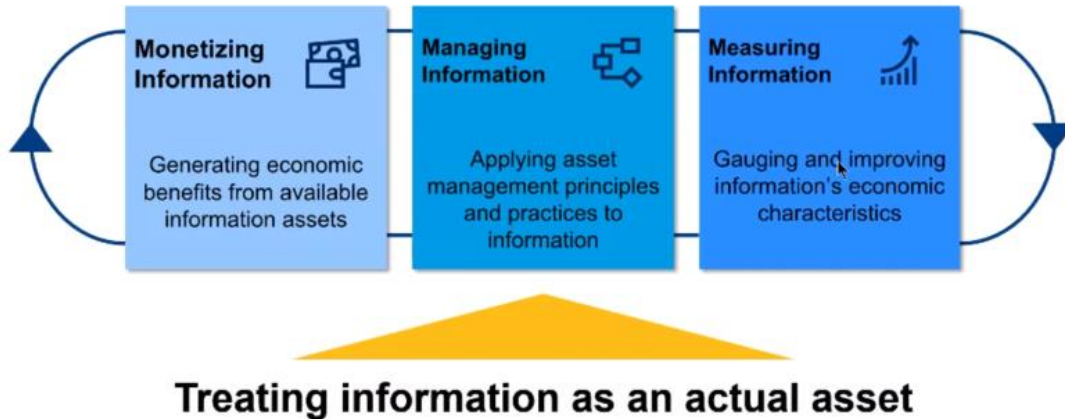
האם אפשרי בכלל למדוד ערך של נתונים?

Infonomics: מונח עליו מדברים יותר ויותר לאחרונה, עוסק בצד הכמותי של עולם הנתונים. ספר מצוין של Doug Laney בשם Infonomics ממפה את הרכיבים השונים של כלכלת הנתונים:

מוניטיזציה, ניהול ומדידה. אף אחד מהם לא יכול להתקיים ללא האחרים.

- **מוניטיזציה:** היכולת לייצר תועלות כלכליות מנתונים (בין אם באופן ישיר – כמו לדוגמה מכירת המידע לגוף צד 3, או באופן עקיף – שיפור איכות המוצר וכד')
- **ניהול:** ניהול הנתונים כנכס, ואף השאלת שיטות ניהול נכסים ותיקות לטובת ניהול נתונים
- **מדידה:** חשוב מאוד! מדידה מתמדת של העלות, ה"שווי" והערך. מדידה יכולה להתבצע על פני מספר פרמטרים ונקודות מבט: שווי הנתונים עבור השוק, עלות אחזקת הנתונים, ערך הנתונים לטובת שיפור אופרציות קיימות ועוד ועוד.

What is Infonomics?



Source: Gartner

למרות שארגונים (עדיין) לא מציינים בספרים את ערך הנתונים שלהם כנכס חשוב, סביר שנגיע לנקודה זו בקרוב, ואחד הסימנים לכך היא ההסתכלות של חברות השקעות וה Valuation הגבוה (בהרבה!) שהן מעניקות לחברות מונעות נתונים. ערכה של חברה מונעת נתונים כפי 2 יותר גבוה מזה של חברה שאינה מונעת נתונים, וחברות שמוצריהן הנם 100% מוצרי דאטה מגיעות לפי 3 יותר (רק תחשבו על הערך של חברות כמו אמאזון, גוגל, נטפליקס וכו').

מהם החסמים?

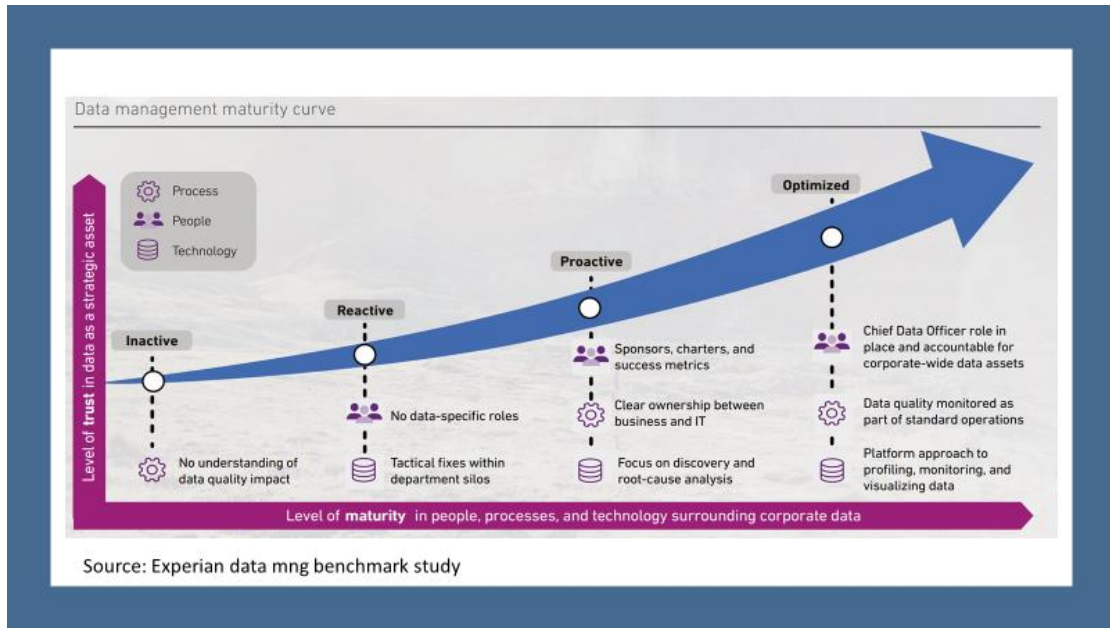
במפגשים קודמים דובר הרבה על הנושא **התרבותי** כחסם עיקרי (התנגדות לשינוי, רמת Data Literacy – אוריינות נתונים נמוכה, ולעתים אף פחד/רתיעה מנתונים). בנוסף לאלה, חסמים נוספים כוללים **חסור התאמת תהליכים** להיות כאלה שהנם מונעי נתונים (מה שמשאיר הרבה יוזמות פיתוח מודלים אנליטיים "על המדף", מכיוון שהיחידה העסקית אינה מוכנה לשנות את תהליכי העבודה הקיימים שלה ולהישען על אותן תובנות – ראינו זאת קורה שוב ושוב). במשולש הידוע של people, technology, process, החסם ללא ספק אינו טכנולוגי, אלא הרבה יותר קשור לאנשים ותהליכים.

<u>Principle Challenge to Becoming Data-Driven</u>	<u>2018</u>	<u>2019</u>
People	48.5%	62.5%
Process	32.4%	30.0%
Technology	19.1%	7.5%

Source: New Vantage

חסם נוסף הוא איכות הנתונים ורמת האמון הנמוכה בהם. על פי מחקר של Experian, 33% מהארגונים ציינו שנתוני חוויית הלקוח שברשותם אינם מדויקים. מחקרים אחרים מדברים על ממוצע של כ 20% בכלל הארגונים של מידע שהוא בלתי שמיש בעליל (Dirty Data).

איפה אתם נמצאים על עקומת הבשלות של ניהול הנתונים? (פסיביים, ריאקטיביים, פרואקטיביים)?



Experian data management benchmark study: מקור

כיצד נוודא שאנחנו תוקפים את הבעיות הנכונות?

אחד התחומים הכאובים ביותר הוא מציאת ה use cases הנכונים. איך מחלצים את הרעיונות הנכונים מהיחידות העסקיות? איך מתעדפים? איך מוודאים שאותן יוזמות אנליטיות אכן יתממשו ושמישהו אכן יעשה איתן משהו? יוזמות אנליטיות רבות בסופו של דבר אינן מתממשות (אפילו אם המודל מדויק) כפי ששמענו מארגונים רבים. אז איך מפצחים את זה?

העלינו שאלה זו לשלב הדיון הפתוח במפגש. להלן הפרספקטיבה שסופקה על ידי המשתתפים במפגש, מנהלי תחום דאטה/אנליטיקה/אנליזה בארגוני Enterprise גדולים בישראל:

אחד מנציגי הארגונים שכבר עשו כברת דרך ארוכה בנושא הדאטה ואנליטיקה, חלק מניסיונו והמליץ תמיד לצאת מה"תוצאה" והרעיון – מה רוצים להשיג? נשמע טריוויאלי וברור, אולם ארגונים רבים נוטים להסתכל על הדאטה הקיים ולשאל "מה ניתן לעשות". התועלת בלהתחיל מהבעיה עצמה ומה רוצים להשיג גם מאפשרת מימוש יותר קל – אם מראש אני יודעת שיש לי פרטנר עסקי טוב, מחויב למשימה, שמוכן גם לשנות תהליכים קיימים ביחידה העסקית על מנת להיסמך על המודל, הסיכוי למימוש יעלה משמעותית, לעומת רעיון מגניב שצמח ביחידת הדאטה פשוט בגלל שהוא אפשרי מבחינת הנתונים הקיימים.

אחד הארגונים סיפר על תהליך שבו הלך ואיתר use cases אנליטיים מהתעשייה שלו, עברו על קייסים מהעולם, עברו בין היחידות העסקיות השונות וחיפשו את הפרטנר המתאים. כשאיתרו את אותו פרטנר, עבדו איתו כל הזמן ביחד כאשר הם מראים לו את ההתקדמות על מנת שיכוון אותם וייתן להם תחושה שהם חלק מהתהליך. בכוונה בחרו 3 קייסים מעולמות תוכן שונים, כאשר המטרה הייתה לגרות את הדמיון ולהראות מה ניתן לעשות (טעימות).

מודלים | use cases שתוארו לדוגמה:

- שיפור מודל השימור
- Next best offer
- מודל ערך לקוח שתומך בגופי השירות
- Customer Lifetime Value



- יצירת "תכניות התערבות" שלא מחכות שמשוהו יקרה אלא משפיעות עליו (לדוגמה, רפואה מונעת בעולם הבריאות)
- נטישת עובדים – זיהוי דפוסים בדאטה שמצביעים על איתות נטישה
- חיזוי של lead time בשרשרת האספקה

האם ארגונים מודדים ערך?

מסתבר שמעט מאוד, לפחות על פי הדיון שקיימנו. רק למספר מועט מאוד של ארגונים שנכחו במפגש הייתה מתודולוגיה סדורה למדידת הערך המתקבל מיוזמות אנליטיות.

מילון נתונים – האם זה תנאי מקדים?

בדיון עלה שוב ושוב נושא מילון הנתונים, כנושא חשוב שיש להיערך אליו מחד, אולם נושא שעדיין מאוד קשה לעיכול ועם הרבה בעיות סביב המימוש מצד שני.

לעתים הדיון הורחב אל מעבר לניהול מילון נתונים, אלא לאסטרטגיה היותר כללית של Data Governance. מבין כלל המשתתפים בדיון, אף אחד מהם עדיין לא עוסק בפועל ב Data Governance אולם הייתה הסכמה כללית על החשיבות, ואחד הארגונים סיפר על התארגנות ומינוי מנהל/ת לנושא תחת יחידת ה CDO.

ארגונים גם סיפרו על ה"סבב הקודם" כבר לפני כמה שנים ביישום מילון מונחים ארגוני. בשלב זה של בשלות הטכנולוגיה, הכלים הסתמכו לחלוטין על יכולות ידניות ללא אוטומציה כלל (כיום מדובר על כ-50% יכולות של discovery עצמאי/אוטומטי על ידי הטכנולוגיה, והשאר – עדיין ידני), מה שהביא למצב שבו המילון כל הזמן דורש תחזוקה. זה דרש מאמץ ממנתחי מערכות ובצד העסקי לתחזק, ובאופן טבעי הרלוונטיות שלו הלכה וירדה. בסופו של דבר זה לא עבד ואנשים לא עשו בו שימוש כי זה לא היה מתוחזק.

כיום, ארגונים שנכחו באותו סבב קודם מספרים על הנמכת ציפיות וצמצום ה SCOPE. הם כבר לא מנסים למפות הכל, אלא לוקחים ישויות בעייתיות / נושאים כואבים, ומסדרים ומסדירים אותם כמונח, מגדירים כיצד ערך זה מחושב.

לחלק מהארגונים יותר קל להתחיל משלב ה-LINEAGE כשלב ראשון – היכולת להסביר "מה עבר על הנתון" עד שהגיע למצבו הנוכחי, כלומר איך הגיעו לדוח. אחד הארגונים סיפר על יישום של Octapi ל data lineage לטובת תיקון באגים והבנה על מה שינוי של נתון מסוים ישפיע (כלומר – פרויקט IT).

השלב הבא יהיה לקחת את המוצר יותר לכיוון האנליסטים ששואלים מה הלוגיקה ומה המקור.

משתתפי המפגש סיפרו על יישומי Informatica data catalog וכן Octapi, כמו כן - Atlas – לסביבת ה data lake פפציפית.

בסופו של דבר המטרה היא לאפשר לאנליסטים את היכולת להגיע לנתון הנכון, ושתהיה לנו היכולת לעשות trace back לאיך הגענו לנתון ובסופו של דבר לתמוך במעבר לתצורת עבודה של Self Service. אצל אחד הארגונים הטריגר היה לאפשר להנגיש את הדאטה החוצה לסטארטאפים.

במפגש עלתה שוב ההתלבטות האינסופית לגבי היכן צריך לשבת גוף דאטה בארגון. 2 ארגונים סיפרו על מיקום של יחידת הדאטה תחת השייוק, ארגונים אחרים סיפרו לרוב על יחידת דאטה ייעודית שהוקמה/תקום בקרוב. אין המלצה או תובנה חד משמעית, אחד המשתתפים סיפר על יוזמה עתידית להקים גוף משותף של IT + דאטה "עסקי" כדי לגשר על החסרונות שבכל אחת מהאופציות.