



מושגי יסוד מסביבת דיגיטל - ודאטה

נתונים

נתונים - DATA

פריטי תוכן גולמיים: מספרים אותיות וסמלים שניתנים לקידוד, שמירה או קלט לתוך מערכת מידע.

מידע גולמי

נתוני מקור שלא עברו עיבוד, כדוגמת שעות עבודה/שכר לשעה.

מידע מעובד

מתייחס למניפולציה של מידע דיגיטלי על ידי אלגוריתמים, כדוגמת חישוב משכורת לעובד על בסיס המידע הגולמי של שעות העבודה.

מידע מובנה ומידע לא מובנה

נתונים מובנים STRUCTURED DATA הם נתונים שיש להם מבנה מוגדר היטב, תואם מודל נתונים אשר עוקב אחרי סדק מתמשך, וניתן לגשת אליהם בקלות על ידי בני אדם ותוכנות מחשב. למשל, טבלת נתונים באקסל.

מידע לא-מובנה UNSTRUCTURED DATA הם מידע שאינו מסודר על ידי מודל נתונים או סכמה מוגדרת מראש, ולכן לא ניתן לאחסן אותו במסד נתונים. למשל, הודעות דואר אלקטרוני, הקלטות קוליות, סרטונים, תמונות, דפי אינטרנט.





בעלי התפקידים המרכזיים בתחום מערכות מידע

מנמ"ר - CHIEF INFORMATION OFFICER - CIO

מנהל טכנולוגיות דיגיטליות ומערכות מידע בארגון. למנמ"ר אחריות כוללת לכלל מערכות המידע בארגון, לשירותי טכנולוגיית המידע למשתמשים במערכות המידע של הארגון, לשירותי טכנולוגיית המידע למשתמשים במערכות המידע של הארגון ולתהליכי הפיתוח, ההקמה והתחזוקה של מערכות התשתית.

מנהלת טכנולוגיה ראשית - CHIEF TECHNOLOGY OFFICER - CTO

תפקיד בדרג הניהולי הבכיר של חברה או ארגון. האדם הנושא במשרה זו מטפל בנושאים המדעיים והטכנולוגיים שבתוך הארגון.

תחומי הטכנולוגיה של CTO נקבעים במידה רבה בהתאם לגודל ומבנה של החברה ולתחומי הפעילות שלה. בחלק מהארגונים המנמ"ר ממלא את התפקיד הזה, ובאחרים מדובר בתפקיד העומד בפני עצמו.

ראש תחום תשתיות מחשב/יישומים

האחראי על היישומים והתשתיות להפעלת טכנולוגיות המידע שבארגון. תחומי אחריותו הם רכיבי החומרה והתוכנה, מערכות התקשורת, השרתים, תחנות הקצה אבטחת המידע ותמיכת מערכות המחשוב בעובדי הקצה, תוך דאגה לזמינותם ולתפעולם.

ראשת תחום אבטחת המידע - CIEF INFORMATION SECURITY OFFICER- CISO

האחראית בארגון להתוויית תוכנית אבטחת המידע הארגונית השנתית וליישומה. תוכנית אבטחת מידע מסודרת תעזור לארגון להתכונן לאיומים עתידיים בתחומי המידע והסייבר ולדעת איך להתמודד איתם בצורה היעילה ביותר.

מנהלת אבטחת המידע אחראית להסדרת פתרונות וטכנולוגיות מתאימים לצורכי הארגון, לניהול הספקים והמיישמים של מערכי הגנת הסייבר, ולהטמעת תקנים חדשים ורגולציות נדרשות. מנהלת אבטחת המידע אחראית לתגובה לאירועי אבטחת מידע וסייבר וכן אחראית לרכז את כל המשאבים על מנת לתת מענה כולל ויעיל.





DATABASE ADMINISTRATOR- DBA

המתכנתת האחראית למסד הנתונים, ובתוך כך לאפיון מודל הנתונים בחברה באופן שיאפשר להסתנכרן עם כל המערכות.

כמו כן עליה למצוא פתרונות להתמודדות טכנית עם עומסי מידע. בארגונים שונים קיימת חלוקת בין **DBA תשתיתי**, שתפקידה הוא לוודא שמסד הנתונים מתפקד יעיל ומדויק, לבין **DBA אפליקטיבי** המתמחה בתוכנות לניהול מסדי הנתונים.

CHIEF DATA OFFICER - CDO

מנהל נתונים ראשי - האחראי על ניהול המידע הארגוני ובסיסי הנתונים לצורך קידום מטרות הארגון. תפקידו של מנהל הנתונים הראשי לקבוע את דרכי ניהולו של המידע הארגוני, טיובו, שיתופו/מידורו, וכן להבטיח את מיצוי השימוש בו לצורך קבלת החלטות וקביעת מדיניות הארגון.

מנהלת התחום ביחידה המקצועית

מנהלת התחום או מומחית מידע המכירה את הנתונים המבוקשים ודרכה ניתן לקבל את הפילוחים והמידע הנדרש.

לעיתים הדרך היעילה לקבלת המידע תהיה באמצעות מנהלת התחום שעליו מתבצעת הביקורת, ולא באמצעות המנמ"ר.

היכן נחפש את המידע?

חומרה - HARDWARE

אוסף כל הרכיבים הפיזיים במחשב או בהתקן אלקטרוני אחר, כמו טלפון סלולרי או ציוד תקשורת. זאת להבדיל מהידע שמחשב או ההתקן מכיל.

מחשב אישי - PERSONAL COMPUTER - מידע ומסמכים נשמרים בזיכרון הפנימי של המחשב, גם במקרים שבהם יש בארגון שרת מרכזי. להלן הרכיבים המרכזיים במחשב האישי:

כוננים - אלו הם כל רכיבי הזיכרון במחשב אשר אינם הזיכרון הפנימי. הדיסק הקשיח נחשב אף הוא לכונן, ואליו מצטרפים במחשבים אישיים רבים כוננים נוספים, הנבדלים מהדיסק הקשיח בכך שהנתונים נשמרים בהם על גבי מדיות חיצוניות, אותן ניתן להוציא מהכונן.





דיסק קשיח - גם הדיסק הוא רכיב זיכרון, אלא שבניגוד לזיכרון המחשב (המכונה "זיכרון פנימי"), על גבי הדיסק נשמרים הנתונים גם לאחר כיבוי המחשב.

שרת - SERVER - שרת הוא מחשב המריץ תוכנה (אחת או יותר), ומספק שירותים למחשבים אחרים. כדוגמת שרת קבצים, שרת דואר אלקטרוני, שרת בסיס נתונים, שרת הדפסה.

רשת מחשבים - COMPUTER NETWORK - רשת טלקומוניקציה המאפשרת העברת נתונים. מכשירים המחוברים לרשת מחשבים מעבירים זה לזה נתונים על גבי קשרים או חיבורים להעברת נתונים. רשתות מחשבים נפוצות כיום, והמוכרת ביותר היא רשת האינטרנט.

כונן רשת - כונן הרשת מאפשר לשתף מסמכים ברשת המקומית עם משתמשים נוספים בארגון/באותה רשת. הרשת יכולה להיות רשת מחשבים מקומית או רשת מקומית בענן. המשתמש יכול לקרוא את הקבצים ואף לבצע בהם שינויים בהתאם להרשאות שהוקצו עבורו.

לדוגמה, ברשת המשרדית של מבקר המדינה קיימים מספר כונני רשת:

M - כונן שיתופי פתוח לכל העובדים.

H - כונן אישי של העובד.

F - כונן אגפי.

בכל ארגון מוגדרים כונני רשת אחרים בהתאם למאפיינים ולהרשאות הנדרשות.

מחשוב ענן - מחישוב ענן או מחשוב עננים הוא מודל המאפשר גישה נוחה, על גבי רשת תקשורת ועל פי דרישה, למאגר שיתופי של משאבי מחשב.

הדבר נעשה בדרך כלל על ידי ריכוז המשאבים בחוות שרתים המקושרת לרשת תקשורת והקצאתם למשתמשים באופן וירטואלי. השירות למשתמשים יכול להינתן באמצעות ענן מקומי הנמצא ברשת הארגונית, או באופן מרוחק במרכז מחשבים חיצוני שהגישה אליו נעשית דרך רשת האינטרנט.

תוכנה - SOFTWARE

אוסף של הוראות ומיגע הניתנות לביצוע על ידי מחשב. התוכנה משמשת להפעלת המחשב ותוכנות נלוות, לביצוע משימה או לאיסוף משימות.

התוכנה מורכבת מאוסף מאורגן של תוכניות מחשב המשרתות כולן יישום מסוים. באמצעות התוכנה המחשב מספק את שירותיו למשתמשים בו.





דואר אלקטרוני - אמצעי תקשורת אלקטרוני המאפשר העברת מסרים כתובים ממחשב אחד למחשב אחר (או לכמה מחשבים אחרים) דרך שרתים, אם מחשבים אלו מחוברים ברשת. דואר אלקטרוני ארגוני לרוב מנוהל בשרת מרכזי.

מערכת לניהול מסמכים - מערכות לניהול מסמכים (DOCUMENT MANAGEMENT SYSTEMS) הן מערכות מבוססות תוכנת מחשב (במקרה של ניהול מסמכים דיגיטליים) המשמשות כדי לעקוב, לנהל ולאחסן מסמכים.

רוב המערכות מסוגלות לשמור את הרשומה יחד עם הגרסאות השונות שנוצרו או שונו על ידי משתמשים שונים במהלך הזמן (מונח הנקרא "היסטוריית מעקב").

בממ"ה מערכת לניהול המסמכים היא "נתי"ב", המבוססת על תוכנת DOCUMENTUM. ברוב משרדי הממשלה המערכת לניהול המסמכים היא סע"ר (סביבת עבודה רוחבית).

מערכת תפעולית - קיימים בגופים המבוקרים מערכות ליבתיות לניהול תהליכים ומשימות ליבה של הגוף. לדוגמה, מערכת "נט המשפט" של הנהלת בתי המשפט, מערכת "תבל" של הביטוח הלאומי ומערכת "כלים שלובים" בהוצאה לפועל. בנציבות תלונות הציבור מערכת "מנתה" היא מערכת תפעולית.

ERP - ENTERPRISE RESOURCE PLANING SYSTEM - מערכת לתכנון משאבי ארגון - מערכת מידע המשמשת לעיבוד נתונים מינהלי המטפלת בכל צורכי עיבוד הנתונים בארגון (למשל, משאבי אנוש, כספים, תקציב, שיווק, לוגיסטיקה, תפעול ועוד) כמקשה אחת.

גישת ה-ERP שונה מהגישה הקלאסית בכך שהיא מאחדת את כלל מערכות המידע של הארגון למערכת אחת באופן המאפשר שימוש במידע משותף באופן אינטגרטיבי ובעל ממשק אחיד

מרכב"ה - מערכת רוחבית כוללת במשרדי הממשלה - מערכת ERP ממשלתית בפלטפורמת SAP לניהול נכסי ומשאבי המדינה, המצויה בשימוש כל משרדי הממשלה ויחידות הסמך (למעט המגזר הביטחוני). המערכת משמשת עשרות אלפי משתמשים פנימיים במשרדי הממשלה וכמו כן אזרחים, ספקים, עמותות וגופים נתמכים.

CRM - CUSTOMRT RELATIONSHIP MANAGEMENT - ניהול קשרי לקוחות - תחום העוסק בשירות לקוחות ובהבנה וניתוח של צורכיהם באמצעות כלים של טכנולוגיית המידע. התחום כולל מתודולוגיות, תוכנות מחשב וחומרות מחשב המסייעות לארגון לנהל את הקשר עם לקוחותיו.





גיבוי - עותקים ("גיבויים") של נתונים כך שאלו יכולים לשמש לשחזור המידע המקורי לאחר אירוע (פיזי או לוגי) הגורם לאובדן הנתונים המקוריים מאמצעי אחסון נתונים בו הם הוחזקו. תדירות הגיבויים תלויה בחשיבות השינויים שהמידע עובר, ובתדירות התרחשותם.





סוגי קבצים עיקריים

DOC\DOX - סיומת המאפיינת פורמט של קובצי טקסט שבשימוש גרסאות שונות של מעבד התמלילים WORD.

PDF - סיומת המאפיינת את הפורמט שמטרתו היא הצגה מדויקת של מסמכים ושאר תוכן ללא תלות בפלטפורמה.

PORTABLE DOCUMENT FORMAT - PPTX\PPPT - פורמט לשמירת מצגת שקפים. סוג פורמט זה ניתן לעריכה בתוכנת POWER POINT.

XLSX/XLS - קבצים המשמשים ב-MICROSOFT EXCEL. יישום גיליונות אלקטרוניים המשתמש בטבלאות לארגון, ניתוח ואחסון נתונים.

XLSM - הם קבצי גיליון אלקטרוני EXCEL התומכים בפקודות מאקרו (MACRI ENABLED).

מאקרו הוא רצף של פקודות בתוכנת האקסל ליצירת תהליכים אוטומטיים.

MDB\ACCDB - מנהלי בסיסי נתונים טבלאיים ומחוללי יישומים בעלי ממשק משתמש גרפי.

ניתנים לעריכה בתוכנת ACCESS /

EML - קבצים הכוללים הודעת אימייל אחת.

PERSONAL STORAGE TABLE - PST - קובצי יצוא נתונים בפורמט של OUTLOOK. מכילים את ההודעות, יומן ופריטים אחרים.

COMMA SAPRTATED VALUES - CSV - פורמט לשמירת טבלאות בצורת טקסט פשוט. בפורמט זה כל רשומה מופרדת באמצעות שבירת שורה וכל רשומה מורכבת משדות המופרדים באמצעות תו מוגדר, באופן שבדרך כלל נעשה שימוש בפסיק או בטאב. לכל התאים מספר שדות זהה.

BMP ,GIF ,PNG ,JPEG ,TIFF - קובצי תמונות וגרפיקה.

MHTM ,HTML - קובץ הכולל מידע לגבי תצורת דף האינטרנט.

ZIP\BAR - קובץ היכול להכיל מספר רב של קבצים בצורה דחוסה.





מטה דאטה - מידע טכני אודות קובץ הכולל בין היתר את שם המחשב עליו נכתב הקובץ, שם המשתמש, שם האפליקציה, סוג מערכת ההפעלה, מיקום הקובץ ופרטים נוספים כדוגמת תאריכי יצירה ושינויים.

שליפת נתונים

- USER INTERFACE - UI

הצורה וויזואלית שבה המשתמש צורך/רואה את המידע. לדוגמה, הנראות של דף האינטרנט.

- USER EXPERIENCE - UX

מונח המתייחס לחוויה הכוללת של המשתמש במוצר כלשהו, בין אם מדובר באתר, באפליקציה או במוצר פיזי. הרעיון ב-UX הוא להתמקד בפונקציונליות ולגרום למוצר לא רק להיראות טוב, אלא להיות שמיש ונוח עבור המשתמש.

- מסכי קלט ופלט

קלט - הוא התקן המאפשר למשתמש לימור קשר עם המחשב, להכניס לו מידע ולתת לו פקודות (לדוגמה, מקלדת, עכבר, מצלמה).

פלט - פלט הוא התקן המאפשר למחשב ליצור קשר עם המשתמש על ידי הצגת המידע על המסך, ליצור הדפסה או יצירת קול (לדוגמה, מסך המחשב, רמקולים, מדפסות).

דוחות מובנים וחולל דוחות -

דוח מובנה - דוחות שהוגדרו מראש הן בהיבט התצוגה והן בהיבט מועד ההפקה.

מחולל דוחות - מחולל הדוחות מאפשר גמישות בשליפת הנתונים, בשונה מהדוח המובנה.

המחולל מייצר דוח "מותאם אישית" לפי הגדרות שהוזנו על ידי המשתמש.

BI - BUSINESS INTELLIGENCE - כלי ה-BI מאפשר לראות בתצוגה ידידותית וויזואלית נתונים ומגמות, ובעזרתו ניתן לבצע חקירה לאחור לשליפת נתוני המקור.

שליפת מידע ממסד נתונים - שליפת מידע ישירות ממסד הנתונים נעשית באמצעות שפות תכנות, לדוגמה SQL.





העברת חומרים מהגוף המבוקר

DOK - (דיסק און קי), תקליטור ודיסק קשיח נייד.

כספת וירטואלית - כספת וירטואלית היא טכנולוגיה המשמשת לקליטה והפצה של קבצים בצורה מאובטחת בין ארגון לארגון.

הלבנה והשחרה - קבצים המוכנסים לרשת המשרדית נדרשים לעבור תהליך של הלבנה (בדיקה וניקוי) טרם הכנסתם כדי לוודא שאינם מכילים תוכנה זדונית כלשהי.

מסדי נתונים

DATABASE - **מסד נתונים** - אמצעי המשמש לאחסון מסודר של נתונים במחשב לשם אחזור ועיבודם.

מסד נתונים מאוחסן באמצעות אמצעי אחסון נתונים, בדרך כלל על גבי דיסק קשיח. מאפשר גישה ישירה לנתונים.

מסד הנתונים בנוי לפי מודל לאחסון נתונים, כמו מנגנונים פנימיים למיון ולחיפוש.

RELATIONAL DATABASE - RDB - מסד נתונים יחסי מבנה אופטימלי של טבלאות ושדות היוצרים בסיס נתונים, באופן שהמידע מאוחסן באופן אמין ויעיל ובדרך החסכונית ביותר.

SQL - STRUCTURED QUERY LANGUAGE - שפת מחשב הצהרתית לטיפול במידע ועיבודו בבסיסי נתונים יחסיים. השפה מאפשרת שליפת נתונים ועדכוןם ויצירת טבלה ושינויה.

רשומה (שורה) - מכילה נתונים של פריט אחד, כדוגמה פרטים על עובד אחד.

שדה (תא) - הרשומה מתחלקת לשדות. השדה מכיל פרט אחד של מידע, כדוגמת שם המשפחה של העובד.

מפתח ראשי - מונח מעולם בסיסי הנתונים היחסיים. בעיצוב בסיסי הנתונים משתמשים במפתח ראשי כזה כדי לזהות שורה באופן חד-ערכי בטבלה, כך שהיא תקבל ערך פעם אחת לכל היותר, כדוגמת מספר ת.ז.; אין שני עובדים בעלי מספר ת.ז. זהה.

מפתח משני (עמודה) - מתאר שדה בטבלה אשר ערכיו האפשריים נשאבים משדה מקביל בטבלה אחרת אשר בה משמש השדה כמפתח ראשי. למשל, תפקיד. ייתכנו מספר עובדים שיש להם תפקיד זהה.





טבלה - כוללת רשומות רבות, כדוגמת טבלת עובדי הארגון. לכל טבלה מפתח ראשי.

בסיס נתונים - מידע כולל של טבלאות שונות.

ניתוח נתונים

תוכנות נפוצות לניתוח נתונים ולהצגת מידע

MICROSOFT EXCEL - תוכנה ליצירת גיליונות אלקטרוניים, שפותחה ומופצת על ידי מיקרוסופט וכוללת כלי חישוב, כלים גרפיים, טבלאות ציר, וכן שפת תכנות מאקרו בשם

VBA - VISUAL BASIC FOR APPLICATIONS.

MICROSOFT ACCESS - תוכנה לניהול בסיס נתונים טבלאי (DBMS) ומחולל יישומים בעל ממשק משתמש גרפי. התוכנה מיועדת למשתמשים מתקדמים, לסביבת עבודה קטנה ובינונית, ומאופיינת בקלות הפעלה ובממשק ידידותי.

היא יכולה לייבא או לקשר ישירות נתונים המאוחסנים ביישומים ובמסדי נתונים אחרים.

SPSS - **STATISTICAL PACKAGE FOR SOCIAL SCIENCE** - אחת התוכנות הנפוצות בעולם לניתוח סטטיסטי של נתונים.

ACL/IDEA - תוכנה ייעודית לביקורת בעלת יכולת רחבה לקלוט מידע מסביבות טכנולוגיות מגוונות ולניתוח תוכנו ואמינותו על פי מגוון רחב של משתנים.

POWER BI - שירות תוכנתי המאפשר אנליזת נתונים, המפותח על ידי מיקרוסופט. השירות מספר המחשות חזותיות אינטראקטיביות עם שירות-עצמי של יכולות החזיית מידע, שבו משתמשי קצה יכולים ליצור בעצמם דוחות ולוחות מחוונים שיתעדכנו באופן אוטומטי, בלי להסתמך על צוות טכנולוגיית מידע או מנהל מסד נתונים.

OPTICAL CHARACTER RECOGNITION-OCR - טכנולוגיה להמרת תמונה, טקסט מודפס וכתב יד שנסרקו על ידי סורק, למידע דיגיטלי הניתן לעיבוד.

ABYX - כלי להמרת PDF למידע דיגיטלי הניתן לעיבוד (לדוגמה, WORD או EXCEL בטכנולוגיית OCR). כלי זה קיים במשרד מבקר המדינה.





מושגים בניתוח נתונים

כריית נתונים - DATA MINING -

שימוש באלגוריתמים לסקירה, חקירה וניתוח מידע הטמון בבסיס נתונים קיים במטרה להסיק ולזהות דפוסים חדשים, שימושיים ובעלי משמעות.

כריית נתונים היא בעצם תהליך שנועד לחקור כמות גדולה של נתונים באמצעים אוטומטיים ככל הניתן במטרה להפוך מידע לידע.

התנאים לאיכות המידע -

דיוק - מידת התאימות בין סט של נתונים לבין סט של ערכים נכונים בהתאמה. משמע, נתונים שאינם עומדים בסתירה לנתונים אחרים.

עדכניות (CURRENT) - ישות או ישויות מרובות המסופקות בזמן הנדרש או בנקוב מראש וקובעות את מידת עדכניות הנתונים.

עקביות וסטנדרטיות (CONSISTENCY AND STANDARDIZATION) - בארגון עם נתונים איכותיים כל המידע חייב להיות מובנה באותה צורה, על בסיס סטנדרטים שנקבעו ברחבי הארגון כולו, וכך ליצור עקביות פנימית בין הנתונים.

שלמות (INTEGRITY) - הצגת תמונה מלאה על הישויות שהמידע עוסק בהן.

נכונות ומהימנות - עקביות לוגית המאפשרת קשר הגיוני בהשוואה לנתונים דומים, בין השאר מניעת כפילויות והגדרה של חוקים עסקיים המגדירים את המידע באופן ברור.

ייחוד (UNIQUENESS) - מצב שבו יש אילוצים על הנתונים באופן שבו יהיו ייחודיים, ללא חזרות מיותרות - כל ערך יהי ייחודי ויחיד מסוגו.

פעולות בניתוח נתונים

טיוב - טיוב נתונים הינו תהליך לשיפור איכות הנתונים האגורים בבסיסי נתונים. השיפור נדרש בנתונים שאינם שלמים/מדויקים/נכונים.

קיבוץ - קיבוץ נתונים מאפשר לקבץ נתונים הנמצאים בעמודה אחת או יותר לשורה מקובצת אחת.

פירוק - פירוק הוא פעולה הפוכה מפעולת הקיבוץ. פירוק נתונים חזרה לתצורתם המקורית.





הצפת חריגים - הצפת חריגים יכולה להיות חריגה מערכים הנמצאים במדגם. בחינת הצורה הכוללת של הנתונים בתרשים כלשהו חשובה על מנת לכלול סימטריה וסטיות. בחינת הנתונים לאיתור תצפיות חריגות הרחוקות ממסת הנתונים.

אנומליות - אנומליות הוא מונח מתחום כריית המידע המתייחס לאיתור תצפיות או אירועים אשר אינם תואמים לדפוס הצפוי או ליתר התצפיות בבסיס הנתונים.

נקודות קיצון (MIN, MAX) - נקודות קיצון הן נקודות שבהן ערך אחד הוא הגבוה ביותר והערך השני הוא הנמוך ביותר.

רשומות כפולות - ערכים כפולים החוזרים על עצמם. לעיתים מדובר בשגיאה שיש לטייב לפני ניתוח בסיס הנתונים.

רשומות חסרות - לעיתים קיימת חשיבות למספור רציף, כדוגמת מספרי קבלות, במקרים כאלו יש לבצע בדיקה של רשומות חסרות. לדוגמה, מספרי חשבוניות 1, 2...5...6.

