



מהדורה שנייה מעודכנת
ומורחבת

DATA DRIVEN CONSTRUCTION

ניווט בעידן הנתונים בתעשיית הבנייה

עם דוגמאות לשימוש בבינה
מלאכותית ובמודלים של שפה
גדולה

Artem Boiko



DATA-DRIVEN CONSTRUCTION

ניווט בעידן הנתונים בתעשיית הבנייה

מהדורה שנייה, מתוקנת ומורחבת

ארטם בויקו



זה"ה האחרון של המפורסם במשפט - המידע טכנולוגיית של קרוויל יימס'ג כמו הוא בויקו הנתונים זה". הוזה המפורסם בספר מוחלפת להיות צריכה אחת מילה רק, "טיפש, הכלכלה הרומאים של פתגם, הנתונים בעולם הדרך את למצוא וכדי (התוכנה לא) "טיפש העמקות בכל קוראיו את מנווט המחבר. "הכרחי הוא ניווט": היום תקף שעדיין הקדומים גישה על לדבר לא, מתפשר בלתי ובמצפן בטוחה ביד הנתוני האוקיינוס של והשטחים במבט ניכר רק שלא טוב הומור וחוש מאוד מקורית גרפיקה, ולבסוף, מקיפה היסטורית מה, מרה די לספקנות נלהב אישור בין נעה בויקו של לספרו הבינלאומית התגובה. שני ולא מקורי נתונים חושב הוא בויקו. בגרמניה הספר של השנייה למהדורה טוב שעשה אפילו, גועזות תמיד theses ות מרגשות תובנות לקורא מציע הוא. דוגמטי של הגרמנית למחלה מצוינת תרופה. נוספת מחשבה שמעוררות, פרובוקטיביות לא זה. "הכרחי לא זה לחיות": תוספת יש ל"הנ הלטיני הפתגם, אגב. חבוי קונסנוואליזם לומר לא, הכרחיים וחייהם חיים הנתונים - הנתונים לעולם בויקו של גישתו על חל קריטיים.

המפורסם, BIM: המגזין עבור כולל - עצמאי עורך, טליבירי בורקהרד ר"ד —
2013 מאז וסון ארנסט ידי על שנה מדי

אירלנד של הלאומי התקנים במכון הלאומי

בפורמטים והחזקתו הבנייה מגור עבור קרדינלי נכס הוא מידע, בספר שהודגש כפי של הזמנים לוחות את ומאיצה מדויקת החלטות קבלת על מאוד מקלה נגשים בקבלת זה מקור ולניצול לגישה ויעילה ניטרלית גישה מציע הספר. הפרויקטים מונחה תכנות המשלבת עכשווית גישה משלבת בספר המוצגת המתודולוגיה. החלטות המלאכותית הבינה כוח ניצול ידי על. נגשים פתוחים כלים עם מלאכותית בינה תהליכים לייעל, אוטומציה לשפר שואפת המתודולוגיה, פתוח בקוד בתוכנה ושימוש להבנה וקלה ברורה הספר של השפה. בתחום פעולה ושיתוף נגישות ולקדם

באוניברסיטת ואדריכלות יפות לאמנויות הפקולטה דיקן, אופולוגלו סאליה ר"ד – BIM-ל אסיאני-האירו הפורום ומארגן, ביילים אנטליה

“

ואת, גרפיקה, LLM, והסטוריה שילבת שבה הדרך! וואו, זה לומר יכול שאני מה כל הספר של הוריהם. מרשימה באמת היא שלך הנקודות בהבנת הכללית הנחות זהו. משחק משנה באמת זה; הוזה בספר מבריקים אספקטים הרבה כך כל יש. מדדומה על טוב מול. בו שהשקעת והתשוקה המאמץ על אותך משבח ואני, למידע מצוין מקור! מתרשם מאוד שאני לומר מספיק אבל, להמשיך יכול אני. מרשימה כה עבודה יצירת

energylab-ב דיגטלית פרקטיקה מנהלת, פרינסלו נטשה –

“

הזה הספר, המנוסים המקצוענים ועד מהטירונים, הבנייה בתעשיית שנמצא מי כל עבור ונגיעה, אסטרטגיות, בתוכנות מלא זה-ומעייפת רגילה קריאה לא זה! (משחק משנה טכנולוגיות ועד עתידות נתונים רישום משיטות. מעורב אותך לשמור כדי הומור של כמו זה. בבנייה בנתונים השימוש של האבולוציה את מכסה הוא, מתקדמות דיגטליות מהנדס, אדריכל אתה אם בין. הבנייה נתוני של האבולוציה דרך זמן מכונת לקחת הדרך את revolutionize י הוזה המקף המדריך, נתונים אנליסט או פרויקטים מנהל, ההחלטות קבלת את לשפר, תהליכים לייעל התכונן. לפרויקטים נוגש אתה שבה לעשית לא שמעולם כמו פרויקטים ולנהל

ב-ככר פרויקטים ומנהל, ברומא סאפינצה באוניברסיטת מרצה, ורגני פירפאולו – Fintecna.

“

מצוינת הספר של הייצור איכות. שעות 6-מ בפחות, אחת בנשימה הספר את קראתי כיצד מעשיות דוגמאות של רב מספר. נעים גופן, צבעוניות סכמות, ודחוס מבריק נייר לימוד של, שנים לא אם, חודשים לך יחסוך הבנייה לתעשיית ספציפי LLM עם לעבוד לדרוש מבלי, מורכבות ועד מפשטות החל, מאוד מוגזגות העבודה דוגמאות. עצמי הבנייה תעשיית בכל עסקים לבעלי יאפשר הספר. ויקרה מורכבת תוכנה לרכוש ממך חדשה בצורה פיתוח ופרספקטיבות, דיגטציה, שלהם העסק אסטרטגית על להסתכל וחינוכים נגשים כלים עם העילות את להגדיל, יותר קטנות חברות ועבור

ב-הבנייה בתעשיית דיגטלית טרנספורמציה על ויועץ מרצה, קוסארב מיכאיל – TIM-ASG.

“

- הבנייה בתעשיית הדיגטציה של לדמוקרטיזציה דרך אבן הוא בויקו ארטס של ספרו: דרך פורץ במיוחד. (SMEs) ובינוניים קטנים עסקים עבור אמיתי משחק ושינוי יכולות חברות, קוד ללא או מעט עם פתוח קוד של מודרניים כלים שימוש באמצעות צורך ללא - ברווחיות אותם ולנתח שלהן העבודה בתהליכי ביעילות נתונים לשלב כבר מסחריות תוכנה בחבילות היקר השימוש את מבטל זה. בתכנות מעמיק בידע רק לא שרוצה מי לכל ערך יקר מדריך זהו! לפעולה קריאה הוא זה ספר. מטורבלות - פעיל באופן אותו לעצב רוצה גם אלא, הבנייה בתעשיית הדיגטלי השינוי את להבין הוזה הידע את לשתף כדי יחד לעבוד הזמן עכשיו. ומתקדמת יעילה, פרמטית בצורה. הבנייה תעשיית של הפרודוקטיביות את קיימא בר באופן ולהגביר

בעלים, קונסטנץ HTWG-ב בנייה לניהול פרופסור, ביולר מקס מיכאל ר"ד – GemeinWerk Ventures-ב משותף

“

לגבולות מעבר הראשונים הצעדים אחד הוא DataDrivenConstruction ספ כאשר, שלהן המורכבות והניהול העיצוב מערכות עם, הבונים של הרגיל העולם ולשקופות רדיקלית לפשטות סיכוי נותנת לא הנתונים והשבועות המורכבות, נדמה הדומנויות אילו פשוטה בשפה מראה ארטס, בספרו. הבנייה נתוני עם בעבודה מוגברת קונקרטיים צעדים ומספק, בפנינו נפתחות נתונים עם עבודה של מודרניות טכנולוגיות מערכות יובילו לאן להבין שרוצה מי לכל ממליץ אני. בעבודתכם מיד ליישם שניתן שהמהפכה להבין כדי בקפידה הוזה הספר את ללמוד הבנייה בתעשיית האוטומציה בעוד אבל, עכשיו הגיקים את רק מעניין זה. דלתנו על דופקת כבר בבנייה הנתונות למקום בכל יהיו כאלה ותוכנות גישות, BIM, כמו, שנים כמה

ב-דיגטלית טרנספורמציה BIM & IMT, המומחיות מרכז ראש, ב'רוגצ איגור – RGD, InfraBIM.Pro ומייסד

“

שמציין כפי, שעוסק DataDrivenConstruction ספר על כחום ממליץ אני כדי כרגע בו משתמש אני. AECO עבור נתונים מונחה מידע לניהול בגישה, הכותרת בנוסף. מאוד נגיש כמקור אותו מצאתי. שונות קבוצות עם דיונים מספר ביוזמת לעזור מספר והצגת הנתונים, AECO-ב הכלים של ההיסטורי ההקשר של מעמיקה לסקירה דיקה את שמפרטות, שימושיות מאוד דיאגרמות מספר מכלל הספר, מפתח טכנולוגיות עבודה לזרימות דוגמאות עם הסופי המשתמש של האובייקטים ואת הנתונים מקורות מפתחים כאשר מהם יותר צריכים שאמזנו הסוגים הן הללו שהדיאגרמות לי נראה זה נתוני של הכולל המודל את המגדיר - BEP-ל ותורמים מידע אסטרטגיות ומנטרים i-AIM- PIM עבור הגבול את להניח ניתן שעליו הארגון

ב-מערכות של אינטגרציה ומהנדס, Acmena-ב ראשי ויועץ, רנסלי פול – Transport for London.

“

לחצוב, אותם למצוא, אותם להגדיר ללמוד עלינו, "החדש הנפט הם הנתונים" אם הספר את מצאתי. לערך אותם להפוך כדי, אותם לזקק, אותם מספק הספר. הישראה ומעורר אינפורמטיבי מאוד DataDrivenConstruction שמעוניין מי עבור. פשוטה בשפה נתונים עם עבודה ומסביר שימושי היסטורי רקע מובנים הם איך, עובדים הם איך - נתונים של טובה הבנה נותן הוא, דיגטלי בשינוי בהם להשתמש ניתן ואך

BIM-ה ועדת ר"וי, BIM, מתאמי פסגת מנהל, ArcDox-ב מנהל, מונטגיו ראלף –

“

“

מי כל עבוד משחק משנה הוא "DATA DRIVEN CONSTRUCTION" ספר רק לא ארטס. הנתונים בעידן הבנייה תעשיית הולכת שבהם הכיוונים לגבי שסקרן וההזדמנויות, האתגרים, הנוכחיות ההתפתחויות את לעומק חוקר הוא; בנושא נוגע מסביר ארטס – שלו הנגישות הוא הזה הספר את שמייחד מה. בבנייה המבטחות מצאתי. להבנה לקל התוכן את שהופכות נגישות אנלוגיות באמצעות מורכבים רעיונות שלא ערך יקר משאב יצר ארטס, לסיכום. מרתק גם אך אינפורמטיבי מאוד הספר את בתחום חדש או מנסה מקצוען אתה אם בין. השראה מעורר גם אלא מידע מספק רק לגבי שלך ההבנה את ויעמיק שלך הפרספקטיבה את ירחיב הזה הספר, הבנייה !בחום מומלץ. התעשייה הולכת שבהם הכיוונים

BIM TMM GROUP ב- יישום ומנהל אדריכל, סלה מועיאר – Gesamtplanungs GmbH.

“

לימוד כספר נלמד להיות ראוי "Data-Driven Construction"-ש לומר חייב אני "Data-Driven Construction" ספר. המתפתח BIM-ה לתחום רבות שיתרום ספר והוא באוניברסיטאות טובה בצורה המונחים את שמסביר טכני מילון מכיל "Data-Driven Construction" מאוד ויוזאליה בשפה מובנים פשוטים נעשים להסביר מאוד שקשה נושאים. מאוד אם גם, לקורא מועבר להיות צריך ביוזאליה להסביר שמתכוונים שמה חושב אני. יפה מידע דורשת הוויזואליה קריאת, אחרות במילים, מהוויזואליה כמה של ההבנה. בקצרה בויק ארטס של היקרה העבודה את להציג שמה שאני לומר רוצה גם אני. נוסף באוניברסיטאות שלי ובסמינרים בהרצאות

בבנייה פרויקטים בניהול ומרצה, ArchCube-ב בעלים, יזיווגלו אדיו ר"ד – מדיפול ובאוניברסיטת באיסטנבול טכניון באוניברסיטת לאדריכלות במחלקה

“

המידעית העבודה יסודות את חיה בצורה מעביר "Data Driven Construction" את מבדל וכך בסיסיים כלכליים ובמונחים מידע בזרמי שעוסק ספר. בנייה נתוני עם תוכנה יצרן של המבט נקודת את מציג רק לא הוא כי, BIM-ה ספרי משאר עצמו ולראות לקרוא ששווה ספר. בסיסיים מושגים להעביר מנסה גם אלא

פורום ומייסד, Build Informed GmbH של שותף ומייסד ל"מנכ, דירן יאקוב – "On Top With BIM".

“

צריכים הכורים או הנתונים מחפשי ולכן, שאומרים כפי, "החדש הנפט הם נתונים" 21-ה המאה של מהמשאב ערך להפיק כדי הנכונים ובמנטליות בכלים מצוידים להיות מידע" מבוססי תהליכים של חלקלק מדרון על רב זמן הייתה הבנייה תעשיית. הזה מישו ידי על שהוכן מידע סמך על מתבצעת הפרויקטים מסירת שבהם, "ממדי-תלת "נתונים"שה בעוד (העוגה תרשים את או הגרף את תכננו כבר הם, למשל) אחר מיזוג כי במיוחד, יותר הרבה לספק יכולים (הגולמי הנתונים גליון, למשל) הבסיסיים מספק אתה אם. מוגבלים בלתי פוטנציאליים מביאים מלאכותית ובינה מרובים נתונים - כה עד והיחיד - שלך ביותר הטוב המשאב הוא הזה הספר, בנייה (חוקר/מלמד או) עצמנו את מציאו שבו נתונים המונע בעולם לניווט

ב-בבנייה אסטרטגיות מידע בטכנולוגיות חבר פרופסור, אדמו זולפיקאר ר"ד – LSBU, בריטניה.

“

שמציעה מרשימה עבודה הוא בויק ארטס מאת "Data-Driven Construction" וגדלות הולכות מידע וצורות טכנולוגיות של בעידן הבנייה לתעשיית מוצק בסיס חזוניים רעיונות הצגת כדי תוך מובנת בצורה מורכבים נושאים להציג מצליח בויק גם אלא הנוכחיות ההתפתחויות את רק לא שמדגיש דיטב מתוכנן אוסף הוא הספר תכנון את להבין שרוצה מי לכל בחום מומלץ הוא. חידושים על לעתיד מבט מספק נתונים המונעת הבנייה וביצוע

מנהל, שימושיים למדעים שטוטגרט באוניברסיטת מרצה, ר'אברג מרקוס – Konstruktionsgruppe הבנייה בקבוצת סניף מנהל וסגן בכיר פרויקטים BIM Cluster Baden- עמותת של המנהלים במועצת חבר, Württemberg.



מהדורה שנייה, מרץ 2025 © 2025 | ארטם בויק | קרלסרוהה

ISBN 978-3-912002-21-8



ארטם בויק זכויות יוצרים

boikoartem@gmail.com

info@datadrivenconstruction.io

שום חלק מהספר הזה לא יכול להיות משוכפל או מועבר בכל צורה ובאמצעים כלשהם, אלקטרוניים או מכניים, כולל צילום, הקלטה או כל מערכת אחסון וחיפוש מידע, ללא רשות בכתב מהמחבר - למעט הפצה לא מסחרית בצורה לא משתנה. הספר מופץ בחינם וניתן להעבירו בחופשיות למשתמשים אחרים למטרות אישיות, חינוכיות או מחקריות, בתנאי שיש לשמור על זכויות היוצרים והקישורים למקור. המחבר שומר על כל הזכויות הלא חומריות על הטקסט ואינו מעניק שום ערבויות מפורשות או משתמעות. החברות, המוצרים והשמות המוזכרים בספר עשויים להיות דמיוניים או בשימוש כדוגמאות. המחבר אינו נושא באחריות לכל תוצאה של שימוש במידע המובא. המידע המופיע בספר מוצג "כפי שהוא", ללא ערבויות של שלמות או עדכניות. המחבר אינו נושא באחריות להפסדים אקראיים או עקיפים הנובעים משימוש במידע, בקוד או בתוכניות המופיעות בספר. הדוגמאות לקוד המוצגות בספר מיועדות אך ורק למטרות חינוכיות. הקוראים משתמשים בהן על אחריותם. המחבר ממליץ לבדוק את כל הפתרונות התוכנתיים לפני השימוש בסביבה ייצורית. כל הסימנים המסחריים ושמות המוצרים המוזכרים בטקסט הם סימנים מסחריים, סימנים מסחריים רשומים או סימני שירות של החברות המתאימות והם רכוש בעלי הזכויות המתאימים. השימוש בשמות אלה בספר אינו מעיד על קשרים עם בעליהם או על אישור מצידם. אזכור של מוצרים או שירותים של צדדים שלישיים אינו מהווה המלצה ואינו מרמז על תמיכתם. שמות החברות והמוצרים המוזכרים כדוגמאות עשויים להיות סימנים מסחריים של בעליהם. הקישורים לאתרי אינטרנט חיצוניים ניתנים לנוחות בלבד ואינם מעידים על כך שהמחבר מאשר את המידע המוצג באתרים אלה. כל הטקטיסטיות, הציטוטים והמחקרים המובאים היו עדכניים במועד כתיבת הספר. הנתונים עשויים להשתנות עם הזמן.

Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) ספר זה מופץ תחת רישיון

CC BY-NC-ND 4.0

ארטיום בויק. מהדורה ראשונה © 2025. ארטיום בויק. מהדורה שנייה, מעודכנת ומורחבת. כל הזכויות שמורות © 2024

הקדמה למהדורה השנייה

ספר זה הוא תוצאה של דיאלוג חי עם הקהילה המקצועית. הוא מבוסס על דיונים מקצועיים רבים בנושאי עבודה עם נתונים בתעשיית הבנייה, שהתקיימו בפלטפורמות מקצועיות שונות וברשתות חברתיות. דיונים אלה הפכו לבסיס למאמרים, פרסומים וחומרים ויזואליים שזכו לתגובה רחבה בקהילה המקצועית. החומרים של המחבר מושכים מדי שנה מיליוני צפיות בפלטפורמות שונות ובשפות שונות, מאחדים אנשי מקצוע בתחום הדיגיטציה של הבנייה

במהלך השנה שלאחר פרסום המהדורה הראשונה, הזמינו מומחים מיותר מ-50 מדינות את הספר - מברזיל ופרו ועד מאוריציוס ויפן. המהדורה השנייה של הספר, אשר בידיכם כעת, עברה עיבוד והרחבה בהתבסס על משוב של מומחים הערות ביקורתיות על המהדורה הראשונה ודיאלוגים במעגלים מקצועיים. בזכות המשוב, המהדורה השנייה הורחבה יעילים. כמו כן, מספר ETL ועל יצירת תהליכי CAD (BIM) באופן משמעותי: נוספו פרקים חדשים על טכנולוגיות הדוגמאות המעשיות והמקרים גדל באופן משמעותי. ערך מיוחד טמון במשוב ממובילי תעשיית הבנייה, חברות ייעוץ הגדולות, שפנו למחבר עם שאלות על דיגיטציה ואינטרופראביליות הן לפני פרסום הגרסה הראשונה והן IT וחברות. לאחר מכן. רבים מהם כבר מיישמים את הגישות המתוארות בספר או מתכננים לעשות זאת בקרוב.

אתם מחזיקים בידיכם ספר שנוצר בזכות דיונים והחלפת דעות פעילה. ההתקדמות נולדת בדיאלוג, בהתנגשות דעות ובפתיחות לגישות חדשות. תודה שאתם חלק מהדיאלוג הזה. הביקורת הבונה שלכם היא הבסיס לשיפורים עתידיים. אם תגלו שגויות בטקסט או אם תרצו לשתף רעיונות והצעות, כל משוב יתקבל בברכה. פרטי הקשר ליצירת קשר מופיעים בסוף הספר.

מדוע הספר חיוני?

ספר זה נועד להיות משאב חיוני פתוח, המכוון להפצת גישות מודרניות לניהול נתונים בתעשיית הבנייה. הגרסה הראשונה של הספר שימשה כבסיס לאיסוף תגובות והצעות מהקהילה המקצועית, מה שאפשר לשפר את המבנה והתוכן של החומר. כל ההערות, ההצעות והרעיונות נבחנו בקפידה ונלקחו בחשבון בגרסה זו, שעברה שיפוץ. מטרת הספר היא לסייע למומחי תעשיית הבנייה להבין עד כמה חשוב לעבוד עם נתונים: בצורה מערכתית, מודעת ובמבט לעתיד על ערך המידע. המחבר אסף דוגמאות, איורים ותצפיות מעשיות במשך יותר מ-10 שנים בעבודתו בתחום הדיגיטציה של הבנייה רוב החומרים הללו נוצרו במהלך פרויקטים אמיתיים, דיונים עם מהנדסים ומפתחים, השתתפות ביוזמות בינלאומיות וקיום סדנאות הכשרה. הספר הוא ניסיון למקד את הניסיון שנצבר ולשתף אותו בצורה נגישה. אם תרצו לתמוך בהפצת רעיונות הספר ולקבל פורמט נוח לקריאה, עבודה עם דוגמאות וחומרים ויזואליים - תוכלו לרכוש את הגרסה המודפסת.

זכויות שימוש בחומרים

כל החומרים, האיורים והקטעים מתוך ספר זה יכולים להיות משוכפלים, מצוטטים או בשימוש בכל פורמט ובכל מדיה, תודה על הכבוד לעבודה. "Data-Driven Construction" בתנאי שצוין מקור: מחבר ארטם בויקו ושם הספר. ולהפצת הידע.

בהכרת תודה כנה אני מקדיש ספר זה למשפחתי, שהעניקה לי מגיל צעיר
אהבה עמוקה לבנייה, לעיר המכרות שלי - על שיעורי החוסן ולרעייתי
מודדת המידע שלי, שהייתה תמיכתה הבלתי משתנה מקור הראה מתמיד
עבורי.

למי מיועד ספר זה

ספר זה, שנכתב בשפה נגישה, מיועד לקהל רחב בתעשיית הבנייה - מסטודנטים ומתחילים המעוניינים להבין את יסודות תהליכי הבנייה המודרניים, ועד מקצוענים הזקוקים למתודולוגיה עדכנית לניהול נתונים בבנייה. אם אתה אדריכל, מהנדס, מנהל עבודה, מנהל בנייה או אנליסט נתונים, מדריך מקיף זה עם איורים וגרפים ייחודיים מציע תובנות יקרות ערך כיצד להשתמש בנתונים בעסק כדי לייעל ולמכן תהליכים, לשפר את קבלת ההחלטות ולנהל פרויקטים בבנייה ברמות שונות באמצעות כלים מודרניים.

הספר מהווה מדריך מקיף, המשלב יסודות תיאורטיים והמלצות מעשיות על שילוב שיטות ניהול נתונים בתהליכי הבנייה. הספר מתמקד בשימוש אסטרטגי במידע כדי לייעל את הפעילות התפעולית, למכן תהליכים, לשפר את קבלת ההחלטות ולנהל פרויקטים ביעילות באמצעות כלים דיגיטליים מודרניים.

בעמודי ספר זה נבחנים היבטים תיאורטיים ומעשיים של עבודה עם מידע בתעשיית הבנייה. באמצעות דוגמאות מפורטות נבחנת המתודולוגיה של פרמטריזציה של משימות איסוף דרישות, עיבוד נתונים לא מובנים ומגוונים והפיכתם לפתרונות יעילים עבור חברות בנייה.

הקורא עובר בהדרגה מהגדרת דרישות ופיתוח מודלים בסיסיים של נתונים לתהליכים בניית ETL, מורכבים יותר של אינטגרציה של מקורות מידע שונים, יצירת תהליכי צינורות מידע ומודלים של למידת מכונה. הגישה ההדרגתית מאפשרת להדגים בצורה ברורה את המנגנונים של ארגון ומיכון תהליכי עסקים ומערכות תמיכה בקבלת החלטות, בתחום הבנייה. כל חלק בספר מסתיים בפרק מעשי, המכיל הוראות שלב אחר שלב. המאפשרות ליישם מיד את הידע שנרכש בפרויקטים אמיתיים.

סקירה קצרה של חלקי הספר

ספר זה בנוי סביב הקונספט של טרנספורמציה של נתונים בשרשרת יצירת הערך: מאיסוף הנתונים ואבטחת האיכות שלהם ועד לעיבוד אנליטי והפקת פתרונות מעשיים יקרי ערך באמצעות כלים ומתודולוגיות מודרניות.

חלק 1: האבולוציה הדיגיטלית בבנייה - עוקבת אחרי השינוי ההיסטורי בניהול נתונים מהלוחות החרס עד למערכות דיגיטליות מודרניות, מנתחת את הופעת המערכות המודולריות ואת עליית החשיבות של דיגיטציה של מידע בהקשר של מהפכות תעשייתיות.

חלק 2: האתגרים המידעיים של תעשיית הבנייה - בוחן את בעיות הפיצול של הנתונים, "סילו מידע", השפעת AI ו-LLM על קבלת ההחלטות והמגבלות של פורמטים קנייניים, מציע לשקול מעבר למערכות HiPPO גישת LLM.

חלק 3: מערכת נתונים בבנייה - מפתח טיפולוגיה של נתוני בנייה, מתאר שיטות לארגון, אינטגרציה עם מערכות ארגוניות ודן בהקמת מרכזי מומחיות לסטנדרטיזציה של תהליכי מידע.

חלק 4: הבטחת איכות הנתונים - חושף מתודולוגיות להמרת מידע מפורז לנתונים איכותיים ומסודרים, כולל LLM חילוץ נתונים ממקורות שונים, אימות ומידול באמצעות.

חלק 5: חישובי עלות וזמן - מוקדש לדיגיטציה של חישובי עלות ותכנון, אוטומציה של קבלת כמויות מדגמי של פרויקטי בנייה ESG וחשוב מדדי D-8D טכנולוגיות מודל 4, CAD- (BIM-).

חלק 6: מנתח באופן ביקורתי את האבולוציה של טכנולוגיות תכנון, בעיות תאימות של BIM ו-CAD: חלק 6 מערכות, מגמות המעבר לפורמטים פתוחים של נתונים ופרספקטיבות של שימוש בבינה מלאכותית בתכנון.

חלק 7: אנליטיקה של נתונים ואוטומציה - בוחן את העקרונות של ויזואליזציה של מידע, מדדי ביצוע מרכזיים כלים לאורקסטראציה של תהליכי עבודה ושימוש במודלים לשוניים לאוטומציה של משימות ETL, תהליכי שגרתיות.

חלק 8: אחסון וניהול נתונים - חוקר את פורמטי אחסון הנתונים, מושגי מחסנים ואגני נתונים, עקרונות ניהול DataOps ו-VectorOps נתונים וגישות חדשות, כולל בסיסי נתונים וקטוריים ומתודולוגיות.

חלק 9: נתונים גדולים ולמידת מכונה - מוקדש למעבר לניתוח אובייקטיבי המבוסס על נתונים היסטוריים האינטרנט של הדברים באתרי בנייה ושימוש באלגוריתמים של למידת מכונה לחיזוי עלויות וזמנים של פרויקטים.

חלק 10: תעשיית הבנייה בעידן הנתונים הדיגיטליים - מציג מבט על עתיד תעשיית הבנייה, מנתח את המעבר מניתוח סיבתי לעבודה עם מתודולוגיות קורלציה, המושג "אובריזציה" של הבנייה ואסטרטגיות של טרנספורמציה דיגיטלית.

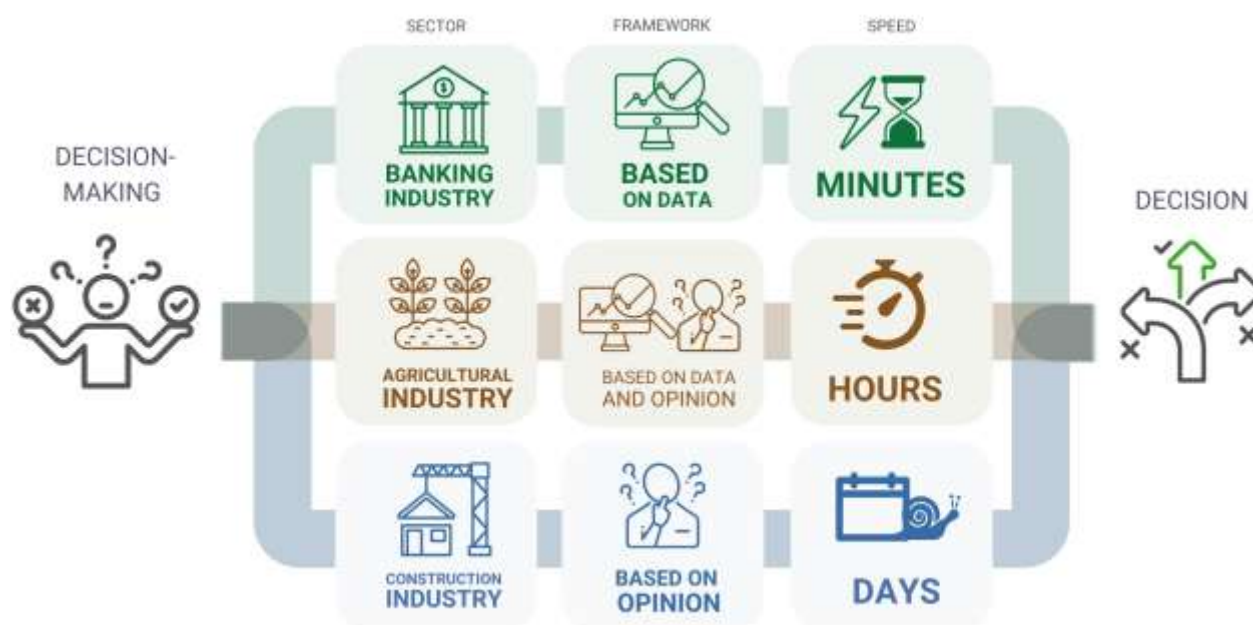
What is meant by **data-driven construction** ?



מבוא

כמה זמן תוכל החברה שלך לשמור על תחרותיות בעולם שבו הטכנולוגיות מתפתחות במהירות וכל היבט בעסק מחייב זמנים ועלויות ועד ניתוח סיכונים, מתאמת באמצעות מודלים של למידת מכונה

תעשיית הבנייה, הקיימת מאז ומעולם, עומדת בפני שינויים מהפכניים המבטיחים לשנות לחלוטין את תפיסותינו לגבי בנייה מסורתית. כבר עכשיו, במגוון תחומים אחרים בכלכלה, הדיגיטציה לא רק משנה את הכללים המוכרים, אלא גם דוחקת החוצה מהשוק חברות שלא הצליחו להסתגל לתנאים החדשים של עיבוד נתונים ואינן מסוגלות להאיץ את קצב קבלת ההחלטות (איור 1)



איור 1 קצב קבלת ההחלטות בתעשיית הבנייה תלוי יותר בגורם האנושי מאשר בתעשיות אחרות

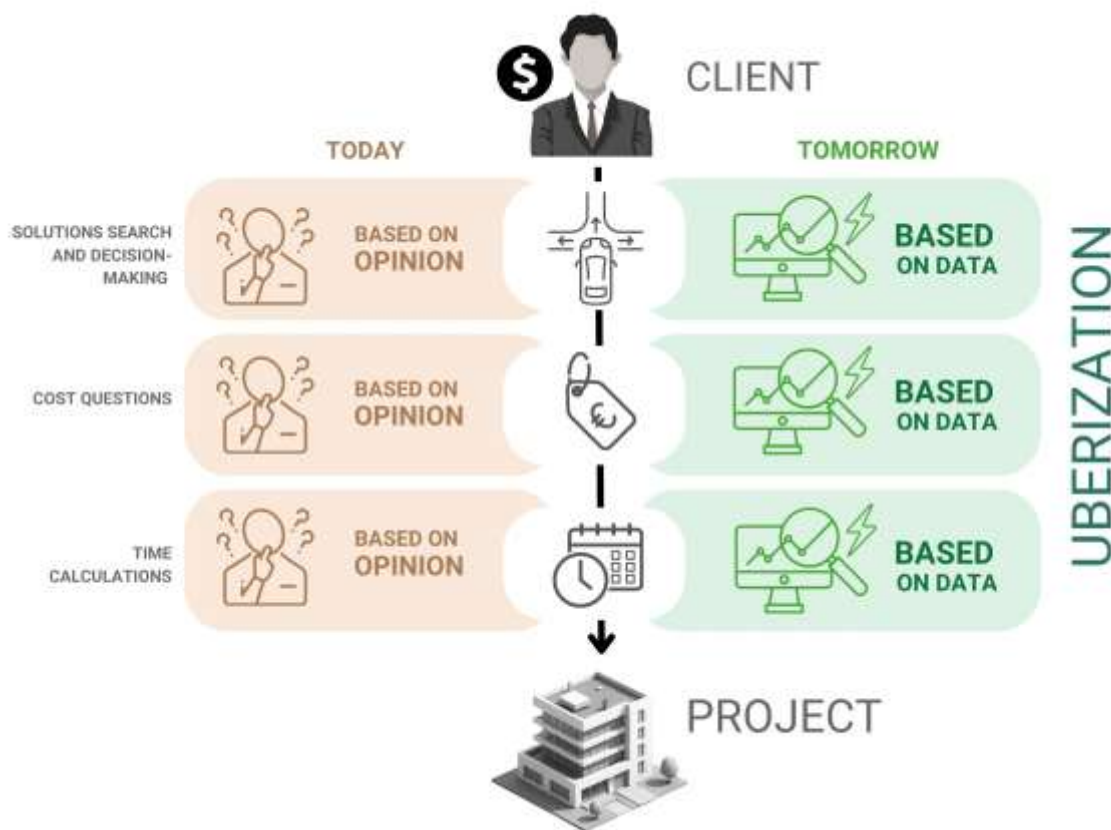
המגזר הבנקאי, הקמעונאות, הלוגיסטיקה והחקלאות מתקדמים במהירות לעבר דיגיטציה מלאה, שבה אין מקום לאי-דיוקים ודעות סובייקטיביות. אלגוריתמים מודרניים מסוגלים לנתח כמויות עצומות של נתונים ולספק ללקוחות תחזיות מדויקות - בין אם מדובר בסיכוי להחזר הלוואה, מסלולי משלוח אופטימליים או חיזוי סיכונים

הבנייה היא אחת מהתעשיות האחרונות שעומדות בפני המעבר הבלתי נמנע מפתרונות המבוססים על דעות של מומחים בעלי שכר גבוה, לפתרונות המבוססים על נתונים. המעבר הזה נובע לא רק מהזדמנויות טכנולוגיות חדשות, אלא גם מהדרישות המוגברות של השוק והלקוחות לשקיפות, דיוק ומהירות

רובטיזציה, אוטומציה של תהליכים, נתונים פתוחים וחיזויים על בסיסם - כל אלה כבר לא סתם הזדמנויות, אלא הכרח רוב החברות בתעשיית הבנייה, שהיו אחראיות לאחרונה על חישובי כמויות, עלויות, זמני פרויקטים ובקרת איכות, עכשיו בסיכון להפוך לביצועיסטים פשוטים, שלא מקבלים החלטות מפתח

עם התפתחות הכוח החישובי, אלגוריתמים של למידת מכונה ודמוקרטיזציה של גישה לנתונים, הפך האיחוד האוטומטי של נתונים ממקורות שונים לאפשרי, מה שמאפשר ניתוח מעמיק יותר של תהליכים, חיזוי סיכונים ואופטימיזציה של

עלויות כבר בשלב הדיון על פרויקט הבנייה. טכנולוגיות אלו יוצרות פוטנציאל לשיפור רדיקלי של היעילות והפחתת העלויות בכל המגזר.



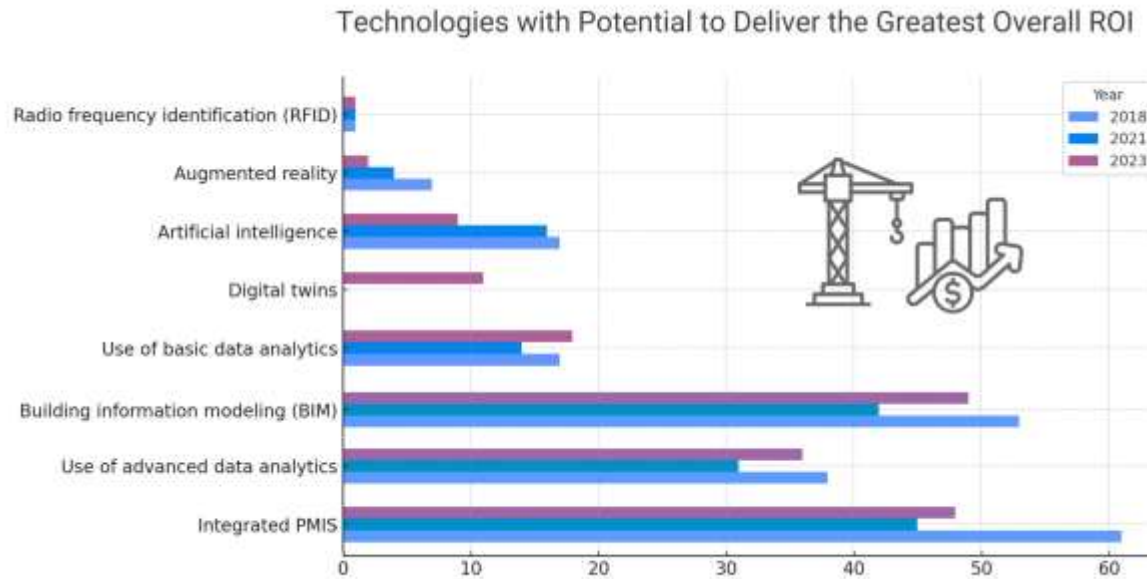
לקוח אינו מעוניין בגורם אנושי מיותר בדרך להגשמת הפרויקט שלו.

למרות כל היתרונות של הכלים והקונספטים החדשים, תעשיית הבנייה מפגרת משמעותית אחרי תעשיות אחרות בכלכלה ביישום טכנולוגיות חדשות.

בין IT תעשיית הבנייה נמצאת במקום האחרון בהוצאות על "IT Metrics Key Data 2017" לפי דוח תעשיות אחרות בכלכלה 19.

הצמיחה המהירה של כמות הנתונים ומורכבות התהליכים הופכת לבעיה עבור ההנהלה של החברות, והבעיה המרכזית בשימוש בטכנולוגיות חדשות היא שהנתונים, למרות שפעם, נשארים מפוזרים, לא מובנים ולעיתים קרובות לא תואמים בין מערכות ומוצרים תוכנה שונים. לכן, רבות מהחברות במגזר הבנייה מודאגות בראש ובראשונה מבעיות איכות הנתונים, שניתן לפתור רק באמצעות יישום של מערכות ניהול וניתוח אוטומטיות ויעילות.

בקרב מנהלי חברות בנייה בשנת 2023, הפוטנציאל הגבוה ביותר KPMG® לפי סקר שנערך על ידי ניתוח נתונים מתקדם ובסיסי (PMIS), לשיפור החזר ההשקעה בפרויקטים נמצא במערכות ניהול פרויקטים (BIM) ודימוי מידע של בניינים.



סקר בקרב מנהלי חברות בנייה: אילו טכנולוגיות יבטיחו את החזר ההשקעה הגבוה ביותר בפרויקטים קפיטליים?

פתרון הבעיות הקשורות לאינטגרציה של נתונים בתהליכי העסקים טמון בהבטחת איכות גבוהה של המידע, בשימוש בפורמטים מתאימים של נתונים וביישום שיטות יעילות ליצירה, אחסון, ניתוח ועיבוד נתונים.

ההכרה בערך הנתונים מאלצת תעשיות שונות להפסיק להשתמש ביישומים מבודדים ובמבנים בירוקרטיים מורכבים. לניהול. במקום זאת, הדגש עובר ליצירת גישות חדשות לארכיטקטורת מידע, שהופכות חברות למודרניות, מונחות נתונים גם תעשיית הבנייה תנקוט בצעד זה, ותעבור מהתפתחות דיגיטלית הדרגתית למהפכה דיגיטלית, sooner or later, אמיתית, שתשפיע על כל החברות.

המעבר לתהליכי עסקים מבוססי נתונים לא יהיה פשוט. חברות רבות יתמודדו עם קשיים, מכיוון יים של נתונים כדי לשפר את היעילות ואת $xaot$ שהמנהיגים לא תמיד מבינים כיצד להשתמש במאגרים. הצמיחה של העסק.

בספר זה אנו צוללים לעולם הנתונים, שבו המידע הופך למשאב אסטרטגי מרכזי, המגדיר את היעילות והעמידות של תהליכי העסקים. בתנאים של גידול מהיר בכמויות המידע, חברות מתמודדות עם אתגרים חדשים. טרנספורמציה דיגיטלית מפסיקה להיות מונח אופנתי – היא הופכת לצורך.



ריש. 4. נתונים ותהליכים הם הבסיס לבנייה

להבין את הטרגנספורמציה – פירושו להיות מסוגל להסביר את המורכב בפשטות. לכן הספר נכתב בשפה נגישה ומלווה באיורים מקוריים, שנוצרו במיוחד להסבר ויזואלי של מושגים מרכזיים. סכמות, גרפים וויזואליזציות אלו נועדו להסיר את מחסומי ההבנה ולהפוך את החומר לנגיש גם לאלה שחשבו בעבר שהנושאים הללו מורכבים מדי. כל האיורים, הסכמות והגרפים בספר זה נוצרו על ידי המחבר ופותחו במיוחד כדי להמחיש את המושגים המרכזיים המתוארים בטקסט.

תמונה אחת שווה אלף מילים

פרד ר. ברנרד, מאייר אנגלי, 1927 –

כדי לקשר בין התיאוריה לפרקטיקה, נשתמש בכלים של אינטליגנציה מלאכותית (בפרט, מודלים לשוניים, (המאפשרים לפתח פתרונות מבלי צורך בידע מעמיק בתכנות. אם אתם ממוקדים בחומר מעשי ומעניין אתכם יותר לעבוד עם נתונים אתם יכולים לדלג על החלק ההקדמתי ולעבור מיד לחלק השני של הספר, שבו מתחיל תיאור של דוגמאות ומקרים ספציפיים.

עם זאת, אין להטיל ציפיות מופרזות על אינטליגנציה מלאכותית, למידת מכונה וכלים של מודלים לשוניים גדולים באופן כללי. ללא נתונים איכותיים והבנה מעמיקה של התחום, גם האלגוריתמים המתקדמים ביותר לא יכולים להבטיח תוצאות אמינות ומשמעותיות.

מנכ"ל מיקרוסופט, סאטיה נדלה, בתחילת 2025, מזהיר מפני הסיכון להיווצרות בועה בתחום האינטליגנציה המלאכותית, ומשווה את ההתרגשות הנוכחית ל"בועת הדוט-קום". הוא מדגיש כי הצהרות על השגת שלבים של אינטליגנציה כללית מלאכותית ללא בסיס מתאים הן "מניפולציה חסרת משמעות של נתונים". נדלה סבור שההצלחה האמיתית של אינטליגנציה מלאכותית צריכה להימדד בתרומתה לצמיחה של התוצר המקומי הגולמי העולמי, ולא בתשומת לב מופרזת להצהרות רועשות.

מאחורי כל המילים הרועשות על טכנולוגיות חדשות ומושגים מסתתרת עבודה מורכבת ומדוקדקת להבטחת איכות

הנתונים, פרמטריזציה של תהליכי עסקים והתאמת כלים למשימות אמיתיות.

גישה מבוססת נתונים אינה מוצר שניתן פשוט להוריד או לקנות. זו אסטרטגיה שיש לבנות. היא מתחילה במבט חדש על תהליכים ובעיות קיימות, ולאחר מכן דורשת תנועה ממושמעת בכיוון הנבחר.

מפתחים מובילים של תוכנה וספקי יישומים לא יהיו הקטר של השינויים בתעשיית הבנייה; עבור רבים מהם, הגישה המבוססת נתונים מהווה איום על מודל העסקים הקיים.

תעשיות אחרות, בניגוד לבנייה, כמו תעשיית הרכב, כבר עברו שינויים רדיקליים ומזיקים, והטרנספורמציה הדיגיטלית שלהן כבר מתקדמת במלוא הקצב. חברות הבנייה צריכות לפעול במהירות ובנחישות: חברות זריזות ייהנו מרווחים עצומים, בעוד שהחברות המתנגדות יעמדו בפני סיכונים חמורים. זכרו את ההלם שהצילום הדיגיטלי גרם לתעשייה הזו. — דוח הפורום הכלכלי העולמי "עיצוב עתיד הבנייה, 2016"

החברות שיבינו בזמן את ההזדמנויות והיתרונות של הגישה החדשה יוכלו להשיג יתרון תחרותי בר קיימא ויוכלו להתפתח ולצמוח מבלי להיות תלויות בהחלטות של ספקים גדולים.

זו ההזדמנות שלכם לא רק לשרוד את הסערה הקרבה של דיגיטיזציה של מידע, אלא גם לקחת אותה תחת שליטה. בספר תמצאו לא רק ניתוח של המצב הנוכחי בתעשייה, אלא גם המלצות קונקרטיות על רה-חשיבה ורה-ארגון של התהליכים. והעסק שלכם, כדי להפוך למובילים בעידן החדש של הבנייה ולשפר את הניסיון המקצועי שלכם.

העתיד הדיגיטלי של הבנייה אינו רק שימוש בטכנולוגיות חדשות ובתוכנות, אלא רה-חשיבה יסודית על עבודה עם נתונים ומודלי עסקים.

האם החברה שלכם מוכנה לשינויים אסטרטגיים אלה?

תוכן עניינים

1.....	מבוא
I.....	תוכן עניינים
2.....	חלק מגלי חימר ועד מהפכה דיגיטלית: כיצד התפתחה המידע בבנייה II
3.....	פרק 2.1 התפתחות השימוש בנתונים בתעשיית הבנייה
3.....	לידת עידן הנתונים בבנייה
3.....	מחירם ופירוטם לטכנולוגיות דיגיטליות
5.....	תהליך ככלי לניסיון, מנוהל על ידי נתונים
7.....	דיגיטיזציה של מידע בתהליך הבנייה
11.....	פרק 2.2 טכנולוגיות ומערכות ניהול בבנייה המודרנית
11.....	מודולריות MRP/ERP המהפכה הדיגיטלית והופעת מערכות
13.....	מערכות ניהול נתונים: מהפקת נתונים למשימות עסקיות
16.....	מיצול של תאגידים: כיצד נתונים מתחברים לתהליכי עסקים
19.....	פרק 2.3 מהפכה דיגיטלית והתפוצצות הנתונים
19.....	תחילת עליית נפחי הנתונים כגל אבולוציוני
20.....	נפח הנתונים המיוצרים בחברה המודרנית
22.....	עלות אחסון הנתונים: היבט כלכלי
24.....	גבולות צבירת הנתונים: מהמון למשמעות
25.....	צעדים נוספים: מתיאוריה של נתונים לשינויים מעשיים
27.....	חלק כיצד עסקי הבנייה שוקעים בכאוס של נתונים III
28.....	פרק 3.1 פיצול וסילו של נתונים
28.....	ככל שיש יותר כלים, כך העסק יעיל יותר
30.....	סילו של נתונים והשפעתם על היעילות של החברה

33.....	שכפול וחוסר איכות הנתונים כתוצאה מהנפרדות
34.....	או הסכנה של דעות בקבלת החלטות HiPPo
37.....	עלייה מתמדת במורכבות ובדינמיות של תהליכי עסק
	המהפכה התעשייתית הרביעית) תעשייה 4.0 (והמהפכה התעשייתית החמישית) תעשייה 5.0)
39.....	בבנייה
43.....	פָּרָק 3.2. הפיכת הכאוס לסדר והפחתת המורכבות
43.....	קוד מיותר ומערכות סגורות כמחסום לשיפור הפרודוקטיביות
44.....	מסילו לאחסון נתונים מאוחד
46.....	AI מערכות אחסון משולבות מאפשרות לעבור לשימוש בסוכני
48.....	מאוסוף נתונים לקבלת החלטות: הדרך לאוטומציה
50.....	צעדים נוספים: הפיכת הכאוס למערכת ניהולית
52	חֵלֶק מסגרת הנתונים בתהליכי עסקי הבנייה IV
53.....	פָּרָק 4.1. סוגי נתונים בבנייה
53.....	סוגי הנתונים החשובים ביותר בתעשיית הבנייה
57.....	נתונים מובנים
58.....	SQL ושפת השאילתות RDBMS מסדי נתונים רלציוניים
61.....	במסדי נתונים ומגמות חדשות SQL שאילתות
63.....	נתונים לא מובנים
64.....	נתונים טקסטואליים: בין כאוס לא מובנה למבנה
65.....	נתונים חצי-מובנים ונתונים עם מבנה חלש
66.....	נתונים גיאומטריים והשימושים שלהם
69.....	מעיצוב ועד אחסון נתונים: CAD נתוני
72.....	בתהליכים CAD-והשימוש ב BIM (BOM) הופעת המושג
78.....	פָּרָק 4.2. אחידות ומבנה של נתונים
78.....	מילוי מערכות בנתונים בתעשיית הבנייה
80.....	טרנספורמציה של נתונים: יסוד קריטי לניתוח עסקי מודרני

84.....	מודלים של נתונים: יחסים בנתונים וקשרים בין רכיבים
88.....	פורמטים קנייניים והשפעתם על תהליכים דיגיטליים
91.....	פורמטים פתוחים משנים את הגישה לדיגיטציה
92.....	שינוי פרדיגמה: קוד פתוח כסוף עידן הדומיננטיות של ספקי התוכנה
94.....	נתונים פתוחים מובנים: יסוד של טרנספורמציה דיגיטלית
97.....	פֶּרֶק 4.3. מודלים לשוניים גדולים ותפקידם בעיבוד נתונים ובתהליכי עסקים
	ChatGPT, LLaMa, Mistral, Claude, DeepSeek, צ'אט מודלים לשוניים גדולים
97.....	לאוטומציה של תהליכי עיבוד נתונים QWEN, Grok
98.....	מודלים לשוניים גדולים: איך זה עובד
100	שימוש במודלים לשוניים גדולים מקומיים עבור נתונים רגישים של החברה
102	שליטה מלאה על בינה מלאכותית בחברה ואיך להפעיל מודל לשוני גדול משלך
104	עוזרי מודלים לשוניים גדולים אינטליגנטיים עם גישה לנתוני החברה: RAG
106	פֶּרֶק 4.4. עם תמיכה במודלים לשוניים גדולים ושינויים עתידיים בתכנות IDE
106	מניסויים עם מודלים לשוניים גדולים לפתרונות עסקיים: IDE בחירת
108	עם תמיכה במודלים לשוניים גדולים ושינויים עתידיים בתכנות IDE
109	כלי חיוני לעבודה עם נתונים: Python Pandas
113	פורמט אוניברסלי לנתונים טבלאיים: DataFrame
116	צעדים נוספים: יצירת מסגרת נתונים עמידה
118.....	חֶלֶק איכות נתונים: ארגון, מבנה, מודל V
119	פֶּרֶק 5.1. המרת נתונים לפורמט מובנה
119	תמונות וטקסטים לפורמטים מובנים, PDF, לומדים להמיר מסמכים
120	לטבלה PDF דוגמה להמרת מסמך
124	לפורמט מובנה JPEG, PNG המרת תמונת
126	המרת נתוני טקסט לפורמט מובנה
129	לפורמט מובנה CAD (BIM) המרת נתוני
134	עוברים לנתונים מובנים CAD ספקי פתרונות

137	פְּרָק 5.2. סיווג ואינטגרציה: שפה אחידה לנתוני בנייה.....
137	מהירות קבלת החלטות תלויה באיכות הנתונים.....
138	סטנדרטיזציה ואינטגרציה של נתונים.....
140	תאימות דיגיטלית מתחילה בדרישות.....
143	שפה אחידה לבנייה: תפקיד המדרגים בטרנספורמציה דיגיטלית.....
145	אבולוציה של מערכות סיווג: CoClass ו-Masterformat, OmniClass, Uniclass.....
150	פְּרָק 5.3. מודל נתונים ומרכז מצוינות.....
150	מודל נתונים: מודל קונספטואלי, לוגי ופיזי.....
153	מודל נתונים מעשי בהקשר של בנייה.....
155	יצירת בסיס נתונים באמצעות מודלים לשוניים גדולים.....
158	במודל נתונים (CoE) מרכז מצוינות.....
161	פְּרָק 5.4. מערכת דרישות ואימות מידע.....
161	איסוף וניתוח דרישות: המרת תקשורת לנתונים מובנים.....
165	דיאגרמות תהליכים ויעילות של תכניות קונספטואליות.....
167	RegEx דרישות מובנות וביטויים רגולריים.....
172	איסוף נתונים לתהליך האימות.....
174	אימות נתונים ותוצאות האימות.....
179	ויזואליזציה של תוצאות האימות.....
181	השוואת בדיקות איכות נתונים לצרכים חיים של בני אדם.....
183	צעדים נוספים: הפיכת נתונים לחישובים מדויקים ותוכניות.....
185	חֵלֶק חִישוּבֵי עֵלּוֹת וזֶמֶן: הַטְמַעַת נִתּוּנִים בְּתֵהֵלֵיכִי בִּנְיָה VI.....
186	פְּרָק 6.1. חִישוּבֵי עֵלּוֹת וְהַעֲרָכָה פְּרוִיקְטִים בִּבְנִיָּה.....
186	יסודות הבנייה: הערכת כמות, עלות וזמן.....
187	שיטות חישוב עלות פרויקטים.....
188	שיטת משאבים להכנת הערכות וחִישוּבֵי בִּנְיָה.....
188	בסיס נתונים של משאבי בנייה: קטלוג חומרים ועבודות בנייה.....

190	הכנת חישובים וחישוב עלות עבודות על בסיס בסיס המשאבים
194	חישוב סופי של עלות הפרויקט: מהערכות לתקציב
199	פֶּרֶק 6.2. חישוב כמות ואוטומציה של הכנת הערכות ותוכניות לוח זמנים
199	שימוש בפרמטרים נפחיים וכמותיים: D-5 ל-D-4 המעבר מ-3
199	CAD-והפקת נפחים של תכונות מ D תכונות 5
203	קיבוץ נתוני פרויקט לפי תכונות: (QTO) חישוב כמות
208	ונתונים מובנים LLM באמצעות QTO אוטומציה של
211	Excel של כל הפרויקט באמצעות כללים לקבוצות מטבלת QTO חישוב
218	CO ₂ וחישוב פליטות פחמן דו-חמצני D-8, D-6, D-4 פֶּרֶק 4.6.3
218	שילוב זמן בהערכות בנייה: D מודל 4
219	לוח זמנים לבנייה ואוטומציה שלו על בסיס נתוני חישוב
221	מיעילות אנרגטית ועד אבטחת בטיחות: D-8 שכבות תכונה מורחבות 6
224	וחישוב פליטות פחמן דו-חמצני בפרויקטים בבנייה CO ₂ הערכת
228	פֶּרֶק 6.4. בבנייה ERP ו-PMIS מערכות
228	בבנייה לדוגמת חישובים והערכות ERP מערכות
233	לאחר הבנייה ERP הקשר בין: PMIS
234	PMIS ו-ERP ספקולציות, רווח, סגירות וחוסר שקיפות ב
236	הסגורים: תעשיית הבנייה זקוקה לגישות חדשות ERP/PMIS-סוף עידן ה
238	צעדים נוספים: שימוש יעיל בנתוני פרויקטים
242	ובימ: שיווק, מציאות ועתיד נתוני הפרויקטים בבנייה CAD חלק VII
243	פֶּרֶק 7.1 בתעשיית הבנייה BIM הופעת מושגי
243	CAD כמושגים שיווקיים של ספקי BIM ו-open BIM ההיסטוריה של הופעת
246	במקום בסיסי נתונים משולבים - מערכות מודולריות סגורות: BIM מציאות
247	בתעשיית הבנייה IFC הופעת פורמט פתוח
250	כתלות בליבת הגיאומטריה IFC בעיית פורמט
252	הופעת נושא הסמנטיקה והאונטולוגיה בבנייה

254	מדוע טכנולוגיות סמנטיות אינן עומדות בציפיות בתעשיית הבנייה
257	פְּרָק 7.2. פורמטים סגורים של פרויקטים ובעיות אינטרופראביליות
257	CAD (BIM) נתונים סגורים וירידת פרודוקטיביות: פינה ללא מוצא בתעשיית
259	CAD מיתוס האינטרופראביליות בין מערכות
263	ונתונים גרנולריים USD-המעבר ל
267	פְּרָק 7.3. גיאומטריה בבנייה: מקווים לקוביות
267	מתי קווים הופכים לכסף או למה גיאומטריה חשובה לבונים
267	מקווים לנפחים: כיצד שטח ונפח הופכים לנתונים
269	ופוליגונים: שימוש בטסלציה לגיאומטריה MESH, USD-המעבר ל
271	CAD (BIM) - סיווג ייחודי של פירוט ב LOD, LOI, LOMD
273	CAD (BIM) - AIA, BEP, IDS, LOD, COBie - תקנים חדשים ל
277	CAD לעבודה עם LLM-פְּרָק 7.4. פרמטריזציה של תכנון ושימוש ב
277	הדרך לניתוח ופורמטים פתוחים: CAD (BIM) אשליית הייחודיות של נתוני
280	CAD ו-BIM תכנון באמצעות פרמטרים: עתיד
282	פרויקטים CAD בתהליכי עיבוד נתוני LLM הופעת
285	LLM ו-Pandas עם DWG ניתוח אוטומטי של קבצי
291	צעדים נוספים: המעבר מפורמטים סגורים לנתונים פתוחים
293	חֵלֶק קבלת החלטות מבוססות נתונים, ניתוח, אוטומציה ולמידת מכונה VIII
294	פְּרָק 8.1. ניתוח נתונים וקבלת החלטות מבוססות נתונים
294	נתונים כמשאב בקבלת החלטות
297	ויזואליזציה של נתונים: מפתח להבנה ולקבלת החלטות
299	KPI ו-ROI מדדי ביצוע
301	לוחות מידע ודשבורדים: ויזואליזציה של מדדים לניהול יעיל
303	ניתוח נתונים ואומנות השאלת שאלות
306	פְּרָק 8.2. ETL זרימת נתונים ללא מאמצים ידניים: מדוע נדרש
306	הפחתת עלויות והאצת העבודה עם נתונים ETL: אוטומציה של

309	איסוף נתונים.....
312	טרנספורמציה: יישום כללי בדיקה וטרנספורמציה ETL.....
315	ויזואליזציה של התוצאות בצורה של דיאגרמות וגרפים ETL: טעינת.....
320	PDF יצירה אוטומטית של מסמכי ETL: טעינת.....
321	FPDF יצירה אוטומטית של מסמכים עם ETL: טעינת.....
325	הכנת דוחות והעלאה למערכות אחרות ETL: טעינת.....
326	PDF ויזואליזציה של נתונים מתוך מסמכי LLM: באמצעות ETL.....
331	אוטומטי ETL פְּרָק 8.3. צינור.....
331	ETL צינור: קונוויור נתונים אוטומטי של.....
334	LLM באמצעות Pipeline-ETL תהליך בדיקת נתונים של.....
337	(BIM) CAD-בדיקת נתונים ומידע של רכיבי הפרויקט ב: ETL-צינור.....
343	פְּרָק 8.4 ותהליכי עבודה: פתרונות מעשיים ETL אורקסטרציה של.....
343	ואפאצ'י איירפלו: אוטומציה ואורקסטרציה של תהליכי עבודה DAG.....
344	ETL אפאצ'י איירפלו: יישום מעשי של אוטומציה בתהליכי.....
348	אפאצ'י ניפי למיפוי והמרת נתונים.....
349	אורקסטרציה של תהליכים ללא קוד, עם קוד נמוך n8n.....
352	צעדים נוספים: המעבר מפעולות ידניות לפתרונות מבוססי אנליטיקה.....
354	חֵלֶק אַחֲסוֹן וְנִיהוּל נְתוּנִים בְּבִנְיַה IX.....
355	פְּרָק 9.1. תשתית נתונים: מפורמטי אחסון ועד אחסונים דיגיטליים.....
355	אטומי נתונים: יסוד לניהול מידע יעיל.....
356	מאגר מידע: קבצים או נתונים.....
358	אחסון נתונים גדולים: ניתוח פורמטים פופולריים ויעילותם.....
360	Apache Parquet אופטימיזציה של אחסון נתונים עם.....
362	מחסן נתונים: DWH.....
363	מניקוי מסורתי לעיבוד גמיש: ELT-ל ETL אגם נתונים - אבולוציה של.....
365	סינרגיה של מחסני נתונים ואגני נתונים: Data Lakehouse אדריכלות.....

367	 ואגם נתונים DWH או CDE, PMIS, ERP
371	 פְּרָק 9.2. ניהול מחסני נתונים ומניעת כאוס.
371	 מאגרי נתונים וקטוריים ותיבת גבול.
374	 ממשלת נתונים, מינימליזם נתונים וביצת נתונים.
377	 דאטה אופס ווקטור אופס: סטנדרטים חדשים לעבודה עם נתונים.
378	 צעדים נוספים: מאחסון כאוטי לאחסונים מובנים.
380	 חֵלֶק נתונים גדולים, למידת מכונה וחזויים X
381	 פְּרָק 10.1. נתונים גדולים וניתוחם.
381	 נתונים גדולים בבנייה: מהאינטואיציה לחזוי.
382	 שאלת הכדאיות של נתונים גדולים: מתודולוגיה, סטטיסטיקה ודגימת נתונים.
384	 נתונים גדולים: ניתוח נתוני מערך של מיליון היתרי בנייה בסן פרנסיסקו.
390	 CAD (BIM) דוגמה לנתוני גדולים על בסיס נתוני.
394	 אינטרנט של הדברים וחזויים חכמים.
398	 פְּרָק 10.2. למידת מכונה וחזויים.
398	 למידת מכונה ובינה מלאכותית ישנו את האופן שבו אנו בונים.
400	 סובייקטיבית לחזוי סטטיסטי оценка מע.
402	 דאטהסט טיטניק: שלום עולם בעולם ניתוח נתונים ונתונים גדולים.
407	 למידת מכונה בפעולה: מנוסעים בטיטאניק לניהול פרויקטים.
411	 תחזיות וניבויים על בסיס נתונים היסטוריים.
413	 מונחים מרכזיים בלמידת מכונה.
416	 פְּרָק 10.3. חיזוי עלות וזמן באמצעות למידת מכונה.
416	 דוגמה לשימוש בלמידת מכונה לקביעת עלויות וזמנים של פרויקט.
418	 תחזית עלות וזמן הפרויקט באמצעות רגרסיה ליניארית.
421	 K-nearest neighbor (k-NN) תחזיות עלות וזמן הפרויקט בעזרת אלגוריתם.
425	 צעדים נוספים: מאחסון לניתוח ולחיזוי.
429	 חֵלֶק תעשיית הבנייה בעידן הנתונים הדיגיטליים. הזדמנויות ואתגרים XI

430	פָּרֶק 11.1. אסטרטגיות הישרדות: יצירת יתרונות תחרותיים
430	קורלציות במקום חישובים: העתיד של אנליטיקה בבנייה
433	גישה מבוססת נתונים בבנייה: תשתית ברמה חדשה
434	משרד דיגיטלי מהדור הבא: כיצד הבינה המלאכותית משנה את סביבת העבודה
436	נתונים פתוחים ואובריזציה הם איום על עסקי הבנייה הקיימים
439	בעיות לא פתורות של אוברייזציה כהזדמנות האחרונה לנצל את הזמן עבור טרנספורמציה
444	פָּרֶק 11.2. מדריך מעשי ליישום גישה מונחית נתונים
444	מעקרון למעשה: מפת דרכים להמרה דיגיטלית בענף הבנייה
446	מניחים את היסוד הדיגיטלי: 1-5 צעדים לבגרות דיגיטלית
450	מגלה את הפוטנציאל של נתונים: 5-10 צעדים לבגרות דיגיטלית
456	מפת דרכים לשינוי: מהכאוס לחברה מונחית נתונים
459	בנייה בתעשייה 5.0: כיצד להרוויח כאשר כבר אי אפשר להסתיר יותר
461	סיכום
464	על המחבר
465	משוב
465	הערה על התרגום
466	מיומנויות ומושגים נוספים
470	מילון מונחים
476	רשימת ספרות וחומרים מקוונים
494	אינדקס נושאים

מקסימום נוחות עם הגרסה המודפסת

לשימוש נוח יותר וגישה מהירה. Data-Driven Construction אתם מחזיקים בידיכם את הגרסה הדיגיטלית החינמית של לחומרים, אנו ממליצים לשקול את הגרסה המודפסת



■ תמיד בהישג יד: ספר בפורמט מודפס יהפוך לכלי עבודה אמין המאפשר למצוא ולהשתמש במהירות בויזואליזציות ובתרשימים בכל מצב עבודה

■ איכות גבוהה של איורים: כל התמונות והגרפים בגרסה המודפסת מוצגים באיכות מקסימלית

■ גישה מהירה למידע: ניווט נוח אפשרות לעשות הערות, סימניות ולעבוד עם הספר בכל מקום

ברכישת הגרסה המודפסת המלאה של הספר, אתם מקבלים כלי נוח לעבודה נוחה ויעילה עם מידע: אפשרות להשתמש בחומרים ויזואליים במשימות יומיומיות, למצוא במהירות את התרשימים הנדרשים ולעשות הערות. בנוסף, רכישתכם תומכת בהפצת ידע פתוח.

datadrivenconstruction.io/books: ניתן להזמין את הגרסה המודפסת של הספר בכתובת



חלק I

מגלי חימר ועד מהפכה דיגיטלית: כיצד התפתחה המידע בבנייה

בחלק הראשון של הספר נדונה האבולוציה ההיסטורית של ניהול הנתונים בתעשיית הבנייה מהקלטות פרימיטיביות על חומרים פיזיים ועד מערכות דיגיטליות מודרניות. ניתוח – והשפעת פיצול הנתונים על ERP המהפכה בטכנולוגיות ניהול המידע, הופעת מערכות היעילות של תהליכי העסקים. תשומת לב מיוחדת ניתנת לתהליך הדיגיטציה של המידע ולחשיבות הגוברת של ניתוח אובייקטיבי במקום הערכות מומחים סובייקטיביות. נדונה גם, הצמיחה האקספוננציאלית של כמויות המידע שבהן נתקלת תעשיית הבנייה המודרנית, ואתגרים הקשורים לכך עבור מערכות תאגידיות. נבחנת מיקום תעשיית הבנייה בהקשר של המהפכה התעשייתית הרביעית והחמישית, כמו גם הפוטנציאל של שימוש בבינה מלאכותית ובגישות מונחות נתונים ליצירת יתרונות תחרותיים ברי קיימא.

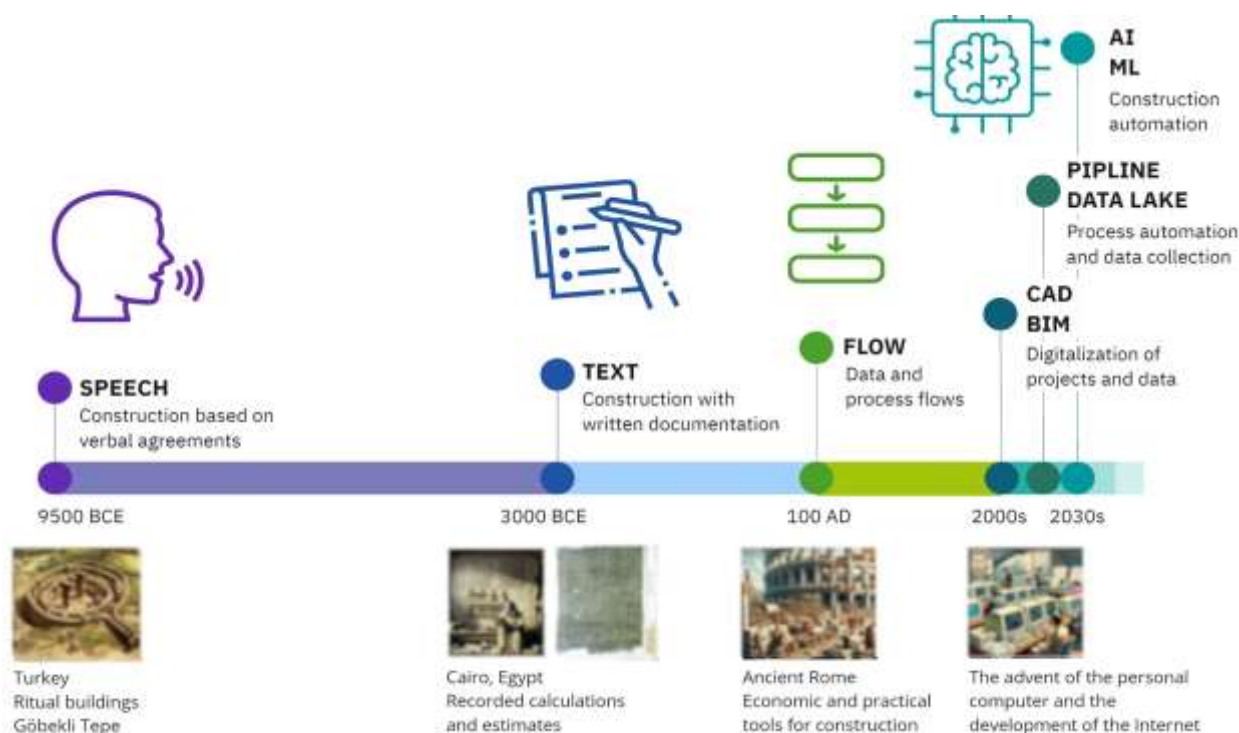
פרק 1.1.

התפתחות השימוש בנתונים בתעשיית הבנייה

לידת עידן הנתונים בבנייה

לפני כ-10,000 שנה, בעידן הניאוליתי, האנושות עשתה מעבר מהפכני בהתפתחותה, כאשר היא נטשה את אורח החיים הנודד לטובת התיישבות, מה שהוביל להופעת הבניינים הראשונים הפרימיטיביים מחרס, עץ ואבן.

עם התפתחות הציביליזציות, האדריכלות הפכה למורכבת יותר, מה שהוביל להופעת המקדשים והבניינים הציבוריים הראשונים. המורכבות הגוברת של הפרויקטים האדריכליים דרשה מהמהנדסים וממנהליהם של העבר ליצור את ההקלטות והחישובים הראשונים. ההקלטות הראשונות על לוחות חרס ועל פפירוסים כללו לעיתים קרובות תיאור של הלוגיקה של חישוב כמות החומרים הבנייתיים הנדרשים, עלותם וחישוב תשלום העבודה שבוצעה. כך החלה עידן השימוש בנתונים בבנייה – הרבה לפני הופעת הטכנולוגיות הדיגיטליות המודרניות.



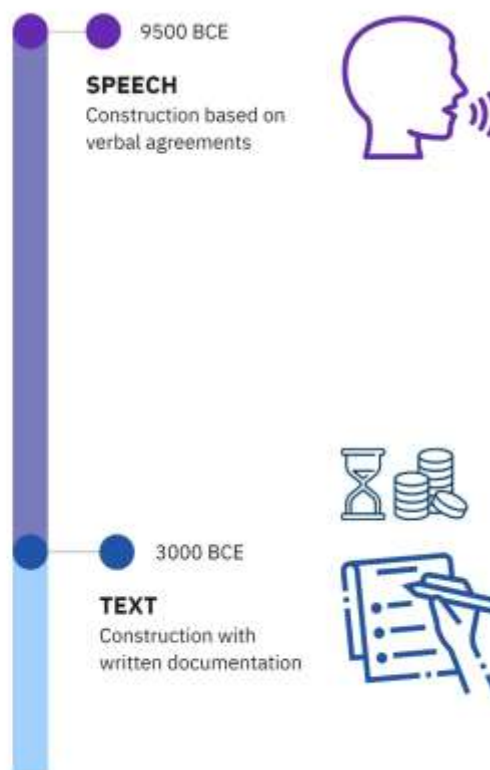
כרונולוגיה של התפתחות טכנולוגיות המידע בבנייה: מידע מילולי ועד אינטליגנציה מלאכותית 1.11 איור

מחירם ופירוס לטכנולוגיות דיגיטליות

העדויות הדוקומנטריות הראשונות בבנייה מתייחסות לתקופת בניית הפירמידות, בסביבות 3000–4000 לפני הספירה.

מאז, ניהול רשומות כתובות הקל והנחה את ההתקדמות בתעשיית הבנייה, מה שאפשר לצבור ולסדר ידע, שהוביל במהלך השנים הבאות לחדשנות משמעותית בשיטות הבנייה ובאדריכלות 10,000

"A0" או נייר בגודל (1.12 איור) השימוש בנושאים פיזיים ראשונים בבנייה, כמו לוחות חימר, פפירוס בן אלף שנה בשנות ה-80, לרישום נתונים לא התכוון במקור לשימוש במידע הזה בפרויקטים חדשים. המטרה העיקרית של רשומות אלו הייתה תיאור מפורט של מצב הפרויקט הנוכחי, כולל חישובים של חומרים נדרשים ועלויות עבודה. באותו אופן בעולם המודרני, קיום נתוני פרויקט דיגיטליים ודגמים לא תמיד מבטיח את השימוש שלהם בפרויקטים עתידיים ולעיתים קרובות הם משמשים בעיקר כמידע לחישובים הנוכחיים של חומרים נדרשים ועלויות הבנייה.



פפירוס מהמאה ה-3 לפני הספירה, המתאר את עלות הציור של סוגי חלונות שונים בארמון המלכותי באמצעות 1.12 איור טכניקת אנקאוסטיקה.

לאנושות נדרשו כ-5,000 שנה לעבור משיחות בעל פה למסמכים כתובים בניהול פרויקטים בבנייה, וכך גם נדרשו כמות זמן דומה לעבור מנושאים נייריים לנתונים דיגיטליים כמקור העיקרי לתכנון ובקרה.

בדומה לכך, כיצד התפתחות הקשרים המסחריים והכספיים עוררה את הופעת הכתיבה והעורכי דין הראשונים, אשר פתרו בעיות שנוי במחלוקת, כך גם הרשומות הראשונות על עלויות חומרים והיקפי עבודה בבנייה הובילו להופעת המנהלים הראשונים בתעשיית הבנייה, אשר היו אחראים על תיעוד, ניטור ואחריות על מידע מפתח לגבי לוחות זמנים ועלויות הפרויקט.

כיום, הנתונים משחקים תפקיד הרבה יותר משמעותי: הם לא רק מתעדים את ההחלטות שהתקבלו, אלא גם הופכים לכלי לחיזוי ודימוי העתיד. על בסיס זה נבנה הגישה המודרנית לניהול פרויקטים - הפיכת הניסיון המצטבר למערכת קבלת החלטות, המבוססת על נתונים מובנים ומאומתים

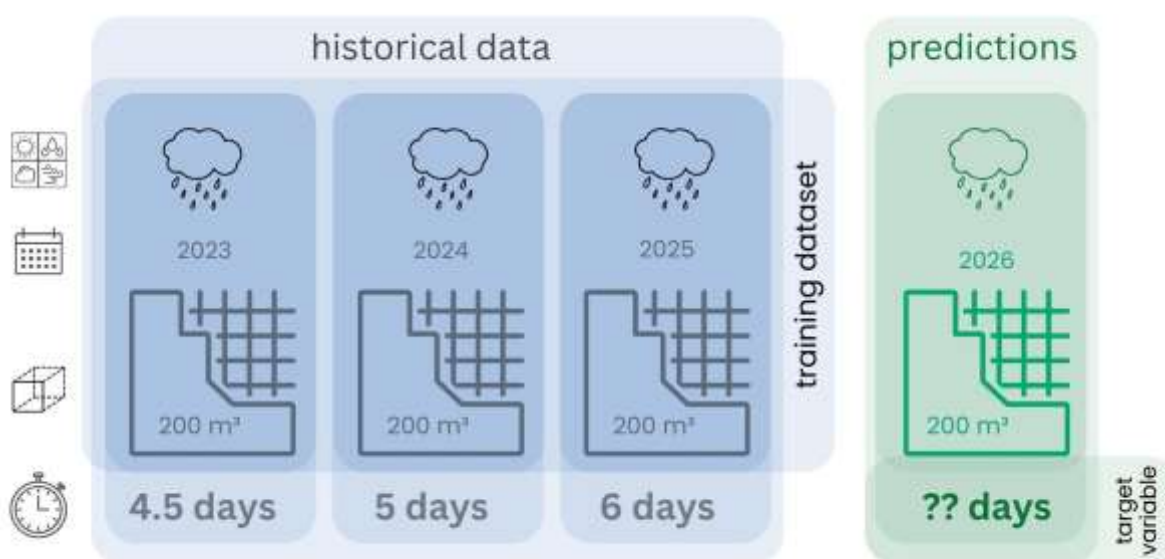
תהליך ככלי לניסיון, מנוהל על ידי נתונים

בבסיס כל תהליך טמונה טרנספורמציה של ניסיון העבר לכלי לתכנון העתיד. ניסיון במובן המודרני מייצג סט נתונים מובנה, שהניתוח שלו מאפשר לבצע תחזיות מבוססות

נתוני העבר הם הבסיס לחיזוי, שכן הם מדגימים בצורה ברורה את תוצאות העבודות שבוצעו ומספקים תובנות לגבי הגורמים המשפיעים על תוצאות אלו

נבחן דוגמה ספציפית מתוך בנייה מונוליתית: בדרך כלל, כאשר מתכננים את לוחות הזמנים של העבודות, מתחשבים בנפח הבטון, במורכבות המבנה ובתנאי מזג האוויר. נניח כי למנהל העבודה באתר הבנייה או לנתונים היסטוריים של pour- (החברה, בשנים האחרונות) 2023–2025 (יש נתונים המצביעים על כך שדרוש בין 4.5 ל-6 ימים במקרה זה, הערכת הזמן - התהליך האנליטי פועל כמנגנון להמרת נתונים מפורזים לניסיון מובנה, ולאחר מכן - לכלי תכנון מדויק. הנתונים והתהליכים הם מערכת אקולוגית אחת, שבה אחד לא יכול להתקיים ללא השני

חשבו את מה שניתן לספור, מדדו את מה שניתן למדוד, ואת מה שאינו ניתן למדידה, הפכו אותו למדיד. גלילאו גליליי



בעתיד מהכמויות אחת לחיזוי אימוני נתונים כסט משמשים היסטוריים נתונים 1.13 איור.

וקבלת תהליכים של אופטימיזציה, יעיל פרויקטים בניהול קריטי לרכיב הופך נתונים ניתוח, המודרני העסקי בעידן

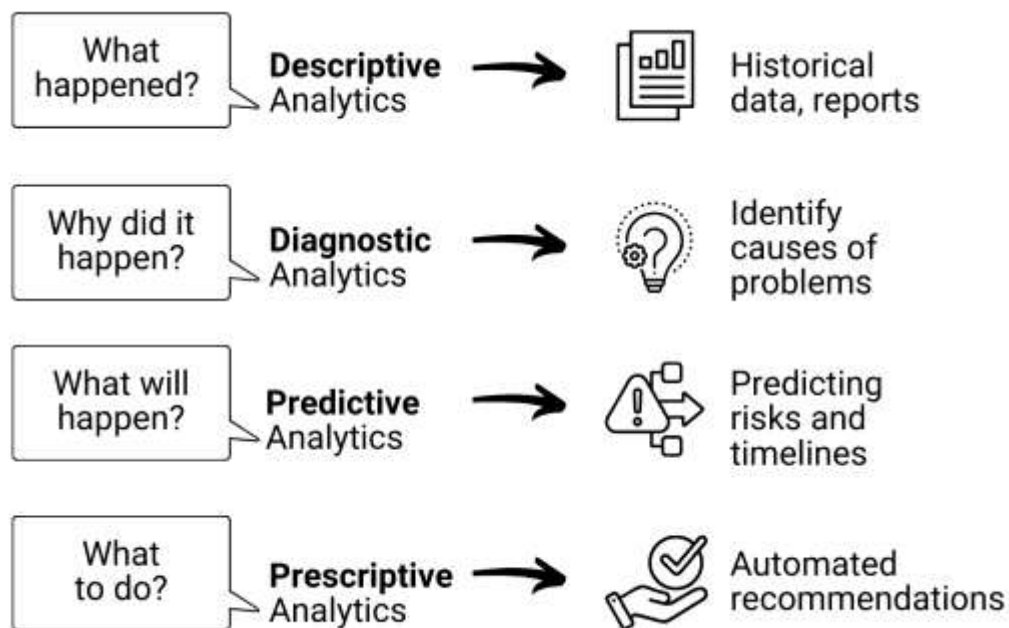
מהן אחת כל כאשר, ניתוח של מפתח רמות ארבעה של בהבנה בהדרגה מתקדמת הבנייה תעשיית. אסטרטגיות החלטות - ייחודיים יתרונות ומספקת ספציפית שאלה על עונה

ותוצאות אירועים על ודוחות היסטוריים נתונים ומספקת "קרה מה" השאלה על עונה - תיאורית אנליטיקה ר"מ 200 של בשטח מונוליטי מבנה יציקת דרשה (2023-2025) האחרונות השנים שלוש במהלך: מהעבר ימים 6-4.5 בין גשום אוויר במזג

הניתוח: בעיות להיווצרות הסיבות את חושפת, "קרה זה למה" השאלה על עונה - דיאגנוסטית אנליטיקה של ההתקשות תהליך את שהאט, הגשום האוויר מזג בגלל התאריך המונוליטי המבנה של היציקה שזמן מראה הבטון.

:"יקרה מה" השאלה על עונה, ביצוע וזמני פוטנציאליים סיכונים מנבאת, לעתיד מכוונת - חזויית אנליטיקה גשום אוויר במזג ר"מ 200 של בשטח דומה מונוליטי מבנה יציקת כי מנבאים, היסטוריים נתונים על בהתבסס הידועים והנטיות הגורמים בכל התחשבות תוך, ימים 5.5-כ תדרוש 2026 בשנת

לחברות שמאפשר מה, "לעשות מה" השאלה על ועונה אוטומטיות המלצות מספקת פרסקריפטיבית אנליטיקה מיוחדים בתוספים להשתמש מומלץ, למשל, העבודה של אופטימיזציה לצורך. האופטימליות בפעולות לבחור ביותר הנמוכה הסבירות עם בתקופות היציקה את לתכנן; גבוהה לחות בתנאי הבטון התייבשות להאצת מזג בתנאי גם ימים 4.5-4-ל העבודה זמן את לקצר שיאפשר מה, למבנה זמניים כיסויים לארגן; לגשמים נוחים לא אוויר



סוגי האנליטיקה העיקריים: מתיאור העבר ועד תהליך קבלת החלטות אוטומטי 1.14. ריש

טרנספורמציה דיגיטלית מלאה, המניחה מעבר לאנליטיקה מערכתית ולניהול מבוסס נתונים, דורשת לא רק גיוס קבלנים, חיצוניים, אלא גם יצירת צוות פנימי מיומן. המשתתפים המרכזיים בצוות כזה צריכים להיות מנהלי מוצר, מהנדסי נתונים

אנליסטים ומפתחים, שעובדים בשיתוף פעולה הדוק עם מחלקות עסקיות. שיתוף פעולה זה הכרחי כדי לנסח שאלות אנליטיות נכונות ולבצע פרמטריזציה יעילה של בעיות עסקיות לקבלת החלטות. בתנאים של חברה מידע, הנתונים הופכים -לא רק לכלי עזר, אלא לבסיס לחיזוי ואופטימיזציה

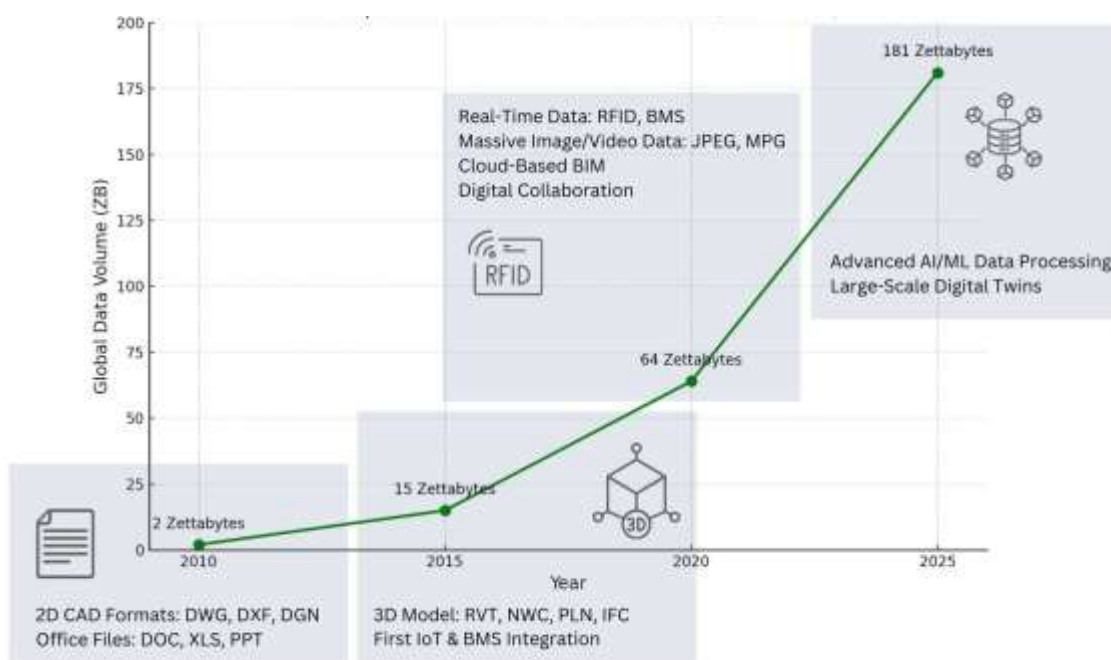
בתחום הבנייה, הטרנספורמציה הדיגיטלית משנה באופן יסודי את הגישות לתכנון, ניהול ותפעול של פרויקטים. תהליך זה נקרא דיגיטליזציה של מידע - כאשר כל ההיבטים של תהליך הבנייה מועברים לצורת נתונים דיגיטלית, המתאימה לניתוח.

דיגיטליזציה של מידע בתהליך הבנייה

במשך אלפי שנים, כמות המידע שנרשמה בתחום הבנייה כמעט ולא השתנתה, אך בעשורים האחרונים היא צמחה במהירות.

ניהול נתונים. מה הסטודנטים צריכים כדי להצליח בעולם העסקים המשתנה? PwC® לפי מחקר של (במהירות) "2015, 90% מכל הנתונים בעולם נוצרו בשנתיים האחרונות) נכון לשנת 2015. (עם זאת רוב החברות לא מנצלות את הנתונים הללו במלואם, מכיוון שהם או נשארים במערכות מבודדות, או פשוט מאוחסנים בארכיון ללא ניתוח ממשי

בשנים האחרונות, הגידול בכמות הנתונים רק האיץ, והוכפל מ-15 זטאבייט בשנת 2015 ל-181 זטאבייט בשנת 2025. מדי יום, השרתים של חברות הבנייה והתכנון מתמלאים במסמכי פרויקט, לוחות זמנים, חישובים ודוחות כספיים. עבור RVT, NWC, PLN - ועבור מודלים תלת-ממדיים, DGN ו-DXF, DWG, משתמשים בפורמטים D/3D שרטוטים 2 בנוסף, ישנם סרטונים PPT ו-XLSX, DOC מסמכים טקסטואליים, גיליונות עבודה ומצגות נשמרים בפורמטים IFC™. זיהוי) RFID תוויות, IoT ובזמן אמת מתקבלים נתונים מרכיבי JPEG ו-MPG ודימויים מאתר הבנייה - בפורמטים (ניטור ושליטה) BMS ומעקב (ומערכות ניהול בניינים



הצמיחה הפרבולית של כמות הנתונים 2010–2025. ריש.

בתנאים של גידול מהיר בכמויות המידע, תעשיית הבנייה מתמודדת עם הצורך לא רק לאסוף ולאחסן נתונים, אלא גם להבטיח את בדיקתם, אימותם, מדידותיהם ועיבוד אנליטי שלהם. כיום, התעשייה חווה שלב פעיל של דיגיטליזציה של מידע - המרה מערכתית של כל ההיבטים של פעילות הבנייה לצורת נתונים דיגיטלית, המתאימה לניתוח, פרשנות ואוטומציה.

דיגיטליזציה של מידע פירושה קבלת מידע על כל הישויות והאלמנטים של פרויקט הבנייה ושל תהליך הבנייה עצמו - כולל אלה שלא נחשבו בעבר למידע - והמרתן לפורמט נתונים, כדי להפוך את המידע לכמותי ומועיל לניתוח.

בהקשר של בנייה, זה מתייחס לתיעוד והבעה במספרים של מידע על כל רכיבי הפרויקטים וכל התהליכים - מהעברת ציוד ואנשים באתר הבנייה ועד לתנאי מזג האוויר והאקלים באתר, מחירי החומרים הנוכחיים וריבית הבנק המרכזי במטרה ליצור מודלים אנליטיים.

אם אתה יכול למדוד את מה שאתה מדבר עליו ולהביע את זה במספרים - זה אומר שאתה יודע משהו על הנושא הזה. אבל אם אתה לא יכול לבטא את זה כמותית, הידע שלך מוגבל מאוד ואינו מספק. אולי זה שלב ראשוני, אבל זה לא רמת ידע מדעי אמיתי. ו. תומסון (לורד קלווין, 1824–1907, מדען בריטי)

דיגיטציה של מידע חורגת הרבה מעבר לגישה המסורתית לאיסוף מידע, שבה נרשמו רק מדדים בסיסיים - כמו שעות אדם או עלויות חומר בפועל. היום כמעט כל אירוע יכול להיות מומר לזרם נתונים שניתן לניתוח מעמיק באמצעות כלים של אנליטיקה מתקדמת ושיטות של למידת מכונה. בתעשיית הבנייה התרחש שינוי מהותי: משרטטים על נייר, תקציבי

- והנחיות בעל פה - למערכות דיגיטליות, שבהן כל רכיב של האובייקט הופך למקור נתונים. אפילו העובדים Excel מהמהנדסים ועד הבנאים באתר - נחשבים כעת כאוסף של משתנים דיגיטליים וסטים של נתונים.

בעיות מוכרות - גישות חדשות: סקירה עולמית של הבנייה לשנת 2023, "דיגיטליים" KPMG לפי נתוני הופכים לדחפים מרכזיים להעלאת רווחיות, Big Data ו-AI) כפולים, אינטליגנציה מלאכותית הפרויקטים.

טכנולוגיות מודרניות לא רק שמפשטות את איסוף המידע, מה שהופך אותו במידה רבה לאוטומטי, אלא גם מפחיתות באופן רדיקלי את עלות אחסון הנתונים. כתוצאה מכך, חברות מוותרות על גישה סלקטיבית ומעדיפות לשמור את כל מאגר המידע לניתוח עתידי, מה שפותח הזדמנויות פוטנציאליות לאופטימיזציה של תהליכים בעתיד.

דיגיטציה של מידע ודיגיטיזציה מאפשרות לחשוף ערך סמוי, שלא נעשה בו שימוש בעבר. כאשר הנתונים מאורגנים כראוי, הם מקבלים חיים חדשים: ניתן לעשות בהם שימוש חוזר, להרהר בהם מחדש ולשלבם בשירותים ובפתרונות חדשים.

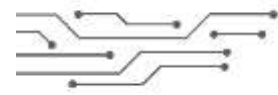
בעתיד, דיגיטציה של מידע תוביל ככל הנראה לאוטומציה מלאה של ניהול המסמכים, ליישום תהליכי בנייה עצמאיים ומהנדסים דיגיטליים AI ולהופעת מקצועות חדשים - אנליסטים של נתוני בנייה, מומחים לניהול פרויקטים באמצעות אובייקטי הבנייה יהפכו למקורות מידע דינמיים, וההחלטות יתבססו לא על אינטואיציה או ניסיון סובייקטיבי, אלא על עובדות דיגיטליות מהימנות וניתנות לשחזור.

מידע הוא הנפט של המאה ה-21, והאנליטיקה היא מנוע הבעירה הפנימית. פיטר סונדרגארד, סגן נשיא בכיר של Gartner®.

לשנת 2024, צפוי כי ההוצאות הגלובליות על ניהול נתונים ואנליטיקה IoT Analytics בהתאם לנתוני יגדלו באופן חד משמעי מ-185.5 מיליארד דולר בשנת 2023 ל-513.3 מיליארד דולר עד שנת 2030. כאשר שיעור הצמיחה השנתי הממוצע יעמוד על 16%. עם זאת, לא כל המרכיבים צומחים באותם קצבים: האנליטיקה מתפתחת במהירות, בעוד שהצמיחה של מערכות אחסון הנתונים מתמתנת. האנליטיקה תספק את הצמיחה המהירה ביותר במערכת ניהול הנתונים: על פי תחזיות, היקפה יגדל מ-60.6 מיליארד דולר בשנת ל-227.9 מיליארד דולר עד שנת 2030, מה שמתאים לשיעור צמיחה שנתי ממוצע של 27% 2023.

עם האצת תהליך הדיגיטציה של המידע והצמיחה המהירה של כמויות המידע, מנהלי פרויקטים וחברות בנייה מתמודדים עם הצורך לאחסון, לנתח ולעבד באופן שיטתי נתונים מגוונים ולעיתים לא אחידים. בתגובה לאתגר זה, החל משנות ה-20 של המאה ה-20, התעשייה החלה לעבור באופן נרחב ליצירה, אחסון וניהול מסמכים באופן אלקטרוני - החל 90 מגיליונות עבודה וחישובים פרויקטיים ועד שרטוטים וחוזים.

מסמכים מסורתיים מודפסים, אשר דרשו חתימות, אחסון פיזי, סקירה קבועה וארכוב בארונות, נדחקים בהדרגה על ידי מערכות דיגיטליות, שבהן הנתונים נשמרים בצורה מסודרת - בבסיסי נתונים של יישומים ייחודיים.



פרק 1.2.

טכנולוגיות ומערכות ניהול בבנייה המודרנית

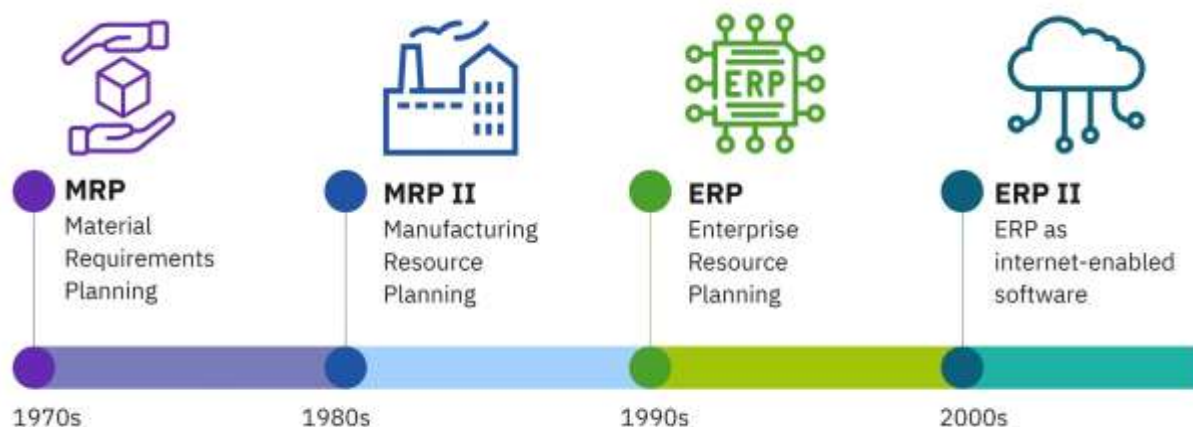
מודולריות MRP/ERP המהפכה הדיגיטלית והופעת מערכות

עידן האחסון והעיבוד המודרני של נתונים בצורה דיגיטלית החל עם הופעת סרטי המגנטי בשנות ה-50, אשר פתחו את האפשרות לאחסן ולהשתמש בכמויות גדולות של מידע. הפריצה הבאה הייתה הופעת כונני הדיסק, ששינו באופן רדיקלי את הגישה לניהול נתונים בתעשיית הבנייה.

עם התפתחות מחסני הנתונים, נכנסו לשוק פתרונות רבים מחברות שונות, אשר החלו לפתח תוכנה מודולרית ליצירה, אחסון, עיבוד נתונים ואוטומציה של משימות שגרתיות.

הצמיחה האקספוננציאלית של כמות המידע והכלים הביאה לצורך בפיתוח פתרונות מודולריים משולבים, אשר אינם פועלים עם קבצים בודדים, אלא מסייעים לנהל ולשלוט בזרימת הנתונים במסגרת תהליכים ופרויקטים שונים.

הכלים הראשונים של הפלטפורמה המורכבת לא רק היו צריכים לאחסן מסמכים, אלא גם לתעד את כל הבקשות לשינויים ופעולות בתהליכים: מי יזם אותן, מה היה היקף הבקשה ומה בסופו של דבר נרשם כערך או מאפיין. למטרות אלו נדרשה תכנון MRP מערכת שיוכלה לעקוב אחרי חישובים מדויקים והחלטות שהתקבלו. פלטפורמות כאלה הפכו למערכות (תכנון משאבים ארגוניים, אשר זכו לפופולריות מתחילת שנות ה-90 ERP-דרישות חומר ו).

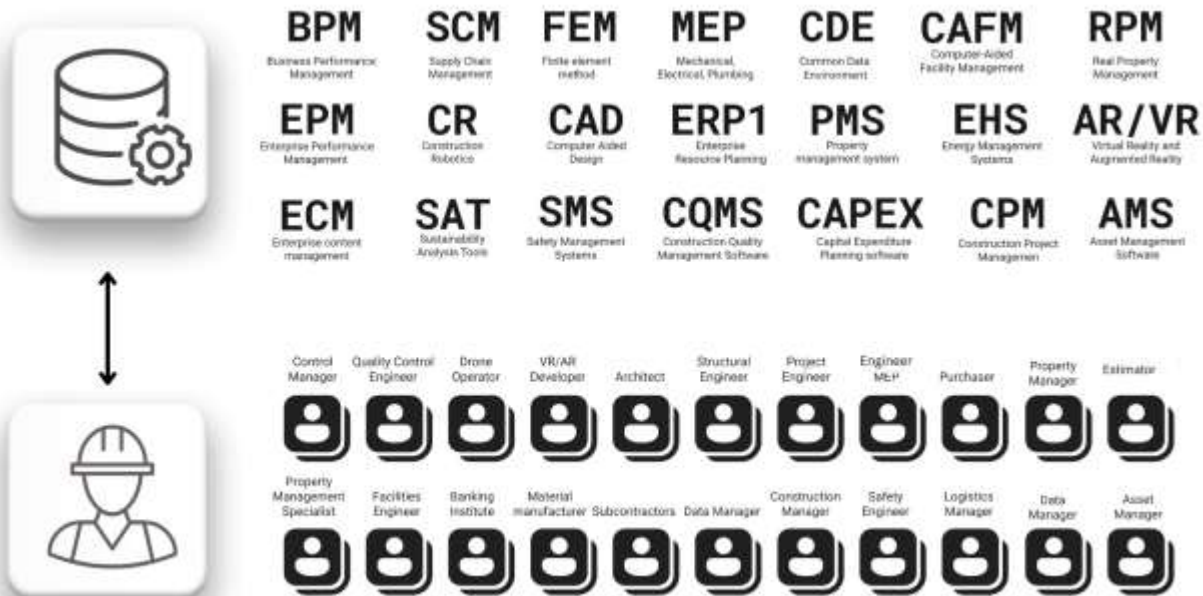


- בשנות ה-80 ERP הישגים בתחום טכנולוגיות האחסון הביאו להופעת מערכות

הראשונות הניחו את היסודות לעידן הדיגיטציה בניהול תהליכי עסקים ופרויקטים בבנייה ERP ו-MRP מערכות מערכות מודולריות, אשר נועדו במקור לאוטומציה של תהליכי עסקים מרכזיים, החלו עם הזמן להשתלב עם פתרונות תוכנה נוספים, גמישים ומסתגלים יותר.

פתרונות נוספים אלו נועדו לעיבוד נתונים ולניהול תוכן פרויקטים, הם או החליפו מודולים מסוימים במערכות הגדולות,

או שהשלימו אותם ביעילות, והרחיבו את הפונקציונליות של כל המערכת.



איור 1.22 פתרונות תוכנה חדשים משכו לעסק צבא שלם של מנהלים לניהול זרמי נתונים.

בעשורים האחרונים חברות השקיעו סכומים משמעותיים במערכות מודולריות, רואות בהן פתרונות אינטגרטיביים לטווח ארוך.

עומד על ERP לשנת 2022, התקציב הממוצע למשתמש במערכת Software Path לפי נתוני דוח דולר אמריקאי. בממוצע, כ-26% מעובדי החברה משתמשים במערכות אלו. כך, עבור ארגון עם 9,000 מגיעות לכ-900,000 דולר ERP משתמשים, העלויות הכוללות ליישום 100

השקעות בפתרונות מודולריים סגורים הופכות לפחות מוצדקות לאור ההתפתחות המהירה של טכנולוגיות מודרניות גמישות ופתוחות. אם השקעות כאלה כבר נעשו, חשוב להעריך מחדש את תפקיד המערכות הקיימות: האם הן עדיין נחוצות בטווח הארוך, או שניתן לשקול את הפונקציות שלהן וליישם אותן בצורה יעילה ושקופה יותר.

אחת הבעיות המרכזיות של פלטפורמות מודולריות עכשוויות לעיבוד נתונים היא שהן מרכזות את ניהול הנתונים בתוך אפליקציות סגורות. כתוצאה מכך, הנתונים - הנכס המרכזי של החברה - תלויים בפתרונות תוכנה ספציפיים ולא להפך. זה מגביל את האפשרויות לשימוש חוזר במידע, מקשה על הגירה ומפחית את הגמישות של העסק בתנאים של נוף דיגיטלי המשתנה במהירות.

אם יש סבירות שהחשיבות או הביקוש לארכיטקטורה מודולרית סגורה יפחתו בעתיד, יש מקום להכיר כבר היום בהוצאות שנעשו כהוצאות בלתי הפיכות ולהתמקד במעבר אסטרטגי למערכת אקולוגית דיגיטלית פתוחה, ניתנת להרחבה ומסתגלת יותר.

תוכנה קניינית מאופיינת בשליטה מוחלטת של חברת הפיתוח על הקוד המקורי והנתונים של המשתמשים הנוצרים במהלך השימוש בפתרונות כאלה. בניגוד לתוכנות קוד פתוח, למשתמשים אין גישה למבנה הפנימי של האפליקציה ואינם יכולים לבדוק, לשנות או להתאים אותה לצרכיהם. במקום זאת, הם מחויבים לרכוש רישיונות המעניקים את הזכות להשתמש בתוכנה בהתאם למגבלות שנקבעו על ידי הספק.

הגישה המודרנית הממוקדת בנתונים מציעה פרדיגמה שונה: נתונים צריכים להישב כנכס אסטרטגי מרכזי - עצמאי ארוך טווח ומופרד מפתרונות תוכנה ספציפיים. האפליקציות, מצידן, הופכות לכלים לעבודה עם נתונים, שניתן להחליפן בחופשיות מבלי לסכן את אובדן המידע הקריטי.

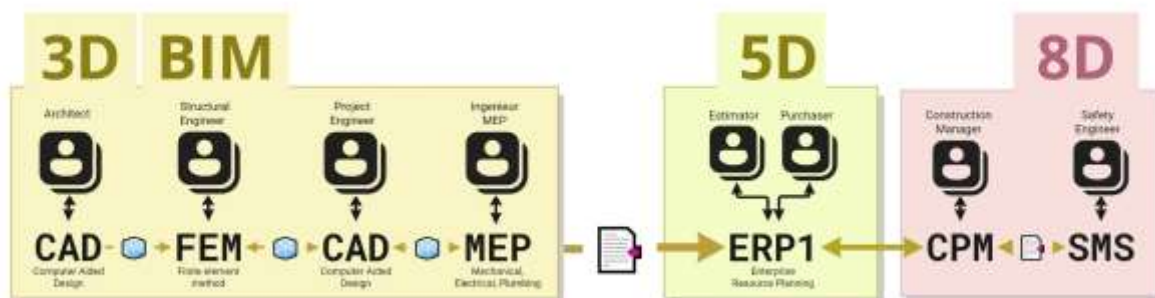
בשנות ה-90 סיפקה לעסקים כלים רבי עוצמה לניהול תהליכים, אך הביאה גם MRP ו-ERP להתפתחות של מערכות לתוצאה בלתי צפויה - עלייה משמעותית במספר העובדים העוסקים בתחזוקת זרמי המידע. במקום לאוטומט ולפשט את IT המשימות התפעוליות, מערכות אלו לעיתים קרובות יצרו רמות חדשות של מורכבות, בירוקרטיה ותלות במשאבי פנימיים.

מערכות ניהול נתונים: מהפכת נתונים למשימות עסקיות

חברות מודרניות מתמודדות עם הצורך לשלב מספר מערכות לניהול נתונים. הבחירה במערכות ניהול נתונים, ניהול נכון של מערכות אלו ושילוב מקורות נתונים מפוזרים הופכים למשימה קריטית עבור היעילות של העסק.

באמצע שנות ה-2020 ניתן למצוא מאות (ובחברות גדולות אלפי (מערכות שונות בחברות בנייה בינוניות) איור 1.23) אשר צריכות לפעול בהרמוניה כדי שכל היבטי תהליך הבנייה יתנהלו בצורה חלקה ומסודרת.

ניהול מבוסס נתונים בפרויקטים קפיטליים דיגיטליים "משנת 2016 - איש" Deloitte® לפי מחקר של מקצוע בתחום הבנייה משתמש במוצע ב-3.3 יישומים תוכנתיים ביום, אך רק 1.7 מהם משולבים זה בזה.



איור 1.23 כל מערכת עסקית דורשת צוות מקצועי ומנהל אחראי לניהול איכותי של הנתונים.

להלן רשימה של מערכות פופולריות עבור חברות בינוניות וגדולות בתחום הבנייה, אשר משמשות לניהול יעיל של פרויקטי בנייה:

- (תכנון משאבים ארגוניים) – מספקת אינטגרציה של תהליכים עסקיים, כולל חשבונאות, רכש וניהול ERP פרויקטים.
- (תוכנת תכנון הוצאות הון) – משמשת להכנת תקציב וניהול השקעות כספיות בפרויקטי בנייה CAPEX מסייעת לקבוע את עלות הנכסים הקבועים וההשקעות בנכסים ארוכי טווח.
- (מודל מידע בנייתי) – משמשת ליצירת שרטוטים טכניים מפורטים (BIM-עיצוב בעזרת מחשב ו CAD ומודלים תלת-ממדיים של פרויקטים. המיקוד במערכות אלו הוא בעבודה עם מידע גיאומטרי.
- (מכני, חשמלי, אינסטלציה) – מערכות הנדסיות הכוללות רכיבים מכניים, חשמליים ואינסטלציה MEP ומפרטות את "מערכת הדם" הפנימית של הפרויקט.
- (מערכות מידע גיאוגרפיות) – משמשות לניתוח ותכנון שטחים, כולל מיפוי וניתוח מרחבי GIS.
- (תוכנה לניהול איכות הבנייה) – מבטיחה שהליכי הבנייה עומדים בסטנדרטים ובתקנות שנקבעו CQMS מסייעת בהסרת פגמים.
- (ניהול פרויקטי בנייה) – כולל תכנון, תיאום ובקרת תהליכי הבנייה CPM.
- (ניהול מתקנים בעזרת מחשב) – מערכות לניהול ותחזוקת מבנים CAFM.
- (ניהול שרשרת אספקה) – נדרשת לאופטימיזציה של זרימת החומרים והמידע בין הספקים לאתר הבנייה SCM.
- (ניהול ביצועים ארגוניים) – מכוונת לשיפור תהליכים עסקיים ופרודוקטיביות EPM.
- (תוכנת ניהול נכסים) – משמשת לאופטימיזציה של השימוש, הניהול והתחזוקה של ציוד ותשתיות AMS במהלך כל מחזור חיי הנכסים.
- (ניהול נכסים ריאליים) – כולל משימות ותהליכים הקשורים לניהול ותפעול מבנים וקרקעות, כמו גם RPM משאבים ונכסים קשורים.

- (הנדסה בעזרת מחשב –) הנדסה ממוחשבת, כוללת מערכות חישוביות וסימולציות, כגון ניתוח בשיטת CAE (CFD) ודינמיקה של נוזלים חישובית (FEA) אלמנטים סופיים.
- (CAE) דינמיקה של נוזלים חישובית – מודל של זרימת נוזלים וגזים. תת-קטגוריה של CFD.
- (תכנון תהליכים בעזרת מחשב –) תכנון תהליכים טכנולוגיים בעזרת מחשב. משמש ליצירת מפות CAPP מסלול ומפות טכנולוגיות.
- (ייצור בעזרת מחשב –) ייצור אוטומטי, יצירת תוכניות בקרה למכונות עם בקרת מספרית CAM.
- (ניהול נתוני מוצר –) ניהול נתונים על מוצר, מערכת לאחסון וניהול תיעוד טכני PDM.
- (מערכת ניהול ביצועי ייצור –) מערכת לניהול תהליכי ייצור בזמן אמת MES.
- (ומערכות PDM, CAPP, CAM) ניהול מחזור חיי מוצר – ניהול מחזור חיי רכיב בפרויקט, מאחד את PLM נוספות לשליטה מלאה על המוצר מהפיתוח ועד ההשגדה.

מערכות אלו ורבות אחרות, הכוללות פתרונות תוכנה מגוונים, הפכו לחלק בלתי נפרד מהתעשייה הבנייתית המודרנית. מהותית, מערכות אלו מהוות בסיס נתונים מיוחדים עם ממשקים אינטואיטיביים, המאפשרים הזנה, עיבוד וניתוח מידע בצורה יעילה בכל שלב של תכנון ובנייה. אינטגרציה של כלים דיגיטליים זה עם זה לא רק מסייעת באופטימיזציה של תהליכי עבודה, אלא גם מעלה באופן משמעותי את דיוק ההחלטות המתקבלות, מה שמשפיע לטובה על זמני הביצוע ואיכות הפרויקטים.

אך האינטגרציה אינה קיימת בחצי מהמקרים. לפי סטטיסטיקה, רק כל אפליקציה או מערכת שנייה משולבת עם פתרונות אחרים. זה מצביע על המשכיות הפיצול בסביבה הדיגיטלית ומדגיש את הצורך בפיתוח תקנים פתוחים וממשקים מאוחדים כדי להבטיח חילופי מידע רציפים במסגרת פרויקט הבנייה.

אחד האתגרים המרכזיים באינטגרציה עבור חברות מודרניות הוא המורכבות הגבוהה של מערכות דיגיטליות והדרישות מהמשתמשים, הנדרשות לחיפוש ופרשנות מידע בצורה יעילה. לצורך תמיכה בכל מערכת המוטמעת בעסק, מוקמת צוות מומחים בראשות מנהל מפתח.

מנהל המערכת המרכזי משחק תפקיד מכריע בכיוונון זרימת הנתונים ואחראי על איכות המידע הסופי בדומה למנהלים הראשונים לפני אלפי שנים שאחראים על המספרים שנרשמו על פפירוס או לוחות חימר.

כדי להפוך זרמי מידע מבודדים לכלי ניהול, נדרשת יכולת לאינטגרציה מערכתית ולניהול נתונים. בארכיטקטורה זו המנהלים צריכים לפעול כאלמנטים ברשת אחת – בדומה למיצילים, המחבר בין חלקים שונים של החברה לאורגניזם חי שלם, המסוגל להסתגל ולהתפתח.

מיציל של תאגידים: כיצד נתונים מתחברים לתהליכי עסקים

תהליך האינטגרציה של נתונים באפליקציות ובסיסי נתונים מתבסס על אגירה של מידע המתקבל ממקורות מגוונים, כולל מחלקות שונות ומומחים. המומחים מחפשים את הנתונים הנדרשים, מעבדים אותם ומעבירים למערכותיהם ואפליקציותיהם לשימוש נוסף.

כל מערכת בחברה, המורכבת מקבוצת כלים, טכנולוגיות ובסיסי נתונים – היא עץ ידע, שורשיה באדמת נתונים היסטוריים וצומחת כדי להביא פירות חדשים בדמות פתרונות מוכנים: מסמכים, חישובים, טבלאות, גרפים ולוחות

מחוננים. מערכות בחברה, בדומה לעצים באותו אזור יער, מתקשרות ומתקשרות זו עם זו, מהוות מערכת מורכבת ומסודרת היטב, הנתמכת ומנוהלת על ידי מנהלי מומחים.

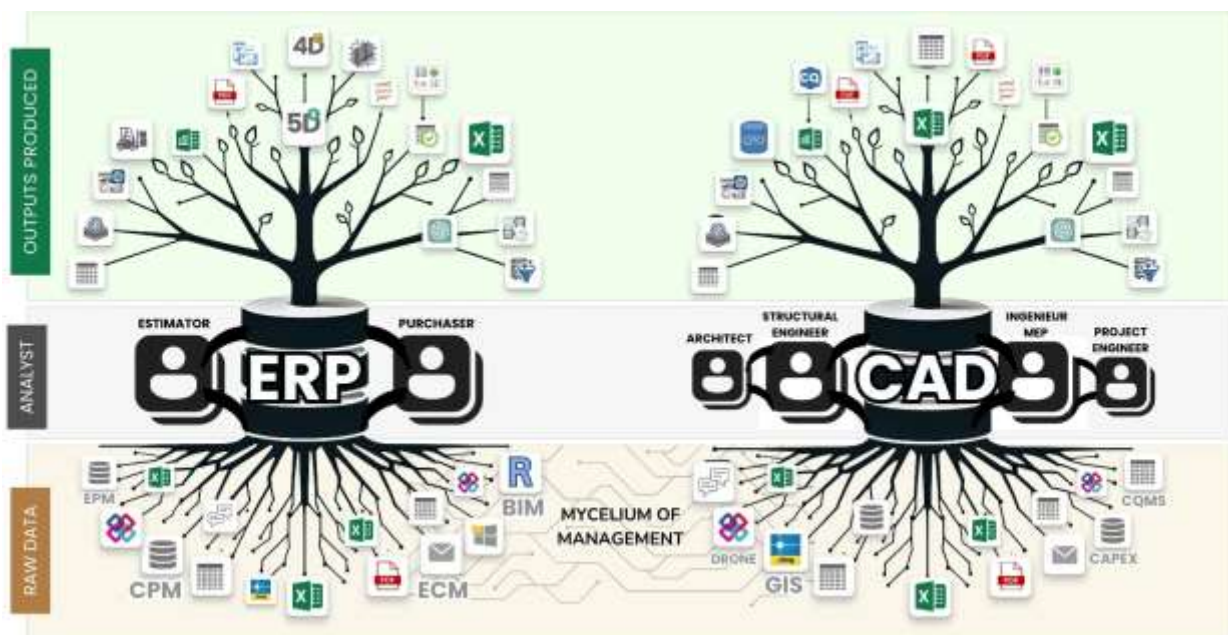
מערכת החיפוש והעברת המידע בחברה פועלת כרשת יער מורכבת, המורכבת מעצים (מערכות) ומיצלויים של פטריות (מנהלים), (המהווים מתווכים ומעבדים, ומספקים את העברת המידע והגעתו למערכות הנדרשות. זה מסייע לשמור על זרימה בריאה ויעילה של נתונים והפצתם בתוך החברה.

מומחים, בדומה לשורשים, סופגים נתונים גולמיים בשלביו הראשונים של הפרויקט, והופכים אותם לחומרי הזנה עבור פועלות כצירי מידע רבי (ועוד ERP, CPM, BIM - 1.24 איור) האקוסיסטמה הארגונית. מערכות ניהול נתונים ותוכן עוצמה, דרכם הידע הזה זורם בכל רמות החברה.

כמו בטבע, שבו כל אלמנט באקוסיסטמה ממלא תפקידו, בנוף העסקי של החברה כל משתתף בתהליך - מהמהנדס ועד אינם רק מנגנונים (1.25 איור) המערכתיים "trees" האנליסט - תורם לצמיחה ולפוריות של הסביבה המידע". עצ -לאיסוף מידע, אלא גם יתרון תחרותי, המבטיח פיתוח בר קיימא של החברה.

אקוסיסטמות יער משקפות במדויק את העקרונות של ארגון מבנים דיגיטליים ארגוניים. בדומה למבנה הרב-שכבתי של היער - מהשיחים ועד לקצוות העצים - הניהול הארגוני מפיץ משימות לפי רמות אחריות ומחלקות פונקציונליות.

שורשים עמוקים ומסועפים של העצים מספקים יציבות וגישה לחומרי הזנה. באופן דומה, מבנה ארגוני חזק ותהליכים יציבים בעבודה עם נתונים איכותיים תומכים בכל האקוסיסטמה המידע של החברה, ומסייעים לצמיחה ולפיתוח בר קיימא גם בתקופות (רוחות חזקות) של חוסר יציבות בשוק ומשברים.



אינטגרציית נתונים דרך מערכות שונות דומה למיצלויים, המחבר בין מנהלים ומומחים לרשת מידע אחת 1.25 איור.

"ההבנה המודרנית של היקף בעסקים התפתחה. כיום, ערך החברה נקבע לא רק על ידי החלק הנראה שלה - "הקצוות בצורת מסמכים ודוחות סופיים - אלא גם על ידי עומק "מערכת השורשים" של נתונים שנאספו באיכות ועובדו בצורה מערכתית. ככל שיותר מידע ייאסף ויעובד, כך ערך העסק יגדל. חברות, המאגדות באופן שיטתי "קומפוסט" של נתונים מעובדים ומסוגלות להפיק מהם תובנות מועילות, מקבלות יתרון אסטרטגי

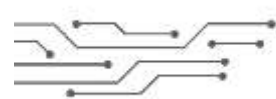
מידע היסטורי הופך לסוג חדש של הון, המבטיח צמיחה, אופטימיזציה של תהליכים ויתרון תחרותי בעולם הממוקד בנתונים, המנצחים אינם אלו שיש להם יותר, אלא אלו שיודעים יותר.

עבור תעשיית הבנייה, זה מציין מעבר לניהול פרויקטים בזמן אמת, שבו כל התהליכים - מעיצוב ורכישות ועד תיאום CAD, מודלים, ERP קבלנים - יתבססו על נתונים מעודכנים מדי יום. אינטגרציה של מידע ממקורות שונים (מערכות תאפשר לבנות תחזיות מדויקות יותר, להגיב במהירות לשינויים ולהימנע מעיכובים (RFID, באתר הבנייה IoT חיישני הנגרמים מחוסר נתונים מעודכנים

בהתאם למחקר "העסק של שנת 2025, המנוהל על ידי נתונים" (מקינזי & חברה, ©2022 [18]), [[חברות מצליחות בעתיד יתבססו על נתונים בכל ההיבטים המרכזיים של פעילותן - מהחלטות אסטרטגיות ועד אינטראקציות תפעוליות

נתונים יפסיקו להיות פשוט כלי אנליטי ויהפכו לחלק בלתי נפרד מכל תהליכי העסק, תוך כדי הבטחת שקיפות, שליטה ואוטומציה בניהול. גישה מונחית נתונים תאפשר לארגונים למזער את השפעת הגורם האנושי, להפחית סיכונים תפעוליים ולהגביר את שקיפות ויעילות קבלת ההחלטות

המאה ה-21 הופכת את הפרדיגמה הכלכלית: אם בעבר נפט כונה "זהב שחור" בזכות יכולתו להניע מכונות ותחבורה, כיום, נתונים היסטוריים, שנדחסים תחת לחץ הזמן, הופכים למשאב אסטרטגי חדש, המניע לא מכונות אלא אלגוריתמים. לקבלת החלטות, שיגרמו לעסק להתקדם



פרק 1.3.

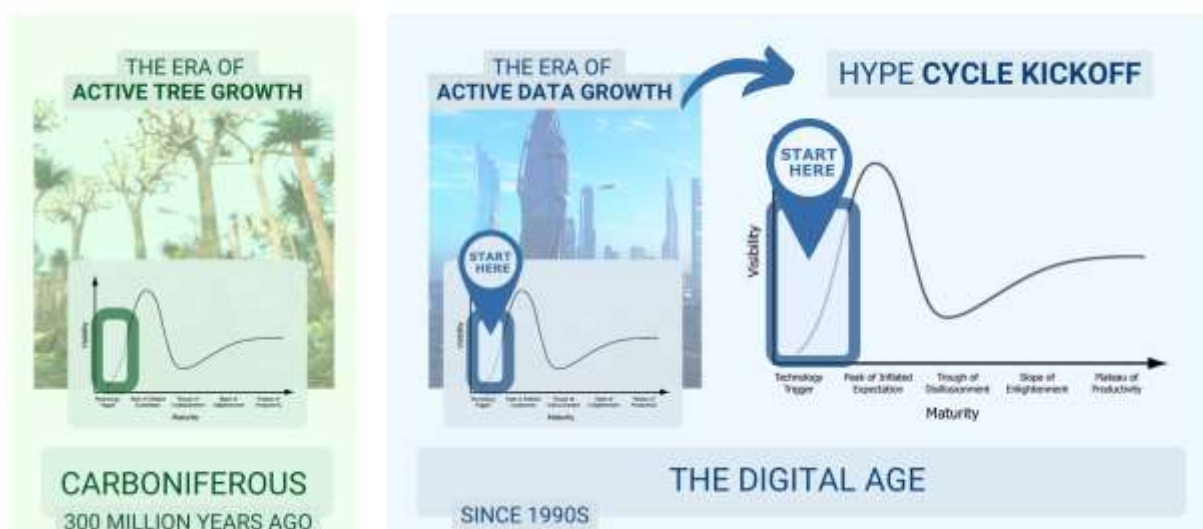
מהפכה דיגיטלית והתפוצצות הנתונים

תחילת עליית נפחי הנתונים כגל אבולוציוני

תעשיית הבנייה חווה התפוצצות מידע חסרת תקדים. אם נציג את העסק כעץ ידע (איור 1.25), (המוזן מנתונים, אז השלב הנוכחי של דיגיטציה ניתן להשוות לצמיחה המהירה של צמחייה בעידן הפחם - תקופה שבה הביוספירה של כדור הארץ - השתנתה בזכות הצטברות מהירה של ביומסה) (איור 1.31)

בתנאים של התפתחות דיגיטלית גלובלית, כמות המידע בתחום הבנייה מוכפלת מדי שנה. טכנולוגיות מודרניות מאפשרות לאסוף נתונים ברקע, לנתח אותם בזמן אמת ולהשתמש בהם בהיקפים שנראו עד לא מזמן כבלתי אפשריים

על ידי גורדון מור (מייסד שותף של אינטל, ©הצפיפות והמורכבות של formul בהתאם לחוק מור, שנ [19] מעגלים משולבים, כמו גם הכמויות של נתונים מעובדים ומאוחסנים, מוכפלות בערך כל שנתיים



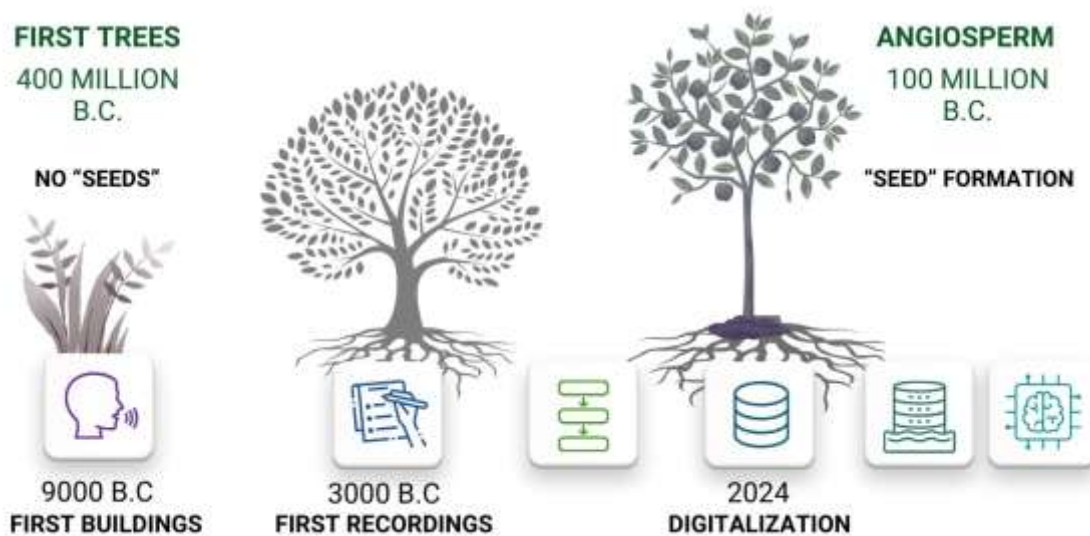
איור 1.31 תחילת הדיגיטציה הביאה לצמיחה אקספוננציאלית של נתונים, בדומה להתפרצות הצמחייה בעידן הפחם.

אם מבנים מגליתיים עתיקים, כמו גובקלי טפה (טורקיה, (לא השאירו אחריהם ידע מתועד שניתן לשימוש חוזר, כיום טכנולוגיות דיגיטליות מאפשרות הצטברות ושימוש חוזר במידע. ניתן להשוות זאת למעבר האבולוציוני מצמחים ספורים לצמחים זרעיים (הופעת הזרע נתנה דחיפה להפצה רחבה של חיים על פני כדור הארץ). איור 1.32).

שניתן להרחיב, DNA באופן דומה, נתונים מפרויקטים קודמים הופכים לסוג של "זרעים דיגיטליים" - נושאי ידע ולהשתמש בהם בפרויקטים ומוצרים חדשים. הופעת כלים מודרניים של אינטליגנציה מלאכותית - למידת מכונה - ChatGPT, LLaMa, Mistral, Claude, DeepSeek, QWEN, Grok - (LLM) ודגמים לשוניים גדולים

מאפשרת לחלץ, לפרש וליישם נתונים בהקשרים חדשים באופן אוטומטי

בדומה לכך, כיצד הזרעים חוללו מהפכה בהפצת חיים על פני כדור הארץ שהיה בתחילה חסר חיים, זרעי הנתונים "הופכים" לבסיס להופעה אוטומטית של מבנים ומידע חדשים, ומאפשרים למערכות אקולוגיות דיגיטליות להתפתח באופן עצמאי ולהתאים לדרישות המשתנות של המשתמשים.



זרעי נתונים "דיגיטליים" משחקים את אותה תפקיד אבולוציוני כמו אנגיוספרמות - צמחים פורחים ששינו את^{1.32} ריס המערכת האקולוגית של כדור הארץ.

אנו עומדים על סף עידן חדש בבנייה, שבו התפוצצות הנתונים והפצת "זרעי הנתונים" - מידע מובנה מפרויקטים קודמים (LLM) ונוכחיים - מעצבים את היסוד של העתיד הדיגיטלי של הענף. האבקה שלהם בעזרת מודלים לשוניים גדולים מאפשרת לא רק לצפות בשינויים דיגיטליים, אלא גם להשתתף באופן פעיל ביצירת מערכות אקולוגיות מתאימות ולומדות. זו לא אבולוציה - זו מהפכה דיגיטלית, שבה הנתונים הופכים לחומר הבניין המרכזי של מציאות חדשה.

נפח הנתונים בענף הבנייה גדל באופן דרמטי בזכות המידע המגיע מתחומים שונים במהלך כל מחזור חיי הפרויקטים הבנייתיים. הצטברות הנתונים העצומה הזו דחפה את תעשיית הבנייה לעידן הנתונים הגדולים^[20]. פרופסור האנג יאן, הפקולטה להנדסה אזרחית ואדריכלות, אוניברסיטת ווהאן הטכנולוגית, ווהאן, סין

עליית נפח הנתונים בעידן המידע מזכירה תהליכים אבולוציוניים בטבע: כפי שהתפתחות היערות שינתה את הנוף הקדום של כדור הארץ, כך התפוצצות המידע הנוכחית משנה את הנוף של כל תעשיית הבנייה.

נפח הנתונים המיוצרים בחברה המודרנית

במהלך השנתיים האחרונות נוצרו 90% מכלל הנתונים הקיימים בעולם^[21]. [נכון לשנת 2023, כל אדם, כולל אנשי מקצוע בענף הבנייה, מייצר כ-1.7 מגה-בייט של נתונים בשנתיים^[22], והנפח הכולל של הנתונים בעולם יגיע ל-64 זטה-

[23] בייט בשנת 2023, וצפוי לעלות על 180 זטה-בייט, או 15×10^{18} מגה-בייט, עד 2025.

התפוצצות המידע הזו יש לה תקדים היסטורי - המצאת הדפוס על ידי יוהאן גוטנברג במאה ה-15. רק חמישים שנה לאחר הופעתו, מספר הספרים באירופה הוכפל: תוך כמה עשורים הודפסו מספר ספרים שווה למספר הספרים שנוצרו בידיים במהלך 1200 השנים הקודמות [24]. [כיום אנו עדים לצמיחה מהירה עוד יותר: נפח הנתונים בעולם מכפיל את עצמו כל שלוש שנים

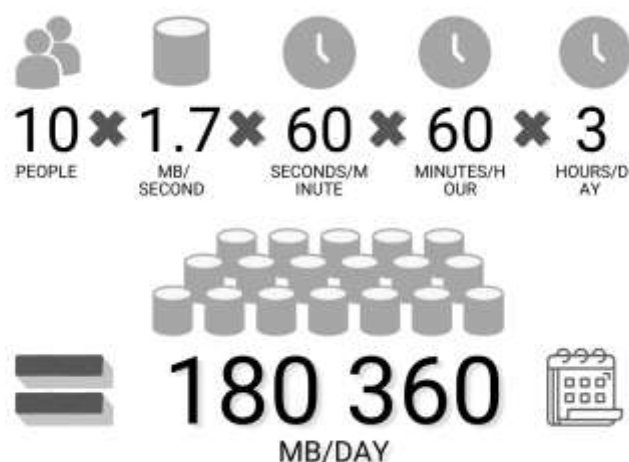
בהתחשב בקצב הנוכחי של צמיחת הנתונים, תעשיית הבנייה עשויה לייצר בעשורים הקרובים נפח מידע שווה לזה שנצבר במהלך כל ההיסטוריה שלה.



אחסון הנתונים היומיומי על ידי כל עובד בשרתים של החברה תורם לצמיחה מתמדת של נפח הנתונים 1.33. ריס.

בעולם המודרני של עסקי הבנייה, אפילו חברות קטנות מייצרות מדי יום כמות עצומה של מידע בפורמטים שונים, והעקבות הדיגיטליים של חברה בנייה קטנה יכולים להגיע לעשרות גיגה-בייט ביום - ממודלים ושרטוטים ועד תיעוד תמונות וחיישנים באתר. אם נניח שכל איש מקצוע מייצר בממוצע כ-1.7 מגה-בייט של נתונים בשנייה, זה שווה בערך (ריס. 1.33) ל-146 גיגה-בייט ביום, או 53 טרה-בייט בשנה.

כאשר צוות פעיל של 10 אנשים עובד במשך 3 שעות ביום, הנפח הכולל של המידע המיוצר ביום מגיע ל-180 גיגה- (ריס. 1.34) בייט.



חברה של 10 אנשים מייצרת כ-50–200 גיגה-בייט של נתונים ביום 1.34. ריס.

אם נניח ש-30% מהנתונים הפעילים הם חדשים (שאר הנתונים נכתבים מחדש או נמחקים), (אז חברה עם 10 עובדים יכולה ליצור כמות של כמה מאות גיגה-בייט של נתונים חדשים בחודש) הנתונים האמיתיים תלויים בסוג הפעילות של החברה).

כך, מתברר בבירור: אנחנו לא רק מייצרים יותר ויותר נתונים - אנחנו מתמודדים עם צורך גובר בניהול, אחסון וזמינות ארוכת טווח של נתונים אלו. ואם בעבר הנתונים יכלו "לשכב" על שרתים מקומיים ללא עלויות מיוחדות, אז בעידן של טרנספורמציה דיגיטלית יותר ויותר חברות מתחילות להשתמש בפתרונות ענן כבסיס לתשתית המידע שלהן.

עלות אחסון הנתונים: היבט כלכלי

בשנים האחרונות יותר ויותר חברות מעבירות את אחסון הנתונים לשירותי ענן. לדוגמה, אם חברה מאחסנת חצי מהנתונים שלה בענן, במחיר ממוצע של 0.015 דולר לגיגה-בייט בחודש, ההוצאות שלה על אחסון עשויות לעלות כ-10–25 [דולר 25] בחודש 50–10.

עבור חברה קטנה עם מודלים טיפוסיים של יצירת נתונים, ההוצאות על אחסון בענן עשויות לנוע ממאות עד פוטנציאלית לאחר מספר שנים, מה שיוצר נטל כלכלי משמעותי (1.35. ריש) יותר מאלף דולר בחודש.

"ארגונים מעבירים את אחסון הנתונים למיקור חוץ, מכיוון שהמורכבות גוברת" Forrester לפי מחקר של שבו השתתפו 214 מנהיגים מקבלי החלטות בתחום התשתית הטכנולוגית, יותר משליש מהארגונים, [26] מעבירים את מערכות אחסון הנתונים למיקור חוץ כדי להתמודד עם הגידול בנפח ובמורכבות של פעולות עם נתונים, כאשר כמעט שני שלישים מהחברות מעדיפות מודל מבוסס מנוי.



העברת נתונים לענן עשויה להגדיל את ההוצאות החודשיות על אחסון עד 2,000 דולר גם עבור חברה עם 10.35 מיליארד ריש עובדים בלבד.

אשר מוסיפות, ERP ומערכות CAD (BIM), CAFM, PMIS המצב מסתבך עם המעבר המהיר לטכנולוגיות ענן, כגון לעלויות האחסון והעיבוד של נתונים. בסופו של דבר, חברות נאלצות לחפש דרכים לאופטימיזציה של ההוצאות ולהפחית את התלות בספקי הענן.

הגישות לאחסון נתונים החלו, (LLM) החל משנת 2023, עם ההתפתחות הפעילה של מודלים לשוניים גדולים להשתנות. יותר ויותר חברות מתחילות לחשוב על החזרת השליטה על הנתונים, מכיוון שעליהן המידע על שרתים פנימיים הופך לבטוח יותר ולכדאי יותר.

בהקשר זה, עולה המגמה להפסיק להשתמש במערכות אחסון ועיבוד נתונים בענן רק עבור נתונים נחוצים, לטובת פריסת ארגוניים מקומיים. כפי שהדגיש מנכ"ל מיקרוסופט באחד הראיונות שלו [27], [במקום להסתמך על AI ופתרונות LLM ינהלו את התהליכים בבסיסי AI בענן לביצוע משימות שונות, סוכני SaaS מספר אפליקציות נפרדות או פתרונות הנתונים, תוך אוטומציה של פונקציות ממערכות שונות.

הגישה הישנה לשאלה זו [עיבוד נתונים] הייתה כדלקמן: אם נזכור כיצד יישומים עסקיים שונים התמודדו עם SAP. אינטגרציה, הם השתמשו בקונקטורים. חברות מכרו רישיונות לקונקטורים הללו, וסביב זה נוצר מודל עסקי, הייתה אפשרות רק בהנחה שיש קונקטור מתאים. לכן SAP הוא אחד הדוגמאות הקלאסיות: גישה לנתוני [ERP] הגישה, לפחות, שאנו מאמצים היא [AI סוכני] נראה לי שמשדה דומה יקרה גם במקרה של אינטראקציה עם כי אם נחשוב על [AI סוכני] כדלקמן: אני חושב שהקונספט של קיום יישומים עסקיים, ככל הנראה, יתמוטט בעידן BG2, זה, הם בעצם מהווים בסיסי נתונים עם הרבה לוגיקה עסקית. — סאטיה נדלה, מנכ"ל מיקרוסופט, ריאיון לערוץ 2024

חורגת ממערכות קלאסיות. אינטליגנציה מלאכותית הופכת LLM-בפרדיגמה זו, הגישה המונעת נתונים עם שימוש ב למתווך בין המשתמש לנתונים (איור 2.23, איור 2.24), (מסירה את הצורך בממשקי תיווך רבים ומעלה את היעילות -"של תהליכי העסקים. נרחיב על גישה זו לעבודה עם נתונים בפרק "הפיכת כאוס לסדר והפחתת מורכבות

בעוד הארכיטקטורה של העתיד רק מתהווה, חברות כבר מתמודדות עם השלכות של החלטות העבר. הדיגיטציה ההמונית בעשורים האחרונים, שהייתה מלווה בהטמעה של מערכות מבודדות ואיסוף נתונים בלתי מבוקר, הביאה לבעיה חדשה — עומס מידע

גבולות צבירת הנתונים: מהמון למשמעות

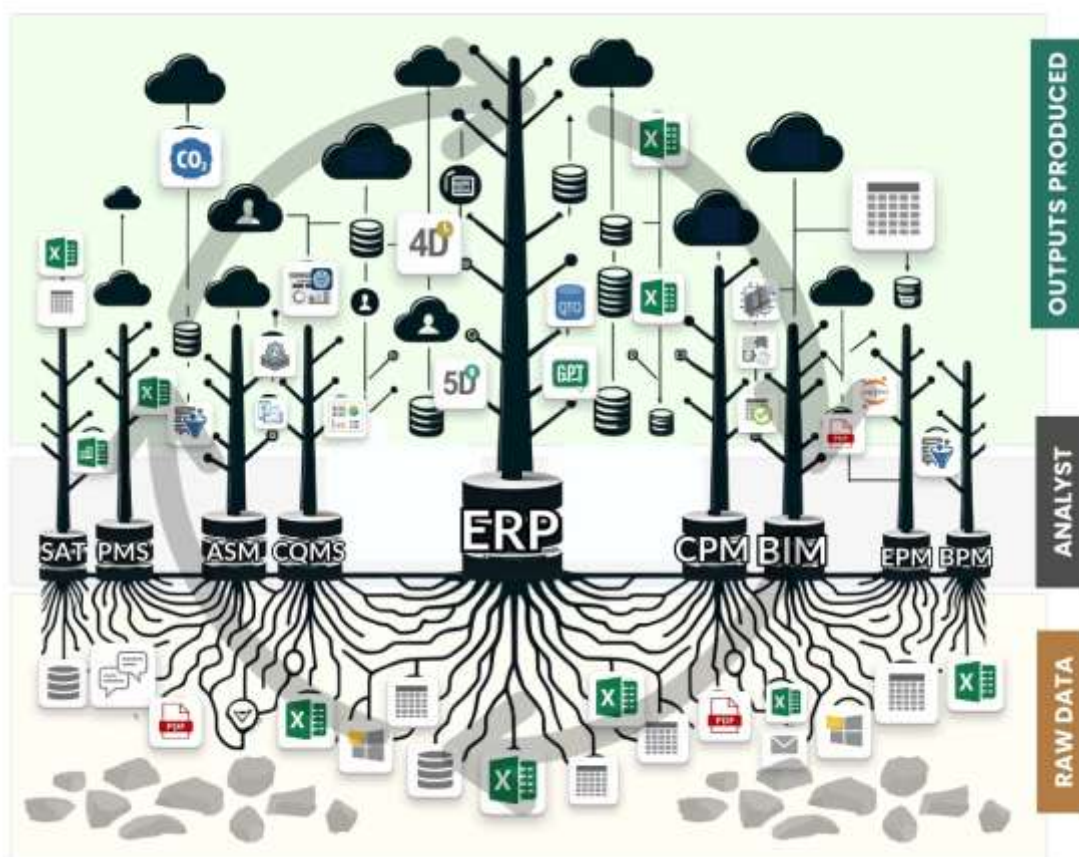
מערכות מודרניות של חברות מתפתחות ופועלות בהצלחה כאשר הצמיחה מנוהלת, כאשר כמות הנתונים ומספר ומנהלים. עם זאת, בעשורים האחרונים, הדיגיטציה הביאה לעלייה IT-היישומים נמצאים באיזון עם יכולות מחלקות הבלתי נשלטת בכמות ובמורכבות הנתונים, מה שגרם לאפקט של רוויה במערכת המידע של החברות.

כיום, שרתים ואחסונים נתונים חשופים לזרם חסר תקדים של מידע גולמי ומגוון, שלא מצליח להפוך לקומפוסט ומאבד במהירות את הרלוונטיות שלו. המשאבים המוגבלים של החברה לא מצליחים להתמודד עם הזרם הזה, והנתונים מצטברים באחסונים מבודדים) מה שנקרא "סילונים", ("שדורשים עיבוד ידני כדי לחלץ מידע שימושי

כתוצאה מכך, בדומה ליער שגדל בו פטל ומכוסה בעובש, מערכות ניהול מודרניות של חברות סובלות לעיתים קרובות מעומס מידע. במקום שהמערכת האקולוגית הארגונית תתבסס על חומרי מידע מזינים נוצרים אזורים מבודדים של נתונים מגוונים, מה שמוביל בהכרח לירידה ביעילות הכוללת של תהליכי העסקים.

תקופה ארוכה של צמיחה אקספוננציאלית בכמויות הנתונים, שנצפתה בעשורים האחרונים, תוחלף בהכרח בשלב של רוויה ולאחר מכן קירור. כאשר האחסונים יגיעו לגבול, יתרחש שינוי איכותי: הנתונים יפסיקו להיות פשוט אובייקט אחסון, ויהפכו למשאב אסטרטגי.

עם התפתחות הבינה המלאכותית ולמידת המכונה, חברות מקבלות אפשרות להפחית את העלויות של עיבוד מידע ולעבור מצמיחה כמותית לשימוש איכותי בנתונים. בעשור הקרוב, תעשיית הבנייה תצטרך לשנות את המוקד - מיצירת כמויות חדשות של נתונים להבטחת המבנה, שלמות והערך האנליטי שלהם.



מקורות נתונים מבודדים מפריעים להחלפת מידע בין מערכות נתונים 1.36 איור.

הערך המרכזי כבר לא נמצא בכמות המידע, אלא ביכולת לפרש אותו אוטומטית ולהפוך אותו לידע יישומי, מועיל לקבלת החלטות ניהוליות. כדי שהנתונים יהיו באמת מועילים, יש לנהל אותם בצורה נכונה. לאסוף, לבדוק, למבנה, לאחסן ולנתח בהקשר של משימות עסקיות ספציפיות.

תהליך ניתוח הנתונים בחברה דומה למחזור חיי והרס העצים ביער והופעת עצים צעירים וחזקים חדשים: עצים בוגרים מתפרקים והופכים לאדמה פורייה עבור נבטים חדשים. תהליכים מוכנים ומושלמים בסיום השימוש שלהם נכנסים לאקוסיסטמה המידע של החברה, והופכים בסופו של דבר לחומס מידע, המזין את הצמיחה העתידית של מערכות ונתונים חדשים.

עם זאת, בפועל, מחזור זה לעיתים קרובות מופר. במקום חידוש אורגני, נוצר כאוס שכבותי - דמוי שכבות גיאולוגיות שבהן מערכות חדשות נערמות על גבי ישנות, ללא אינטגרציה מעמיקה ומבנה. כתוצאה מכך, נוצרים "סילונים" מידע מבודדים, המפריעים למחזור הידע ומקשים על ניהול הנתונים.

צעדים נוספים: מתיאוריה של נתונים לשינויים מעשיים

האבולוציה של הנתונים בבנייה היא הדרך מלווחות חימר לפלטפורמות מודולריות מודרניות. היום, האתגר לא נמצא

באיסוף מידע, אלא ביצירת מבנה שהופך נתונים מבודדים ומגוונים למשאב אסטרטגי. לא משנה מה תפקידך - מנהל חברה או מהנדס רגיל - הבנת ערך הנתונים ויכולת העבודה איתם בעתיד תהפוך לכישור מקצועי מרכזי

לסיכום חלק זה, כדאי להדגיש את הצעדים המעשיים העיקריים שיכולים לעזור ליישם את הגישות שנדונו במשימות היומיומיות שלך

■ ערוך ביקורת אישית על זרמי המידע

- ☐ ערוך רשימה של כל המערכות והיישומים שבהם אתה עובד מדי יום
- ☐ ציין היכן אתה מבזבז את רוב הזמן שלך בחיפוש או בדיקה מחדש של נתונים
- ☐ הגדר את מקורות המידע המרכזיים שלך
- ☐ נתח את הנוף הנוכחי של היישומים שלך מבחינת עודפות וחפיפות פונקציות

■ שאף להתקדם בתהליכים לפי רמות בשלות אנליטית

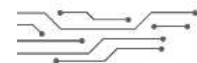
- ☐ התחל את העבודה על המשימות שלך עם אנליטיקה תיאורית (מה קרה?)
- ☐ בהדרגה, הטמע אנליטיקה דיאגנוסטית (למה זה קרה?)
- ☐ חשוב כיצד בתהליכים תוכל לעבור לאנליטיקה חיזוי (מה יקרה?) ואנליטיקה פרסקריפטיבית (מה לעשות?)

■ התחל למבנה את הנתונים העבודה שלך

- ☐ הטמע מערכת אחידה לשמות קבצים ותיקיות שאתה משתמש בהם לעיתים קרובות בעבודתך
- ☐ צור תבניות למסמכים ודוחות בשימוש תדיר
- ☐ ארכבו באופן קבוע פרויקטים שהושלמו עם מבנה ברור

אפילו אם אינכם יכולים לשנות את כל התשתית המידעית בצוות או בחברה שלכם, התחילו עם התהליכים האישיים שלכם ושיפורים קטנים בעבודתכם היומיומית. זכרו, שהערך האמיתי של הנתונים מתגלה לא בנפחם, אלא ביכולת להפיק מהם תועלת מעשית. אפילו מערכות מידע קטנות, אך מאורגנות ומנותחות כראוי, יכולות להניב השפעה משמעותית, אם הן משולבות בתהליכי קבלת ההחלטות

בפרקים הבאים של הספר נעבור לשיטות וכלים ספציפיים לעבודה עם נתונים, נבחן דרכים להמיר מידע לא מובנה לסטים מובנים, נלמד טכנולוגיות אוטומציה לניתוח ונפרט כיצד לבנות מערכת אנליטית יעילה בחברת בנייה





II חלק כיצד עסקי הבנייה שוקעים בכאוס של נתונים

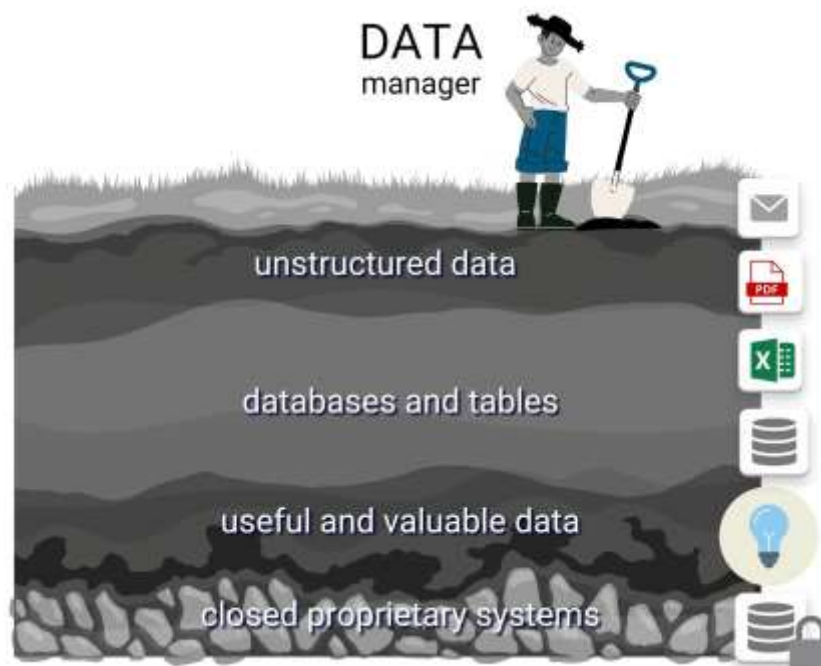
החלק השני מוקדש לניתוח ביקורתי של הבעיות שבהן נתקלות חברות הבנייה בעבודתן עם כמויות הולכות וגדלות של נתונים. נבחנות בהרחבה ההשלכות של פיצול המידע ותופעת הדעה של HiPPO נתונים בסילו, "המונעת קבלת החלטות יעילה. נבדקת בעיית הגישה של" האדם המרוויח ביותר (והשפעתה על איכות ההחלטות הניהוליות בפרויקטים של בנייה מוערכת השפעת הדינמיות של תהליכי העסקים והמורכבות ההולכת וגדלה שלהם על זרמי המידע והיעילות התפעולית. מובאים דוגמאות קונקרטיות כיצד המורכבות המוגזמת של המערכות מגדילה את העלויות ומפחיתה את הגמישות של הארגונים. תשומת לב מיוחדת מוקדשת להגבלות הנגרמות על ידי פורמטים קנייניים ולפרספקטיבות של שימוש בסטנדרטים פתוחים בתעשיית הבנייה. מוצגת הקונספט של המעבר לאקוסיסטמות תוכנה מבוססות בינה הממזערות את המורכבות המוגזמת ואת המחסומים הטכניים, LLM-מלאכותית ו

פרק 2.1. פיצול וסילו של נתונים

ככל שיש יותר כלים, כך העסק יעיל יותר

מבט ראשון עשוי להראות כי הגברת מספר הכלים הדיגיטליים מובילה לעלייה ביעילות. עם זאת, במציאות הדברים שונים. עם כל פתרון חדש, בין אם מדובר בשירות ענן, מערכת מיושנת או דוח אקסל נוסף, החברה מוסיפה שכבה נוספת לנוף הדיגיטלי שלה - שכבה שלעיתים קרובות אינה משתלבת עם השאר.

הנתונים ניתן להשוות לפחם או לנפט: הם מתהווים במשך שנים, נלחצים "תחת שכבות של כאוס, טעויות, תהליכים לא מובנים ופורמטים נשכחים. כדי להפיק מהם מידע שימושי באמת, חברות נדרשות למעשה לחדור דרך שכבות של פתרונות מיושנים ורעש דיגיטלי



איור 2.11 נתונים בגדלים שונים יוצרים שכבות מבודדות - אפילו תובנות "זהב" הולכות לאיבוד בסלעי הגיאולוגיה של מורכבות מערכתית.

כל אפליקציה חדשה משאירה אחריה עקבות: קובץ, טבלה או "סילו" מבודד שלם על השרת. שכבת חימר היא נתונים מיושנים ושכוחים, שכבת חול היא טבלאות ודוחות מפוזרים, ושכבת גרניט היא פורמטים סגורים קנייניים שאינם ניתנים לאינטגרציה. עם הזמן, הסביבה הדיגיטלית של החברה הולכת ומזכירה יותר ויותר מחסן שכבות עם צבירה בלתי מבוקרת של מידע, שבו הערך הולך ואובד בעמקי השרתים של החברה.

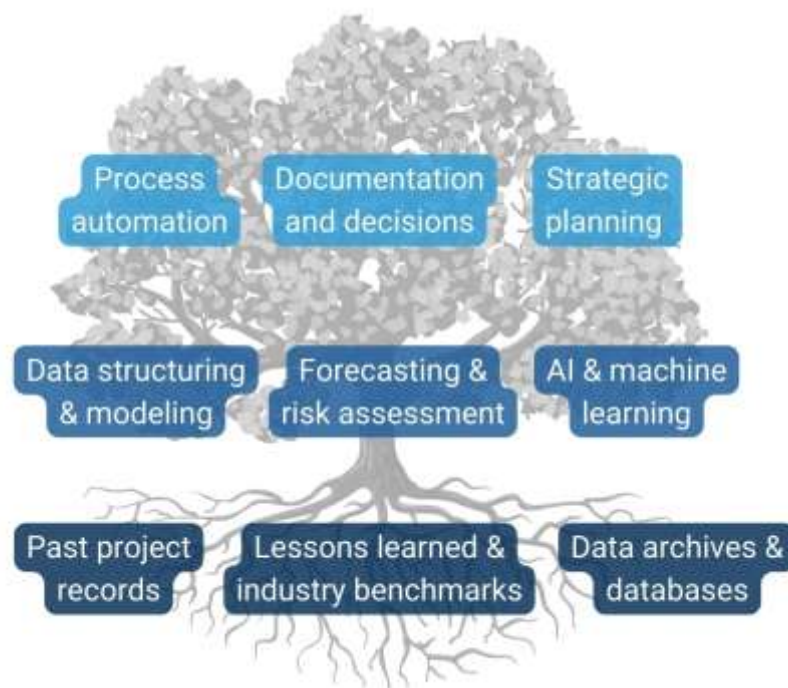
אלא גם הדרך לנתונים איכותיים, complicates עם כל פרויקט חדש וכל מערכת חדשה, לא רק התשתית מת ושימושיים. כדי להגיע ל"מחצבים" יקרי ערך, יש לבצע ניקוי מעמיק, לארגן את המידע, לשבור, "לקבץ אותו לחלקים משמעותיים ולהפיק מידע אסטרטגי חשוב דרך ניתוח ומידול נתונים

נתונים הם דבר יקר ערך, והם ישרדו יותר מהמערכות עצמן.

טים ברנרס-לי, אב האינטרנט וממציא האתר הראשון —

לפני שהנתונים יכולים להפוך ל"דבר יקר ערך" ולבסיס אמין לקבלת החלטות, הם חייבים לעבור הכנה קפדנית. הכנה נכונה הופכת מידע מפוזר לניסיון מאורגן, קומפוסט "מידע שימושי, אשר לאחר מכן הופך לכלי לחיזוי ואופטימיזציה

ישנה טעות נפוצה לחשוב כי לניתוח נדרשים נתונים נקיים לחלוטין, אך במציאות, היכולת לעבוד עם נתונים "מלוכלכים" היא חלק בלתי נפרד מהתהליך.



הנתונים הם מערכת השורש והבסיס לעסק, אשר בתורו מבוסס על תהליכי קבלת החלטות 2.12 איור.

כל עוד הטכנולוגיות לא עומדות במקום, גם העסק שלך צריך להתקדם וללמוד ליצור ערך מנתונים. כפי שחברות נפט ופחם יוצרות תשתית לחציבת מינרלים, כך גם העסק צריך ללמוד לנהל באופן נכון את זרמי המידע החדשים בשרתים שלו ולהפיק מידע יקר ערך מנתונים לא מנוצלים, בפורמטים שונים ובתאריכים ישנים, ולהפוך אותם למשאב אסטרטגי

יצירת מאגרי נתונים היא הצעד הראשון. אפילו הכלים החזקים ביותר לא פותרים את בעיית בידוד הנתונים ונתונים בפורמטים שונים, אם החברות ממשיכות לפעול במערכות מבודדות. כאשר הנתונים קיימים בנפרד זה מזה, מבלי להיפגש או להחליף מידע, העסק נתקל בתופעת "סילואי נתונים". במקום תשתית מאוחדת ומסונכרנת, החברות נאלצות לבזבז משאבים על שילוב וסנכרון נתונים

סילו של נתונים והשפעתם על היעילות של החברה

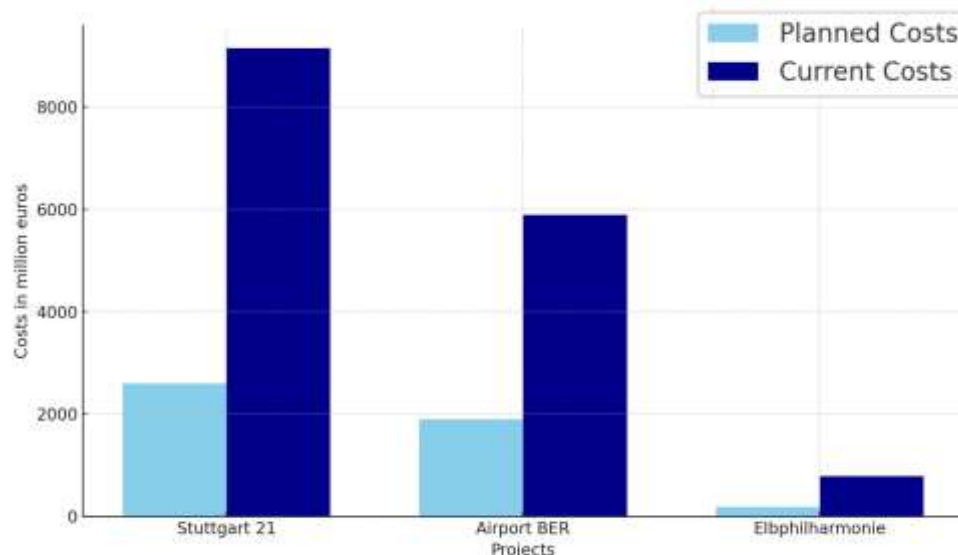
דמיינו שאתם בונים קומפלקס מגורים, אך לכל צוות יש את הפרויקט שלו. חלק בונים קירות, אחרים מניחים תשתיות שלישיים מניחים דרכים, מבלי להתייעץ זה עם זה. כתוצאה מכך, הצינורות לא תואמים לפתחים בקירות, שוחות המעלית לא תואמות לגובה הקומות, והדרכים צריך לפרק ולהניח מחדש.

מצב כזה הוא לא רק תרחיש היפותטי, אלא מציאות של רבים מהפרויקטים הבנייתיים המודרניים. בגלל מספר רב של קבלנים ראשיים ומשניים העובדים עם מערכות שונות וללא מרכז תיאום אחד, התהליך הופך לסדרה אינסופית של אישורים, תיקונים ועימותים. כל זה מביא לעיכובים משמעותיים ולעלויות כפולות של הפרויקטים.

מצב קלאסי שבו באתר הבנייה מתעורר עיכוב: התבנית מוכנה, אך אספקת הברזל לא הגיעה בזמן. כאשר בודקים מידע במערכות שונות, התקשורת מתרחשת בערך כך:

- 🗣️ "המנהל באתר הבנייה ביום ה-20 כותב למנהל הפרויקט: "סיימנו את התקנת התבנית, איפה הברזל
- 🗣️ כתוב שהברזל [PMIS] במחלקת הספקה: — "הקונסטרוקציה מוכנה. במערכת שלי (PMIS) מנהל פרויקט
- 🗣️ "היה צריך להגיע ב-18 לחודש. איפה הברזל
- 🗣️ "שלנו מצוין שהאספקה תהיה ב-25 לחודש ERP-במערכת ה" — (ERP) מומחה הספקה
- 🗣️ 25. — ERP — כתובה התאריך 18, וב PMIS) אחראי על אינטגרציות: — "ב IT מהנדס נתונים או מחלקת
- 🗣️ "ולכן הנתונים לא מסונכרנים. זהו דוגמה טיפוסית לפער מידע, PMIS ל-ERP בין OrderID אין קשר לפי
- 🗣️ "מנהל הפרויקט למנכ"ל — "אספקת הברזל מתעכבת, האתר עומד, ומי נושא באחריות — לא ברור

הסיבה לאירוע הייתה בידוד הנתונים במערכות מבודדות. אינטגרציה ואחידות של מקורות הנתונים, יצירת מאגר מידע מאפשרות להעלים את הפערים בין n8n או Apache NiFi, Airflow ETL אחד, וכן אוטומציה באמצעות כלים המערכות. שיטות וכלים אלו ואחרים יידונו בהרחבה בפרקים הבאים של הספר.

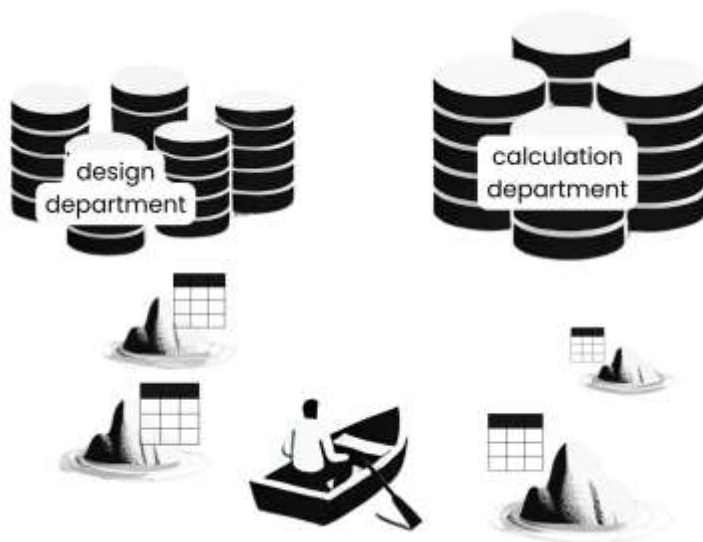


השוואת עלויות מתוכננות ופעילות בפרויקטים תשתיתיים גדולים בגרמניה 2.13. ריש.

אותו דבר קורה עם מערכות ארגוניות: בתחילה נוצרות פתרונות מבודדים, ולאחר מכן יש צורך להוציא תקציבים עצומים על אינטגרציה והתאמה. אם היו מתוכננים מודלים של נתונים וקשרים מהתחלה, הצורך באינטגרציה לא היה קיים כלל.

נתונים מבודדים יוצרים כאוס בעולם הדיגיטלי, בדומה לתהליך בנייה לא מתואם

זמן לפעול או להרוס "משנת 2023, רק 36% מהחברות KPMG "Cue construction 4.0 לפי מחקר
"משתפות נתונים ביעילות בין מחלקות, בעוד ש-61% מתמודדות עם בעיות חמורות בגלל "סילואים
[30]. מבודדים של נתונים



נתונים שנאספים במשך שנים, קשים להוצאה, מצטברים במאגרים מבודדים "סילואים", בסיכון שלא ייעשה בהם 2.14. ריש שימוש לעולם.

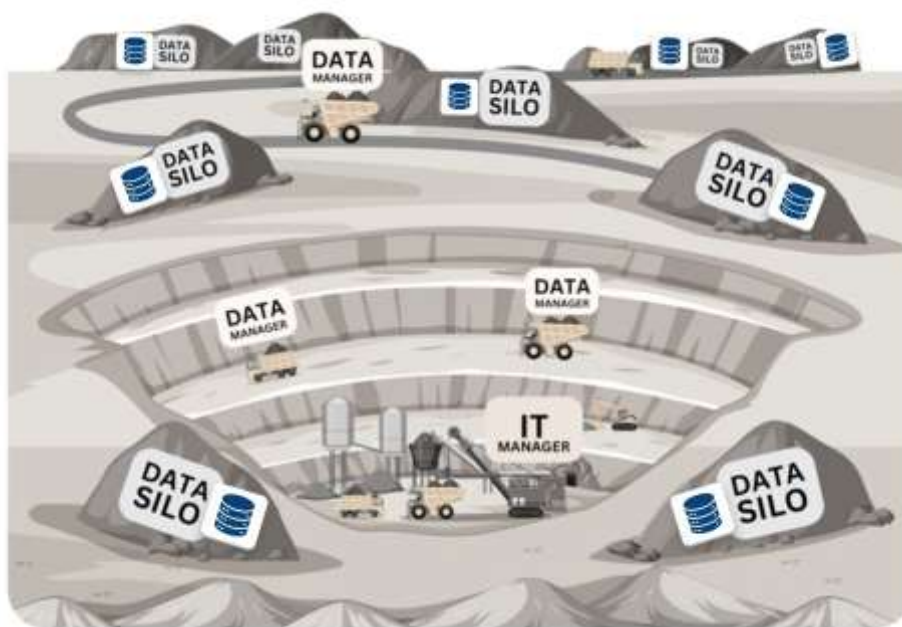
נתוני החברה נשמרים במערכות מבודדות, כמו עצים נפרדים מפוזרים בנוף. כל אחד מהם מכיל מידע יקר ערך, אך חוסר הקשרים ביניהם מונע יצירת מערכת אקולוגית אחת ומקושרת. בידוד זה מפריע לזרימת הנתונים ומגביל את יכולת הארגון לראות את התמונה המלאה. חיבור הסילואים הללו הוא תהליך ארוך ומורכב של גידול מיצל פטריות ברמת הניהול, שצריך ללמוד להעביר פיסות מידע נפרדות בין המערכות.

משנת 2016, אחד המכשולים העיקריים בפני טרנספורמציה דיגיטלית הוא חוסר סטנדרטים אחידים WEF לפי מחקר של נתונים ופיצול

תעשיית הבנייה היא אחת התעשיות המפוצלות ביותר בעולם ותלויה בשיתוף פעולה מסודר של כל המשתתפים [5]. בשרשרת הערך

פורום הכלכלה העולמית 2016: עיצוב עתיד הבנייה –

מעצבים, מנהלים, מתאמים ומפתחים לעיתים קרובות מעדיפים לעבוד באופן עצמאי, נמנעים מקשיים של תיאום. השאיפה הטבעית הזו מביאה ליצירת "סילואים" מידע, שבהם נתונים מבודדים בתוך מערכות נפרדות. ככל שיש יותר מערכות מבודדות כאלה, כך קשה יותר להקים אינטראקציה ביניהן. עם הזמן, כל מערכת מקבלת בסיס נתונים משלה ומחלקת תמיכה מיוחדת ממנהלים, מה שמקשה עוד יותר על האינטגרציה.



כל מערכת שואפת ליצור את הסילו המידע הייחודי שלה, אשר יש לעבד בעזרת כלים מתאימים

המעגל הסגור במערכות ארגוניות נראה כך: חברות משקיעות בפתרונות מבודדים ומורכבים, ואז מתמודדות עם עלויות גבוהות לאינטגרציה שלהם, ומפתחים, המבינים את הקושי של שילוב המערכות, מעדיפים לעבוד באקוסיסטמות ומקשה על המעבר לפתרונות חדשים. מנהלים בסופו של דבר IT-המבודדות שלהם. כל זה מגביר את הפיזור של נוף ה מבקרים את הפיזור של הנתונים, אך לעיתים רחוקות מנתחים את הסיבות לכך ואת הדרכים למנוע זאת. מנהיגים מיושנות, אך החלפתן דורשת השקעות משמעותיות ולעיתים רחוקות מביאה לתוצאה המצופה IT מתלוננים על מערכות כתוצאה מכך, אפילו ניסי להתמודדות עם הבעיה הזו לעיתים קרובות רק מחמירים את המצב.

הסיבה העיקרית לפיזור היא העדפת היישומים על פני הנתונים. חברות קודם כל מפתחות מערכות נפרדות או קונות פתרונות מוכנים מספקים, ואז מנסות לשלב ביניהם, מה שיוצר מאגרי נתונים כפולים ולא תואמים

התמודדות עם בעיית הפיצול דורשת גישה חדשה לחלוטין - העדפת הנתונים על פני היישומים. חברות צריכות קודם כל לפתח אסטרטגיות ניהול נתונים ודגמים של נתונים, ואז ליצור מערכות או לקנות פתרונות שעובדים עם סט מידע אחיד ולא יוצרים מחסומים חדשים

O'Reilly Media, Inc. אנו נכנסים לעולם חדש שבו הנתונים עשויים להיות חשובים יותר מהתוכנה. טים א'ריילי, מנכ"ל

מחקר של מכון מקינזי העולמי "שינוי תפיסה בבנייה: הדרך לשיפור הפרודוקטיביות" (2016) מדגים כי תעשיית הבנייה מפגרת אחרי מגזרים אחרים במעבר לדיגיטל. לפי הדו"ח, יישום ניהול נתונים אוטומטי ופלטפורמות דיגיטליות יכול לשפר משמעותית את הפרודוקטיביות ולהפחית את ההפסדים הקשורים לחוסר התאמה בתהליכים. את הצורך הזה בשינוי דיגיטלי מדגיש גם הדו"ח של איגן (בריטניה, 1998), (המבליט את התפקיד המרכזי של תהליכים משולבים וגישה משותפת בבנייה

בסופו של דבר, אם בעשרת אלפים השנים האחרונות הבעיה העיקרית עבור מנהלי הנתונים הייתה חוסר בנתונים, אז עם הצמיחה המהירה במספר הנתונים ומערכות ניהול הנתונים, המשתמשים ומנהלי הנתונים מתמודדים עם בעיה - עודף נתונים, המקשה על מציאת מידע חוקי ואיכותי.

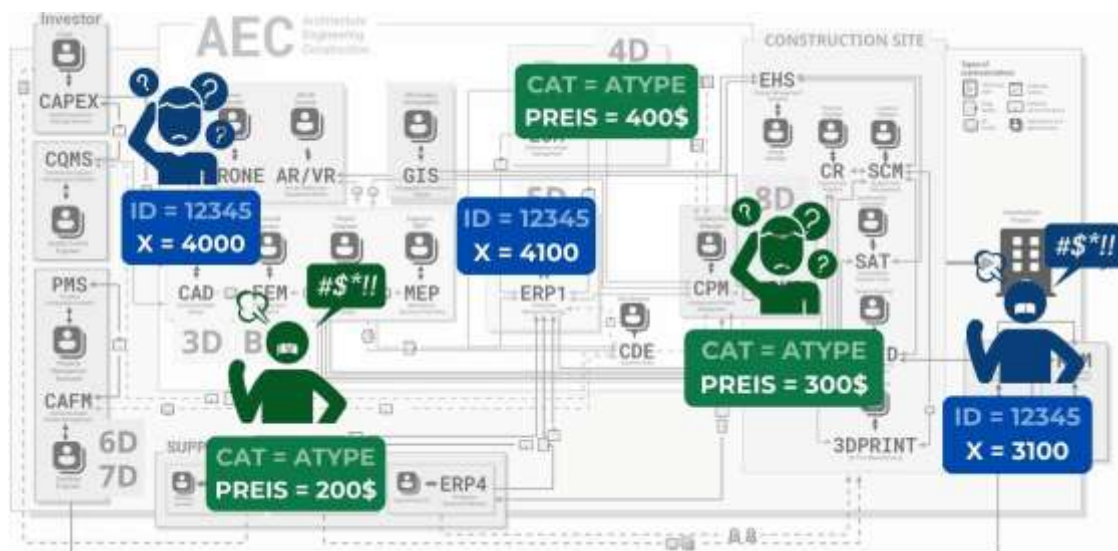
הפיזור של סילונים נתונים מוביל בהכרח לבעיה חמורה של ירידת איכות הנתונים. כאשר ישנן רבות ממערכות עצמאיות אותם נתונים יכולים להתקיים בגרסאות שונות, לעיתים עם ערכים סותרים, מה שיוצר קשיים נוספים למשתמשים, אשר צריכים לקבוע איזו מידע הוא עדכני ואמין.

שכפול וחוסר איכות הנתונים כתוצאה מהנפרדות

בשל בעיית סילו הנתונים, מנהלים נאלצים להקדיש זמן רב לחיפוש ולאיתות נתונים. כדי להימנע מבעיות איכות, חברות יוצרות מבנים מורכבים לניהול מידע, שבהם ההיררכיה של המנהלים אחראית על חיפוש, איתות ותיאום נתונים. עם זאת, גישה זו רק מגבירה את הביורוקרטיה ומאטה את תהליך קבלת ההחלטות. ככל שיש יותר נתונים, כך קשה יותר לנתח ולפרש אותם, במיוחד כאשר אין תקן אחיד לאחסון ולעיבוד שלהם.

עם הופעתם של מספר רב של יישומים ומערכות, שצומחים כמו פטריות אחרי הגשם בעשור האחרון, בעיית הסילו ואיכות הנתונים הלקויה הופכת להיות רלוונטית יותר ויותר עבור המשתמשים הסופיים. ניתן למצוא את אותם נתונים אך עם ערכים שונים, במערכות וביישומים שונים. זה מביא לקשיים עבור המשתמשים הסופיים כאשר הם מנסים לקבוע איזו גרסה של הנתונים היא העדכנית והנכונה מבין רבות הזמינות. זה מוביל לשגיאות בניתוח ובסופו של דבר בקבלת ההחלטות.

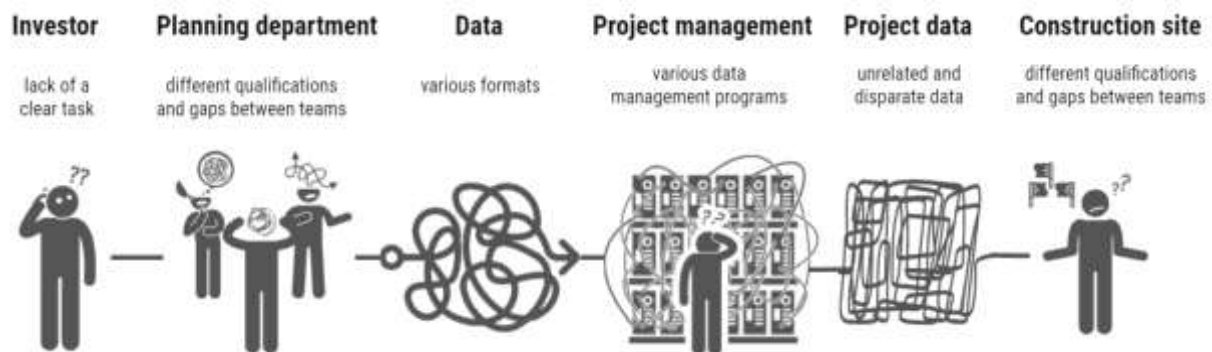
כדי להימנע מבעיות בחיפוש הנתונים הנדרשים, מנהיגי חברות יוצרים ביורוקרטיה רב-שכבתית של מנהלי בדיקה תפקידם הוא לדעת למצוא, לאמת ולשלוח במהירות את הנתונים הנדרשים בצורה של טבלאות ודוחות, תוך כדי ניווט במבוך של מערכות מנותקות.



כאשר מנהלים מנסים למצוא את הנתונים הנדרשים, הם חייבים להבטיח את איכות הנתונים ואת האמינות המשפטית שלהם בין מערכות שונות.

עם זאת, במציאות, מודל זה יוצר קשיים חדשים. כאשר ניהול הנתונים מתבצע ידנית, והמידע מפוזר בין מספר פתרונות לא קשורים, כל ניסיון לקבל מידע מדויק ועדכני דרך פירמידת האחראים הופך לנקודת צוואר בקבוק - בזבזני בזמן - וחשוף לשגיאות.

המצב מחמיר עם העלייה המהירה במספר הפתרונות הדיגיטליים. שוק התוכנה ממשיך להתמלא בכלים חדשים, שנראים מבטיחים. אך ללא אסטרטגיה ברורה לניהול נתונים, פתרונות אלו אינם משתלבים במערכת אחת, אלא להיפך - יוצרים שכבות נוספות של מורכבות וכפילויות. כתוצאה מכך, במקום לפשט את התהליכים, חברות מוצאות את עצמן בסביבה מידע עוד יותר מפורקת וכאוטית.



- מורכבות המערכות ומגוון פורמטי הנתונים מביאים לאובדן עקביות בתהליך הבנייה

כל הבעיות המוזכרות הקשורות לניהול מספר פתרונות מנותקים, בסופו של דבר מביאות את ההנהלה של החברות לתובנה חשובה: העניין אינו בכמות הנתונים ואינו בחיפוש אחר הכלי "הכל יכול" הבא לעיבודם. הסיבה האמיתית טמונה באיכות הנתונים ובאופן שבו הארגון יוצר, מקבל, מאחסן ומשתמש בהם.

המפתח להצלחה מתמשכת אינו במרדף אחרי יישומים "קסומים" חדשים, אלא ביצירת תרבות עבודה עם נתונים בתוך החברה. זה אומר שראיית הנתונים כמשאב אסטרטגי, ושאלות על איכותם, שלמותם ועדכניותם הופכות להיות עדיפות בכל רמות הארגון.

הפתרון לדילמת "איכות מול כמות" טמון ביצירת מבנה נתונים אחיד, שמבטל כפילויות, מסלק סתירות ומאחד את זרמי המידע. ארכיטקטורה כזו מאפשרת ליצור מקור נתונים יחיד, אמין, שעל בסיסו מתקבלות החלטות מבוססות, מדויקות ובזמן.

אחרת, כפי שזה קורה לעיתים קרובות, חברות ממשיכות להסתמך על דעות סובייקטיביות והערכות אינטואיטיביות של ולא על עובדות מהימנות. בתעשיית הבנייה, שבה באופן מסורתי תפקיד משמעותי ממלא הניסיון, HiPPO מומחי המומחיות, זה בולט במיוחד.

או הסכנה של דעות בקבלת החלטות HiPPO

באופן מסורתי, בתעשיית הבנייה, החלטות מפתח מתקבלות על בסיס ניסיון והערכות סובייקטיביות. ללא נתונים מהימנים ובזמן, מנהלי חברות נאלצים לפעול בעיוורון, להסתמך על אינטואיציה של העובדים המרוויחים ביותר

-ולא על עובדות אובייקטיביות, (דעתו של האדם המרוויח ביותר – HiPPO)

NO ANALYTICS?
WELCOME TO THE HIPPO*



*HIGHEST PAID PERSON'S OPINION

בהיעדר אנליטיקה, העסק תלוי בדעתם הסובייקטיבית של מומחים מנוסים 2.18. ריש

גישה כזו, אולי, מוצדקת בתנאים של יציבות ושינויים איטיים, אך בעידן של טרנספורמציה דיגיטלית היא הופכת לסיכון משמעותי. החלטות המבוססות על אינטואיציה והשערות חשופות לעיוותים, לעיתים קרובות מבוססות על השערות לא מאושרות ואינן לוקחות בחשבון את התמונה המורכבת המוצגת בנתונים.

מה שמוצג בחברה ברמת קבלת ההחלטות כוויכוחים רציונליים, לעיתים קרובות אינו מבוסס על דבר קונקרטי. הצלחת החברה לא צריכה להיות תלויה בסמכות ובשכר של המומחים, אלא ביכולת לעבוד ביעילות עם נתונים, לזהות דפוסים ולקבל החלטות מושכלות.

חשוב להפסיק את הקונספט שבו סמכות או ניסיון אוטומטית משמעם נכונות ההחלטה. גישה מבוססת נתונים משנה את כללי המשחק: כעת הבסיס לקבלת החלטות הוא נתונים ואנליטיקה, ולא תפקיד ושכר. נתונים גדולים, למידת מכונה -ואנליטיקה חזותית מאפשרים לזהות דפוסים ולהסתמך על עובדות, ולא על השערות.

ללא נתונים, אתה פשוט עוד אדם עם דעה. 1. אדוארדס דמינג, מדען ויועץ ניהול

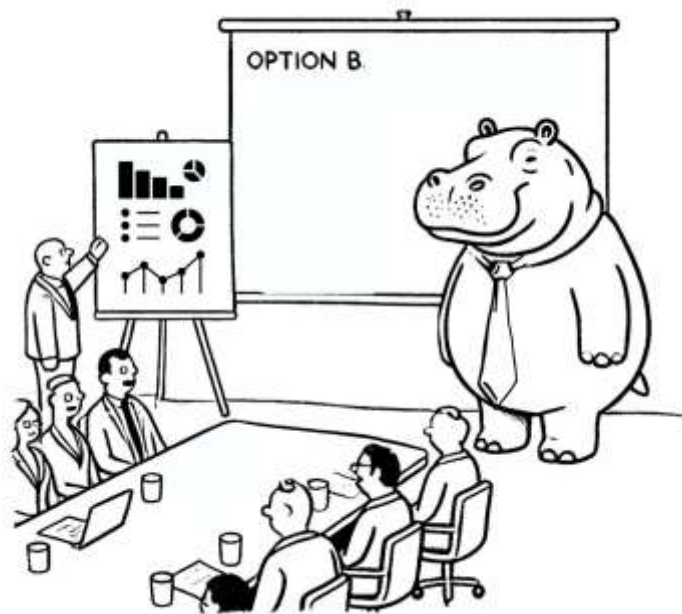
שיטות ניהול נתונים מודרניות גם מבטיחות רציפות ידע בחברה. תהליכים מתוארים בבירור, אוטומציה וגישה מערכתית מאפשרים להעביר אפילו תפקידים מרכזיים מבלי לאבד יעילות.

עם זאת, אמון עיוור בנתונים יכול גם להוביל לטעויות חמורות. הנתונים עצמם הם רק אוסף של מספרים. ללא ניתוח נכון, הקשר ויכולת לזהות דפוסים, הם חסרי ערך ואינם יכולים לנהל תהליכים. המפתח להצלחה אינו בבחירה בין לאנליטיקה, אלא בבניית כלים אינטליגנטיים שממירים מידע מפוזר להחלטות ניהוליות HiPPO אינטואיציה של מבוססות.

בתנאי בנייה דיגיטלית, גורמי הצלחה מכריעים הם לא ותק בעבודה ומקום בהיררכיה, אלא מהירות תגובה, דיוק החלטות ויעילות השימוש במשאבים.

נתונים הם כלי, ולא אמת מוחלטת. הם צריכים להשלים את החשיבה האנושית, ולא להחליף אותה. למרות כל היתרונות של ניתוח נתונים, הם אינם יכולים להדיח לחלוטין את האינטואיציה והניסיון האנושי. תפקידם הוא לסייע בקבלת החלטות מדויקות ומודעות יותר.

יתרון תחרותי יושג לא רק על ידי עמידה בסטנדרטים, אלא על ידי היכולת להקדים את המתחרים ביעילות השימוש במשאבים זהים לכולם. בעתיד, מיומנות העבודה עם נתונים תהפוך לחשובה כמו שהייתה פעם אוריינות או שליטה במתמטיקה. אנשי מקצוע שמסוגלים לנתח ולפרש נתונים יוכלו לקבל החלטות מדויקות יותר, תוך דחיקת אלו התלויים רק על ניסיון אישי.



החלטות צריכות להתבסס על ניתוח אובייקטיבי, ולא על דעתו של העובד המרוויח ביותר.

מנהלים, אנשי מקצוע ומהנדסים יפעלו כאנליסטים של נתונים, חוקרים את המבנה, הדינמיקה והמדדים המרכזיים של פרויקטים. המשאבים האנושיים יהפכו לרכיבים של מערכת, שדורשים התאמה גמישה על בסיס נתונים כדי להשיג את היעילות המרבית.

טעויות בשימוש בנתונים לא הולמים הרבה פחות מאשר בהיעדר נתונים. – צ'ארלס בבג'ידג, 'ממציא מכונית החישוב האנליטית הראשונה'

הופעת נתונים גדולים והטמעת מודלים לשוניים גדולים שינו באופן רדיקלי לא רק את דרכי הניתוח, אלא גם את מהות קבלת ההחלטות. אם בעבר המוקד היה על סיבתיות (למה משהו קרה – ניתוח דיאגנוסטי, (כיום המוקד הוא על היכולת לחזות את העתיד) ניתוח חיזוי, (ובעתיד – גם ניתוח פרסקרפטיבי, שבו למידת מכונה ובינה מלאכותית מציעות את הבחירה הטובה ביותר בתהליך קבלת ההחלטות.

מחקר חדש הראה שכמעט חצי מהמנהיגים סומכים על בינה מלאכותית יותר", SAP™ לפי מחקר חדש של מאשר על עצמם "בשנת 2025, 44% מהמנהיגים הבכירים מוכנים לשנות את החלטותיהם הקודמות על סמך המלצות של בינה מלאכותית, ו-38% היו סומכים על בינה מלאכותית לקבל החלטות עסקיות בשמם עם זאת, 74% מהמנהיגים הצהירו שהם סומכים יותר על עצות של בינה מלאכותית מאשר על חברים ובני משפחה, ו-55% עובדים בחברות שבהן תובנות שהושגו בעזרת בינה מלאכותית מחליפות או לעיתים קרובות עוקפות את השיטות המסורתיות לקבלת החלטות – במיוחד בארגונים עם הכנסות שנתיות של מעל מיליארד דולר. בנוסף, 48% מהנשאלים משתמשים בכלים של בינה מלאכותית גנרטיבית על בסיס יומי 5 מתוכם 15% – מספר פעמים ביום.

עם התפתחות מודלים לשוניים גדולים ומערכות אוטומטיות לניהול נתונים, עולה בעיה חדשה: כיצד להשתמש ביעילות במידע מבלי לאבד את ערכו בתוך הכאוס של פורמטים לא תואמים ומקורות מגוונים, דבר שמחמיר עם המורכבות והדינמיקה הגוברת של תהליכים עסקיים.

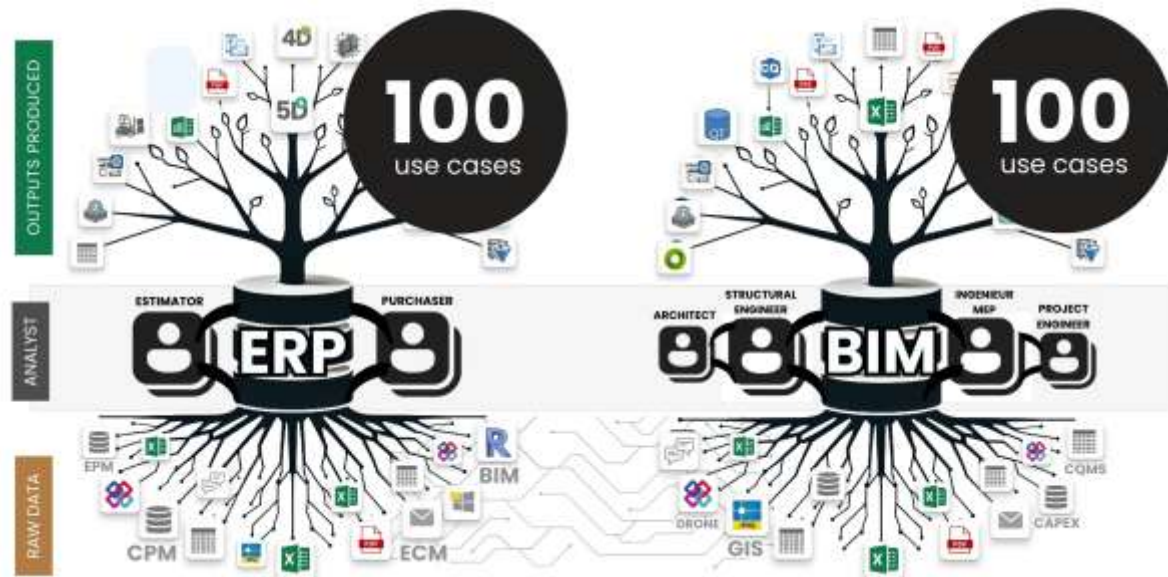
עלייה מתמדת במורכבות ובדינמיות של תהליכי עסק

תעשיית הבנייה היום מתמודדת עם אתגרים משמעותיים בניהול נתונים ותהליכים. הקשיים העיקריים הם פיזור מערכות המידע, בירוקרטיה מופרזת וחוסר אינטגרציה בין הכלים הדיגיטליים. בעיות אלו מחמירות ככל שתהליכי העסקים עצמם הופכים למורכבים יותר - בהשפעת טכנולוגיות, דרישות משתנות של לקוחות ותקנות מתעדכנות.

הייחודיות של פרויקטי הבנייה נובעת לא רק מהמאפיינים הטכניים שלהם, אלא גם מההבדלים בסטנדרטים הלאומיים ובדרישות הרגולטוריות של מדינות שונות. זה דורש גישה גמישה ומותאמת אישית לכל פרויקט, דבר שקשה ליישם במסגרת מערכות ניהול מודולריות מסורתיות. בשל המורכבות של התהליכים והכמות הגדולה של הנתונים, רבות מהחברות פונות לספקים המציעים פתרונות מיוחדים. אך השוק עמוס - מספר רב של סטארטאפים מציעים מוצרים -דומים, מתמקדים במשימות צרות. כתוצאה מכך, לעיתים קרובות מאבדים את הגישה הכוללת לניהול הנתונים.

ההתאמה לזרם המתמשך של טכנולוגיות חדשות ודרישות השוק הופכת לגורם קריטי לתחרותיות. עם זאת, היישומים הקנייניים הקיימים ומערכות המודולים בעלות יכולת התאמה נמוכה - כל שינוי דורש לעיתים קרובות שיפוטים ארוכים ויקרים מצד המפתחים, שלא תמיד מבינים את המאפיינים של תהליכי הבנייה.

חברות מוצאות את עצמן כעבדי הפיגור הטכנולוגי, ממתינות לעדכונים חדשים במקום ליישם גישות אינטגרטיביות חדשניות במהירות. כתוצאה מכך, המבנה הפנימי של ארגוני הבנייה לעיתים קרובות מציג מערכת אקולוגית מורכבת של מערכות היררכיות, ולעיתים קרובות סגורות, שהקישור ביניהן מתבצע דרך רשת ניהול רב-שכבתית.

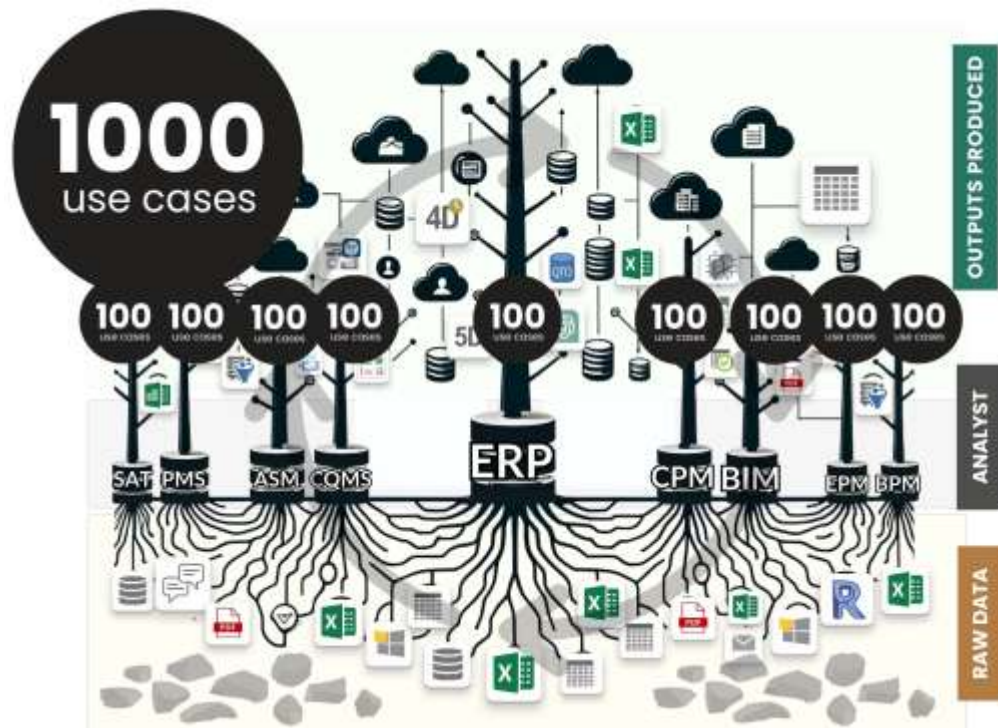


חברות מורכבות ממערכות מקושרות, שהאיחוד שלהן יוצר תהליכים הדורשים אוטומציה.

בקנדה בשנת 2021, רק 25% מהחברות KPMG לפי מחקר שנערך על ידי האגודה הקנדית לבנייה וחברת סבורות שהן נמצאות במצב משמעותי או מצוין בהשוואה למתחרים בתחום יישום הטכנולוגיות או הפתרונות הדיגיטליים. רק 23% מהנשאלים דיווחו שהפתרונות שלהם מבוססים במידה משמעותית או רבה על נתונים. עם זאת, רוב המשתתפים בסקר תיארו את השימוש שלהם בטכנולוגיות אחרות כנסינוני בלבד או הודו שהם כלל לא משתמשים בהן.

חוסר הרצון להשתתף בניסויים טכנולוגיים מתבטא במיוחד בפרויקטים תשתיתיים גדולים, שבהם טעויות עשויות לעלות מיליוני דולרים. אפילו הטכנולוגיות המתקדמות ביותר - דיגיטליים תאומים, ניתוח חיזוי - לעיתים קרובות נתקלות בהתנגדות לא בגלל היעילות שלהן, אלא בגלל חוסר אמינות מוכחת בפרויקטים אמיתיים.

עיצוב עתיד הבנייה, "יישום טכנולוגיות חדשות בתחום" (WEF) בהתאם לדו"ח של הפורום הכלכלי העולמי הבנייה נתקל לא רק בקשיים טכניים, אלא גם במחסום פסיכולוגי מצד המזמינים. רבים מהלקוחות חוששים כי השימוש בפתרונות מתקדמים יהפוך את הפרויקטים שלהם לזירה ניסיונית ויהפוך אותם ל"ארנבות ניסוי", "וההשלכות הבלתי צפויות עשויות להוביל להוצאות נוספות ולסיכונים



עבור כל מקרה שימוש, שוק הפתרונות מציע אפליקציות לאופטימיזציה ואוטומציה של תהליכים.

תעשיית הבנייה מגוונת מאוד: לפרויקטים שונים יש דרישות שונות, מאפיינים אזוריים, תקנות חוקיות וכו'. לכן כמעט בלתי אפשרי ליצור אפליקציה או (5.17 ריש) סטנדרטים לחישובים, (4.210 ריש) -- מערכת קניינית אוניברסלית שתתאים באופן מושלם לכל הדרישות והמאפיינים של הפרויקטים

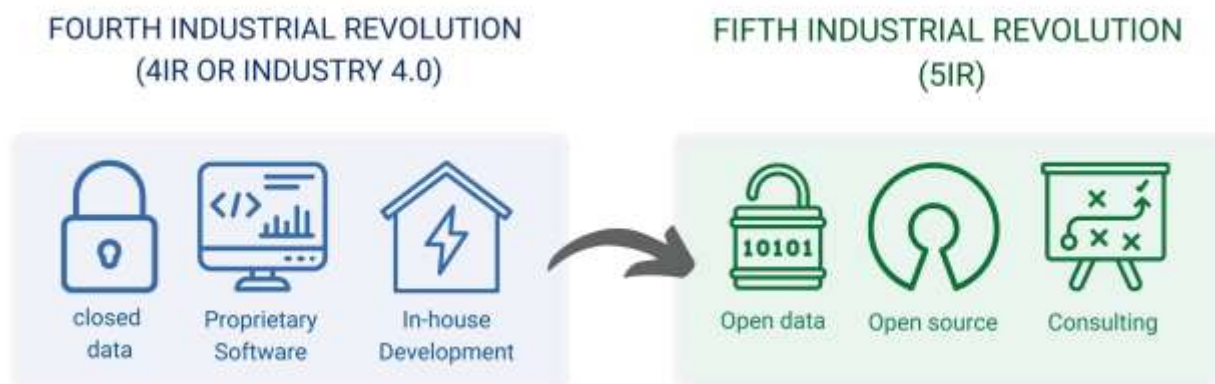
בניסיון להתמודד עם המורכבות ההולכת ומתקדמת של המערכות ותלות בספקי תוכנה, יותר ויותר מגיעים להבנה כי המפתח לניהול נתונים יעיל הוא לא רק שקיפות וסטנדרטיזציה, אלא גם פשטות של הארכיטקטורה של התהליכים עצמם המורכבות והדינמיות ההולכות וגדלות של תהליכי העסקים דורשות גישות חדשות, שבהן העדיפות מועברת מהצטרות נתונים לארגון וסידור שלהם. בדיוק שינוי זה יהווה את הצעד הבא בהתפתחות תעשיית הבנייה, מסמן את סוף עידן שלטון ספקי התוכנה ואת תחילת עידן של ארגון מידע משמעותי

ההבנה של מגבלות הפתרונות האוניברסליים ופגיעות מול העלייה במורכבות מביאה לשינוי עדיפויות: מפלטפורמות סגורות והצטרות נתונים - לשקיפות, אדפטיביות ועבודה מסודרת עם מידע. שינוי זה במחשבה משקף שינויים רחבים יותר בכלכלה ובטכנולוגיה הגלובלית, המתוארים דרך עדשת מה שנקרא "מהפכות תעשייתיות". כדי להבין לאן מתקדמת הבנייה ומהן הכיוונים העתידיים שלה, יש לבחון את מקומה של התעשייה בהקשר של המהפכה התעשייתית הרביעית והחמישית - מהאוטומציה והדיגיטציה ועד אישית, סטנדרטים פתוחים ומודל שירותי נתונים

המהפכה התעשייתית הרביעית) תעשייה 4.0 (והמהפכה התעשייתית החמישית) תעשייה 5.0 (בבנייה

תהליכים טכנולוגיים וכלכליים - אלו הן מושגים תיאורטיים המשמשים לתיאור ולניתוח האבולוציה של החברה והכלכלה בשלבים שונים של ההתפתחות. עם זאת, חוקרים ומומחים שונים עשויים לפרש אותם בדרכים שונות

- קשורה לטכנולוגיות מידע, אוטומציה, דיגיטציה (Industry 4.0 או 4IR) המהפכה התעשייתית הרביעית וגלובליזציה. אחד מהמרכיבים המרכזיים שלה הוא יצירת פתרונות תוכנה קנייניים, כלומר מוצרים דיגיטליים מיוחדים שנועדו למשימות ולחברות ספציפיות. פתרונות אלו לעיתים קרובות הופכים לחלק חשוב מהתשתית הטכנולוגית, אך יחד עם זאת הם מתקשים להתרחב ללא שינויים נוספים.
- נמצאת כיום בשלב מוקדם יותר של קונספטואליזציה ופיתוח מאשר (5IR) המהפכה התעשייתית החמישית העקרונית המרכזיים שלה כוללים הגברת רמת ההתאמה האישית של (4IR). המהפכה התעשייתית הרביעית היא תנועה לעבר פעילות כלכלית יותר מותאמת, גמישה ואישית, עם דגש על IR מוצרים ושירותים. 5. התאמה אישית, ייעוץ ומודלים מונחי שירות. היבט מרכזי של הסדר הכלכלי החמישי הוא השימוש בנתונים - לקבלת החלטות, דבר שאינו אפשרי כמעט ללא שימוש בנתונים פתוחים ובכלים פתוחים) איור 2.112



- איור 2.112 הסדר הרביעי מתמקד בפתרונות, והחמישי - בהתאמה אישית ובנתונים

יצירת אפליקציה עבור חברות בתעשיית הבניין, המיועדת לשימוש בעשרה או מאה ארגונים, אינה מבטיחה את הצלחתה בהרחבה לחברות אחרות, אזורים או מדינות ללא שינויים ושיפורים משמעותיים הסיכוי להרחבה מוצלחת של פתרונות כאלה נשאר נמוך, מכיוון שכל ארגון יש לו תהליכים, דרישות ותנאים ייחודיים שעשויים לדרוש התאמות מותאמות אישית

חשוב להבין שכבר היום, אינטגרציה מוצלחת של פתרונות טכנולוגיים כרוכה בגישה מותאמת אישית עמוקה לכל תהליך פרויקט וחברה. זה אומר שגם לאחר פיתוח מבנה, כלי או תוכנה אוניברסליים, תידרש התאמה מפורטת והגדרה לפי הדרישות והמצבים הייחודיים של כל חברה ופרויקט ספציפי

פיענוח המהפכה התעשייתית החמישית] "38, כ-50% מהמנכ"לים בתעשיות שונות השנה" PwC לפי דוח מהמרים על אינטגרציה של טכנולוגיות מתקדמות וניסיון אנושי. גישה זו מאפשרת להסתגל במהירות לשינויים בעיצוב המוצר או בדרישות הלקוח, וליצור ייצור מותאם אישית

עבור כל תהליך נדרשת פיתוח פונקציה או אפליקציה ייחודית, דבר שמוביל, בהתחשב בהיקף תעשיית הבניין העולמית ייחודית) איור Pipeline ובמגוון הפרויקטים, לקיומם של מספר עצום של מקרים עסקיים, שכל אחד מהם מהווה לוגיקת כל מקרה כזה יש לו מאפיינים משלו ודורש גישה אישית. נבחן ביתר פירוט את המגוון של פתרונות אפשריים. (2.113) לאותה בעיה אנליטית בהקשר של גישות שונות בפרק המוקדש ללמידת מכונה ולניתוח מערך הנתונים "טיטניק" איור -- 9.29.

בהקשר של תהליכים דיגיטליים הוא רצף של פעולות, תהליכים וכלים המבטיחים זרימה Pipeline אוטומטית או מובנית של נתונים ועבודות בשלבים שונים של מחזור חיי הפרויקט



איור 2.113 היחודיות והווריאביליות של מקרים עסקיים מקשה על ניסיונות ליצור פלטפורמות - וכלים סגורים הניתנים להרחבה

חיינו כבר השתנו במידה רבה תחת השפעתה של טרנספורמציה דיגיטלית, וכיום ניתן לדבר על תחילתו של שלב חדש בהתפתחות הכלכלית של תעשיית הבניין. בכלכלה "החדשה" הזו, התחרות תתנהל לפי כללים שונים: מי שיכול להפוך ביעילות ידע ציבורי ונתונים פתוחים למוצרים ושירותים מבוקשים, מקבל יתרון מרכזי בתנאים של המהפכה התעשייתית החמישית.

כפי שמציינת הכלכלנית קייט מסקוס בספרה "זכויות פרטיות ובעיות ציבוריות: הכלכלה הגלובלית של קניין רוחני במאה ה-21" משנת 2012, אנו חיים בכלכלה גלובלית של ידע, והעתיד שייך לאלה שמסוגלים "להפוך גילויים מדעיים למוצר".

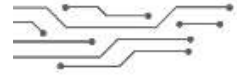
סגורים לסטנדרטים ופלטפורמות פתוחות. חברות יתחילו IT המעבר למבנה כלכלי חמישי שינוי מיקוד מפתרונות להתרחק מהמוצרים התוכנתיים המסורתיים לטובת מודלים מונחי שירות, שבהם הנכס העיקרי יהיה נתונים ולא טכנולוגיות קנייניות.

מחקר של בית הספר לעסקים של הרווארד משנת 2024 מראה את הערך הכלכלי העצום של תוכנה בקוד נוכח ב-96% מכלל הקודים התוכנתיים OSS, לפי המחקר. (Open Source Software, OSS) פתוח חברות היו מוציאות פי 3.5 יותר OSS, ללא OSS וחלק מהתוכנה המסחרית מורכבת ב-99.9% מרכיבי על תוכנה.

האקוסיסטמות של חברות הבנייה, בעקבות מגמות עולמיות, יעברו בהדרגה למבנה כלכלי חמישי, שבו שירותי ניתוח וייעוץ מבוססי נתונים יהיו עדיפים על פני פתרונות סגורים מבודדים עם תרחישים קבועים.

עידן הדיגיטציה ישנה את מאזן הכוחות בתעשייה: במקום תלות בפתרונות של ספקים, חברות יבנו את התחרותיות שלהן על היכולת להשתמש בנתונים ביעילות. כתוצאה מכך, תעשיית הבנייה תעבור ממערכות נוקשות מיושנות לאקוסיסטמות גמישות ואדפטיביות, שבהן סטנדרטים פתוחים וכלים תואמים יהפכו לבסיס לניהול פרויקטים. סוף עידן הדומיננטיות של ספקי היישומים ייצור תנאים חדשים, שבהם הערך ייקבע לא על ידי בעלות על קוד סגור וחיבורים

מיוחדים, אלא על ידי היכולת להפוך נתונים ליתרון אסטרטגי.



פרק 2.2. הפיכת הכאוס לסדר והפחתת המורכבות

קוד מיותר ומערכות סגורות כמחסום לשיפור הפרודוקטיביות

הוגדרו בעיקר על ידי ספקי התוכנה. הם קבעו את כיוון IT-בעשורים האחרונים, השינויים הטכנולוגיים בתחום ה ההתפתחות, קובעים אילו טכנולוגיות על חברות לאמץ ואילו להשאיר בצד. בעידן המעבר מפתרונות מבודדים לבסיסי נתונים מרכזיים ומערכות משולבות, ספקים קידמו מוצרים עם רישוי, מבטיחים שליטה על הגישה וההרחבה. מאוחר שליטה זו הפכה למודל מנוי, מה שמקבע, Software as a Service (SaaS) יותר, עם הגעת הטכנולוגיות בענן ודגמי את המשתמשים בתפקיד של לקוחות קבועים של שירותים דיגיטליים

גישה זו יצרה פרדוקס: למרות כמויות חסרות תקדים של קוד תוכנה שנוצר, בפועל נעשה שימוש רק בחלק קטן ממנו ייתכן שנכתב קוד במאות או אלפי פעמים יותר ממה שנדרש, שכן אותם תהליכי עסקים מתוארים ומועתיקים בעשרות או מאות תוכניות בדרכים שונות - אפילו בתוך חברה אחת. עם זאת, על הפיתוח כבר שולם, וההוצאות הללו אינן ניתנות להחזרה. ובכל זאת, התעשייה ממשיכה לשחזר את המעגל הזה, יוצרת מוצרים חדשים עם ערך מוסף מינימלי למשתמש הסופי, לעיתים קרובות תחת לחץ של ציפיות שוק יותר מאשר צרכים אמיתיים

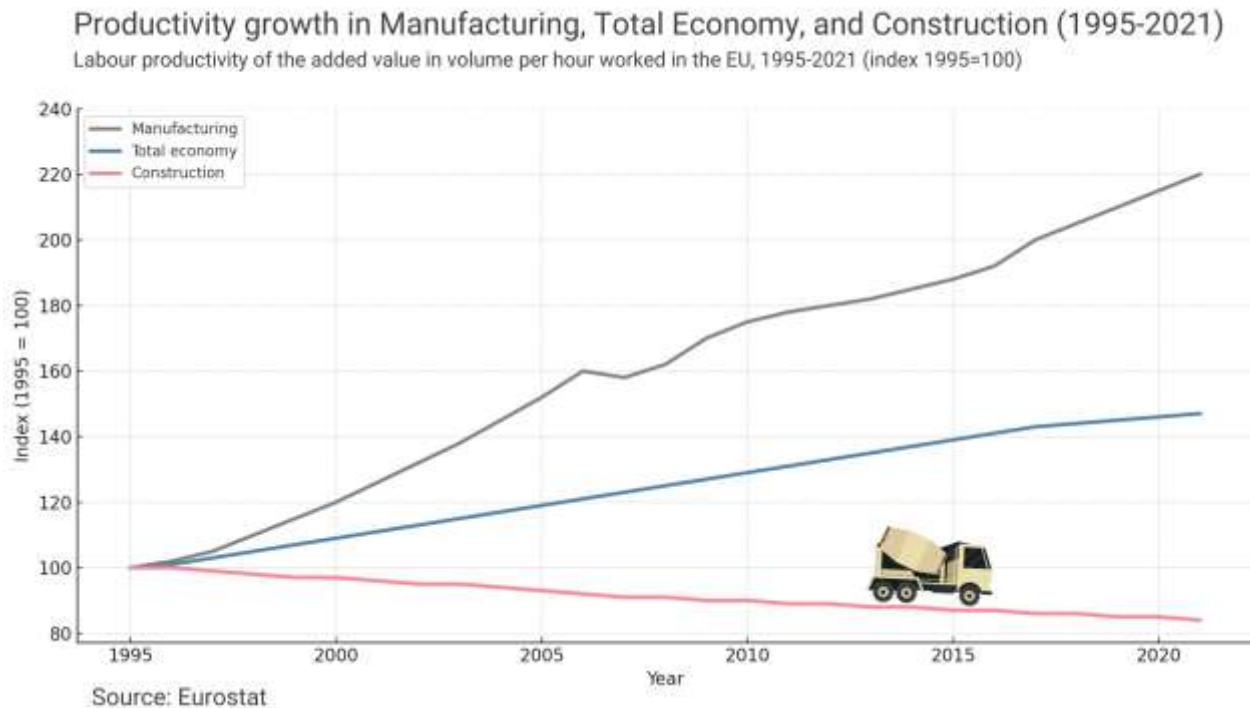
עלות (DAU), בהתאם למדריך להערכת עלויות פיתוח תוכנה שנערך על ידי אוניברסיטת רכש ההגנה פיתוח תוכנה עשויה להשתנות באופן משמעותי בהתאם למספר גורמים, כולל מורכבות המערכת והטכנולוגיות הנבחרות. היסטורית, עלות הפיתוח בשנת 2008 עמדה על כ-100 דולר לכל שורת קוד מקור SLOC. בעוד שעלויות התחזוקה עשויות לעלות עד 4,000 דולר לכל (SLOC).

הליבה הגיאומטרית - עשוי לכלול עשרות מיליוני שורות קוד. מצב דומה קיים גם - CAD רק אחד מרכיבי יישומי שעל מורכבותו נדון בהמשך בפרק החמישי של הספר. עם זאת, בבחינה מעמיקה יותר מתברר: חלק, ERP במערכות ממשק, API, משמעותי מהקוד הזה אינו יוצר ערך מוסף, אלא פשוט פועל כ"שליח" - מעביר נתונים בין מסד הנתונים, המשתמש וטבלאות אחרות במערכת. למרות המיתוס הפופולרי על החשיבות הקריטית של מה שמכונה לוגיקת עסק המציאות הקשה הרבה יותר פרוזאית: בסיסי הקוד המודרניים עמוסים בחלקי קוד ישנים, שמטרתם היחידה היא - להבטיח העברת נתונים בין טבלאות ורכיבים, מבלי להשפיע על קבלת החלטות או על עליית היעילות העסקית "בסופו של דבר, פתרונות סגורים, העוסקים בעיבוד נתונים ממקורות שונים, הופכים בהכרח למערכות "ספגטי מסובכות. עם המערכות המורכבות והמסובכות הללו יכולה להתמודד רק צבא של מנהלים, הפועל במצב חצי ידני ארגון כזה של ניהול נתונים אינו רק לא יעיל מבחינת משאבים, אלא גם יוצר נקודות תורפה קריטיות בתהליכים העסקיים, מה שהופך את החברה לתלויה בקבוצה מצומצמת של מומחים, המבינים כיצד פועל הלבנת הטכנולוגיה הזו

העלייה המתמדת בכמות הקוד, מספר היישומים והמורכבות של המושגים המוצעים על ידי ספקים, הביאה לתוצאה בענף הבנייה. זה הפך את היישום המעשי של דיגיטציה דרך הגדלת מספר IT-צפויה - עלייה במורכבות של המערכת ה היישומים בענף לבלתי יעיל. מוצרים תוכנותיים, שנוצרו מבלי לשים לב לצרכי המשתמשים, דורשים לעיתים קרובות משאבים משמעותיים ליישום ולתחזוקה, אך אינם מביאים את התשואה הצפויה

על פי מחקר של מקינזי "שיפור הפרודוקטיביות בבנייה", "במהלך שני העשורים האחרונים, הצמיחה הגלובלית בפרודוקטיביות העבודה בבנייה עמדה בממוצע על 1% בלבד בשנה, בהשוואה לצמיחה של בכלכלה העולמית ו-3.6% בתעשייה המפעלי. בארצות הברית, פרודוקטיביות העבודה בבנייה לכל 2.8% עובד ירדה בחצי מאז שנות ה-60

עליית המורכבות של המערכות, הבידוד והסגירות של הנתונים פגעו בתקשורת בין המומחים, מה שהפך את ענף הבנייה לאחד הענפים הפחות יעילים.



סגירות ומורכבות הנתונים וככתוצאה מכך תקשורת לקויה בין המומחים הביאו את ענף הבנייה לאחד הענפים הפחות יעילים בכלכלה.

כפי שמדגיש מחקר מקינזי (2024), (הבטחת פרודוקטיביות הבנייה כבר אינה דבר שאינו הכרחי, "בתנאים של מחסור הולך ומתרחק במשאבים ושאפת הענף להכפיל את קצב הצמיחה, הבנייה כבר אינה יכולה להרשות לעצמה להישאר ברמת פרודוקטיביות הנוכחית. על פי תחזיות, ההוצאות הגלובליות על בנייה יגדלו מ-13 טריליון דולר בשנת 2023 ל-22 טריליון דולר עד שנת 2040, מה שהופך את שאלת היעילות לא רק לרלוונטית, אלא גם לקריטית.

אחד מהדרכים המרכזיות לשיפור היעילות יהיה האחידות והפשטה הבלתי נמנעת של מבני היישומים ואדריכלות האקוסיסטמות המעבדות נתונים. גישה זו לרציונליזציה תאפשר להיפטר משכבות אבסטרקציה מיותרות ומסיבוכיות לא נחוצה, שהצטברו במשך שנים במערכות הארגוניות.

מסילו לאחסון נתונים מאוחד

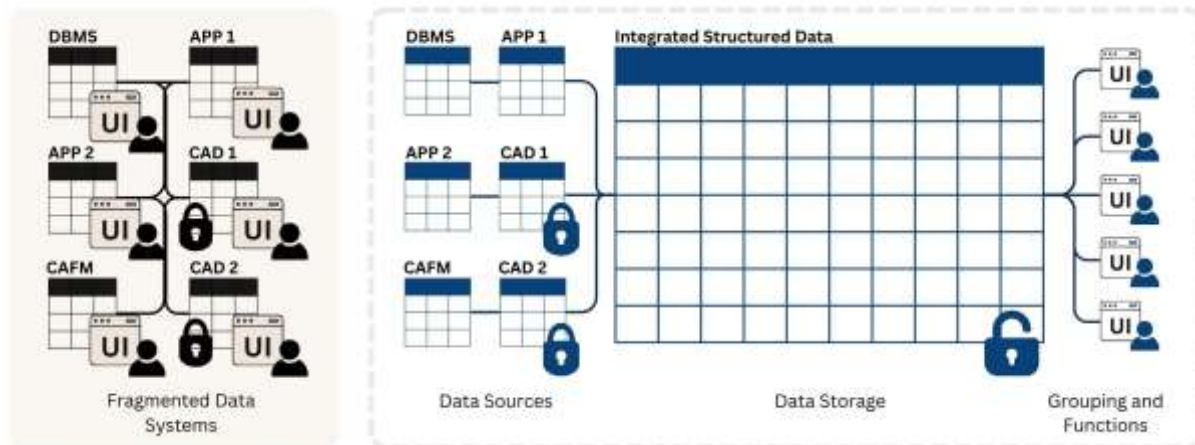
ככל שארגון צובר יותר נתונים, כך קשה יותר להפיק מהם תועלת אמיתית. בשל הפיצול של אחסון המידע בסילונים מבודדים, חברות מודרניות בתהליכי העבודה שלהן מזכירות בונים המנסים לבנות גורד שחקים מחומרים המאוחסנים באלפי מחסנים שונים. עודף המידע לא רק מקשה על הגישה למידע משפטי משמעותי, אלא גם מעכב את קבלת ההחלטות. כל צעד צריך להיבדק ולאשר מספר פעמים.

כל משימה או תהליך קשורים באופן הדוק לטבלה או למסד נתונים נפרד, והחלפת נתונים בין מערכות דורשת אינטגרציות מורכבות. שגיאות וחוסר התאמה במערכת אחת יכולות לגרום לכשלים בשרשרת במערכות אחרות. ערכים לא נכונים, עדכונים מאוחרים ודופליקציות של מידע מאלצים את העובדים לבזבז זמן רב על בדיקות ידניות ואישור נתונים. בסופו של דבר, הארגון מבזבז יותר זמן על תיקון תוצאות הפיצול מאשר על פיתוח ואופטימיזציה של תהליכים.

הבעיה הזו היא אוניברסלית: חלק מהחברות ממשיכות להיאבק עם הכאוס, אחרות מוצאות פתרון באינטגרציה - העברת זרמי המידע למערכת אחסון מרכזית. דמיינו זאת כמו טבלה גדולה אחת, שבה ניתן לאחסן כל ישות הקשורה למשימות, פרויקטים ואובייקטים. במקום עשרות טבלאות ופורמטים מבודדים, מופיעה מערכת אחסון אחת מקושרת (איור 2.22):

- למזער אובדן נתונים;
- להפסיק את הצורך באישור מתמיד של מידע;
- לשפר את זמינות ואיכות הנתונים;
- לפשט את העיבוד האנליטי ולמידת מכונה.

הבאת הנתונים לסטנדרט אחד משמעותה, שבין אם מדובר במקור, המידע מומר לפורמט מאוחד וקריא על ידי מכונה. ארגון כזה של נתונים מאפשר לבדוק את שלמותם, לנתח בזמן אמת ולהשתמש בהם במהירות לקבלת החלטות ניהוליות. נרחיב על מושג מערכות האחסון המוטמעות והשימוש שלהן באנליטיקה ולמידת מכונה בפרק "אחסון נתונים גדולים ולמידת מכונה". נושאי המידול והמבנה של הנתונים יוסרו בפירוט בפרקים "המרת נתונים לצורת מבנה" ו"כיצד הסטנדרטים משנים את המשחק: מקבצים אקראיים למודל נתונים מתוכנן".



איור 2.22 אינטגרציית נתונים מסירה את הבידוד, משפרת את זמינות המידע ומיעלת את תהליכי העבודה.

לאחר ארגון ואיחוד הנתונים, השלב הלוגי הבא הוא בדיקתם. כאשר ישנו מאגר נתונים אינטגרטיבי אחד, התהליך הזה מתפשט באופן משמעותי: אין יותר מספר רב של סכמות לא תואמות, מבנים כפולים וקשרים מורכבים בין הטבלאות. כל המידע מובא למודל נתונים אחד, מה שמסלק סתירות פנימיות ומאיץ את תהליך האימות. בדיקה והבטחת איכות הנתונים הם היבטים מרכזיים בכל תהליכי העסקים, ונעמיק בהם בפרקים המתאימים של הספר.

בשלב הסופי, הנתונים מקובצים, מסוננים ומנותחים. אליהם מוחלות פונקציות שונות (אגרגציה) חיבור, כפל, (חישובים בין טבלאות, עמודות או שורות. העבודה עם הנתונים הופכת לרצף של צעדים: איסוף, ארגון, בדיקה, טרנספורמציה, עיבוד אנליטי וטעינה לאפליקציות הסופיות, שבהן המידע משמש לפתרון בעיות מעשיות. נעמיק בבניית תרחישים כאלה - data pipeline של approach ETL אוטומציה של צעדים ובניית זרמי עיבוד בפרקים המוקדשים לתהליכי

כך, טרנספורמציה דיגיטלית אינה רק פישוט העבודה עם מידע. מדובר בהפסקת המורכבות המיותרת בניהול נתונים במעבר מהכאוס לחיזוי, ממספר מערכות לתהליך מנוהל. ככל שהמורכבות של הארכיטקטורה נמוכה יותר, כך נדרש פחות קוד לתמיכה. ובמבט לעתיד, הקוד עצמו עשוי להיעלם לחלוטין, תוך מתן מקום לסוכנים אינטלקטואליים שמבצעים ניתוח, מסודרים ומעבדים נתונים באופן עצמאי

AI מערכות אחסון משולבות מאפשרות לעבור לשימוש בסוכני

ככל שהמורכבות של הנתונים והמערכות קטנה יותר, כך נדרש פחות קוד לכתיבה ולתחזוקה. והדרך הפשוטה ביותר לחסוך בפיתוח היא להיפטר לחלוטין מהקוד, ולהחליפו בנתונים. כאשר פיתוח קוד האפליקציה עובר מקוד למודלים של מכיוון שברקע של מושגים אלה עומד אופן (data-driven), נתונים, מתרחש בהכרח שינוי לעבר גישה מונעת נתונים. חשיבה שונה לחלוטין

כאשר אדם בוחר בדרך של עבודה עם נתונים במרכז, הוא מתחיל לראות אחרת את תפקידם. הנתונים מפסיקים להיות "חומר גלם" עבור אפליקציות – עכשיו הם הבסיס שעליו נבנית הארכיטקטורה, הלוגיקה והאינטראקציה

הגישה המסורתית לניהול נתונים מתחילה בדרך כלל ברמת האפליקציות ובבנייה מזכירה מערכת ביורוקרטית מסורבלת: אישורים ברמות רבות, בדיקות ידניות, גרסאות אינסופיות של מסמכים דרך מוצרים תוכנתיים מתאימים. עם התפתחות הטכנולוגיות הדיגיטליות, יותר ויותר חברות יאלצו לעבור לעקרון המינימליזם – לאחסן ולהשתמש רק במה שבאמת נחוץ ויהיה בשימוש

את הלוגיקה של המינימיזציה אימצו הספקים. כדי לפשט את תהליכי האחסון והעיבוד של הנתונים, העבודה של הידועים SaaS המשתמשים מועברת מהפונקציות של אפליקציות וכלים לא מקוונים לשירותים בענן ולפתרונות

המודרניות, המאפשר IT) תוכנה כשירות (הוא אחד הכיוונים המרכזיים בתשתיות SaaS המושג למשתמשים לגשת לאפליקציות דרך האינטרנט מבלי הצורך להתקין ולתחזק תוכנה על המחשבים האישיים שלהם

הקל על ההתרחבות, ניהול הגרסאות והפחית את העלויות של תמיכה ותחזוקה, אך מצד שני, בנוסף לתלות SaaS, מגדל, בלוגיקה של אפליקציה ספציפית, הוא גם הפך את המשתמש לתלוי לחלוטין בתשתית הענן של הספק. אם השירות נכשל הגישה לנתונים ולתהליכי העסקים עשויה להיות חסומה באופן זמני או אפילו לאורך זמן. בנוסף, כל הנתונים של נשמרים על השרתים של הספק, מה שיוצר סיכונים מבחינת אבטחת מידע SaaS המשתמש בעת עבודה עם אפליקציות והתאמה לדרישות רגולטוריות. שינוי בתעריפים או בתנאי השימוש עשוי גם להוביל לעלייה בהוצאות או לצורך במיגרציה דחופה

והגישה הממוקדת בנתונים מעמידה בסימן שאלה את עתיד היישומים בצורתם המסורתית LLM-סוכני ה-AI-התפתחות ה-AI-בעבר יישומים ושירותים היו נחוצים לניהול לוגיקת עסק ולעיבוד נתונים, הרי עם הגעת סוכני ה-SaaS ובביצוע פונקציות אלו עשויות לעבור למערכות אינטליגנטיות הפועלות ישירות עם הנתונים, AI,

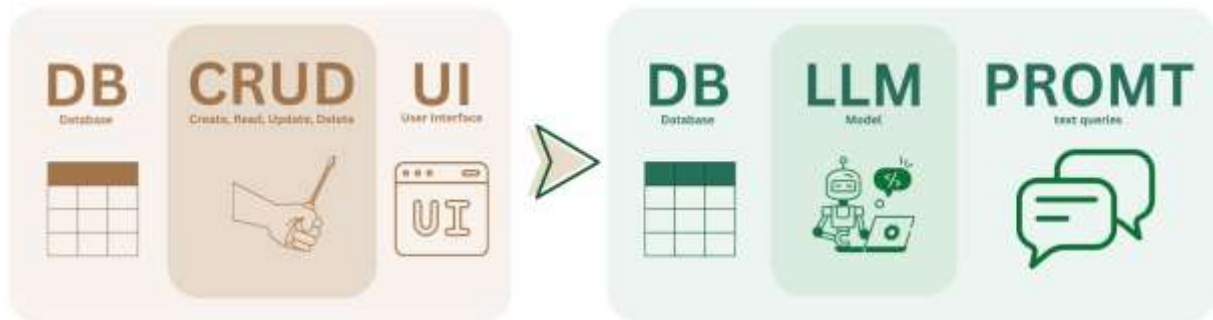
ופתרונות מקומיים AI וברמות הניהול גדונות ארכיטקטורות היברידיות, שבהן סוכני IT לכן, יותר ויותר במחלקות SaaS משלימים את השירותים בענן, ומפחיתים את התלות בפלטפורמות

עשויים להשתנות באופן דרמטי SaaS הגישה שאנו מאמצים מכירה בכך שיישומים עסקיים מסורתיים או יישומי [יצירה, קריאה, עדכון ומחיקה] של בסיסי נתונים עם לוגיקה CRUD בעידן הסוכנים. יישומים אלה, למעשה, מהווים עסקית. אך בעתיד, הלוגיקה הזו תעבור לסוכני בינה מלאכותית

סאטיה נאדלה, מנכ"ל מיקרוסופט, 2024 –

גישה מונחת נתונים ושימוש בסוכני אינטליגנציה מלאכותית/מודלים לשפה גדולים מאפשרים לצמצם את מספר התהליכים המיותרים, ובכך להפחית את העומס על העובדים. כאשר הנתונים מאורגנים כראוי, קל יותר לנתח, להמחיש וליישם אותם לצורך קבלת החלטות. במקום דוחות ובדיקות אינסופיות, המומחים מקבלים גישה למידע עדכני בכמה באופן אוטומטי בצורה של מסמכים מוכנים ולוחות מחוונים LLM קליקים או באמצעות סוכני

(LLM). בעבודה עם נתונים יעזרו לנו כלים של אינטליגנציה מלאכותית (א"מ וצ'אטים של מודלים לשוניים גדולים המסורתיים) יצירה, קריאה, עדכון, מחיקה (לשימוש במודלים CRUD בשנים האחרונות ישנה מגמה של מעבר מפעולות מסוגלים לפרש שפה טבעית וליצור באופן אוטומטי שאילתות מתאימות LLM. לניהול נתונים (LLM) לשוניים גדולים -למסד הנתונים, מה שמפשט את האינטראקציה עם מערכות ניהול הנתונים

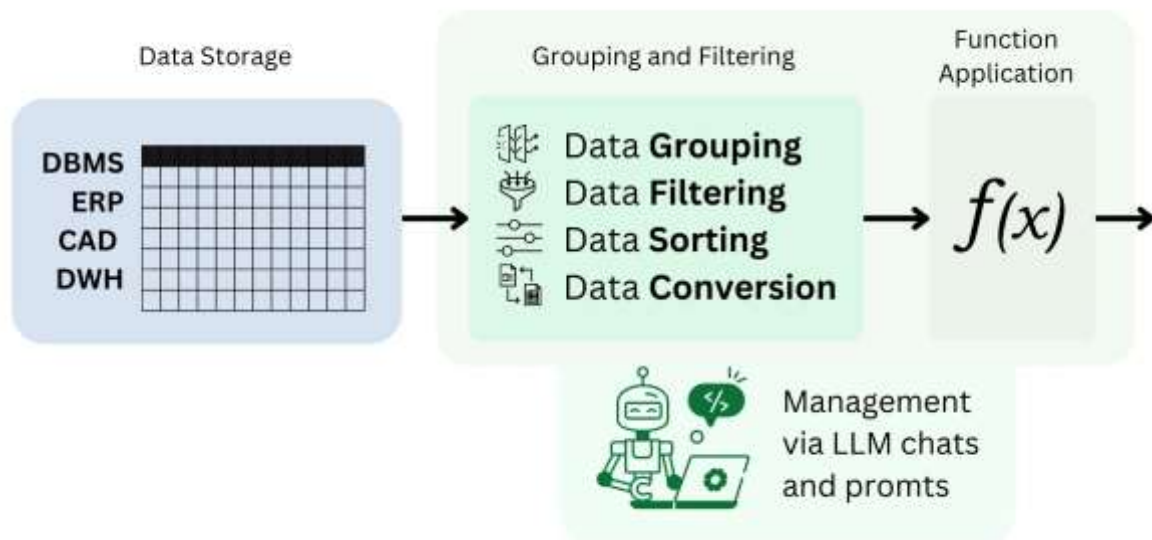


אי.אי. יחליף את ההחלטות בנוגע לניהול מאגרי מידע ובסיסי נתונים ובאינטגרציה שלהם, תוך שהוא דוחק בהדרגה את היישומים CRUD. המסורתיים ואת פעולות

בשלושת עד ששת החודשים הקרובים, הבינה המלאכותית תכתוב 90% מהקוד, ובתוך 12 חודשים כמעט כל הקוד אנתרופיק, מרץ 2025 LLM עשוי להיות מיוצר על ידי בינה מלאכותית. דאריי אמודאי, מנכ"ל חברת

בשנת 2025 המפתחים (למשל) למרות ההתפתחות המהירה של כלים לפיתוח בינה מלאכותית עדיין משחקים תפקיד מרכזי בתהליך זה. סוכני בינה מלאכותית הופכים לעוזרים מועילים יותר ויותר: הם מפרשים (פרטים נוספים על כך בפרקים הבאים) SQL ו-Pandas אוטומטית את הבקשות של המשתמשים, מייצרים שאילתות

או כותבים קוד לניתוח נתונים. כך, הבינה המלאכותית מחליפה בהדרגה את ממשקי המשתמש המסורתיים של היישומים התפשטות מודלים של בינה מלאכותית, כמו מודלים לשוניים, תמריץ את הפיתוח של ארכיטקטורות היברידיות. במקום אנו עשויים לראות אינטגרציה של שירותי ענן עם מערכות ניהול SaaS, לוותר לחלוטין על פתרונות מבוססי ענן ומוצרי מאפשרת להשתמש במודלים רבי עוצמה של (federated learning) נתונים מקומיות. לדוגמה, למידה פדרטיבית בינה מלאכותית מבלי להזיז נתונים רגישים לענן. כך, חברות יוכלו לשמור על שליטה על הנתונים שלהן, תוך כדי גישה לטכנולוגיות מתקדמות.



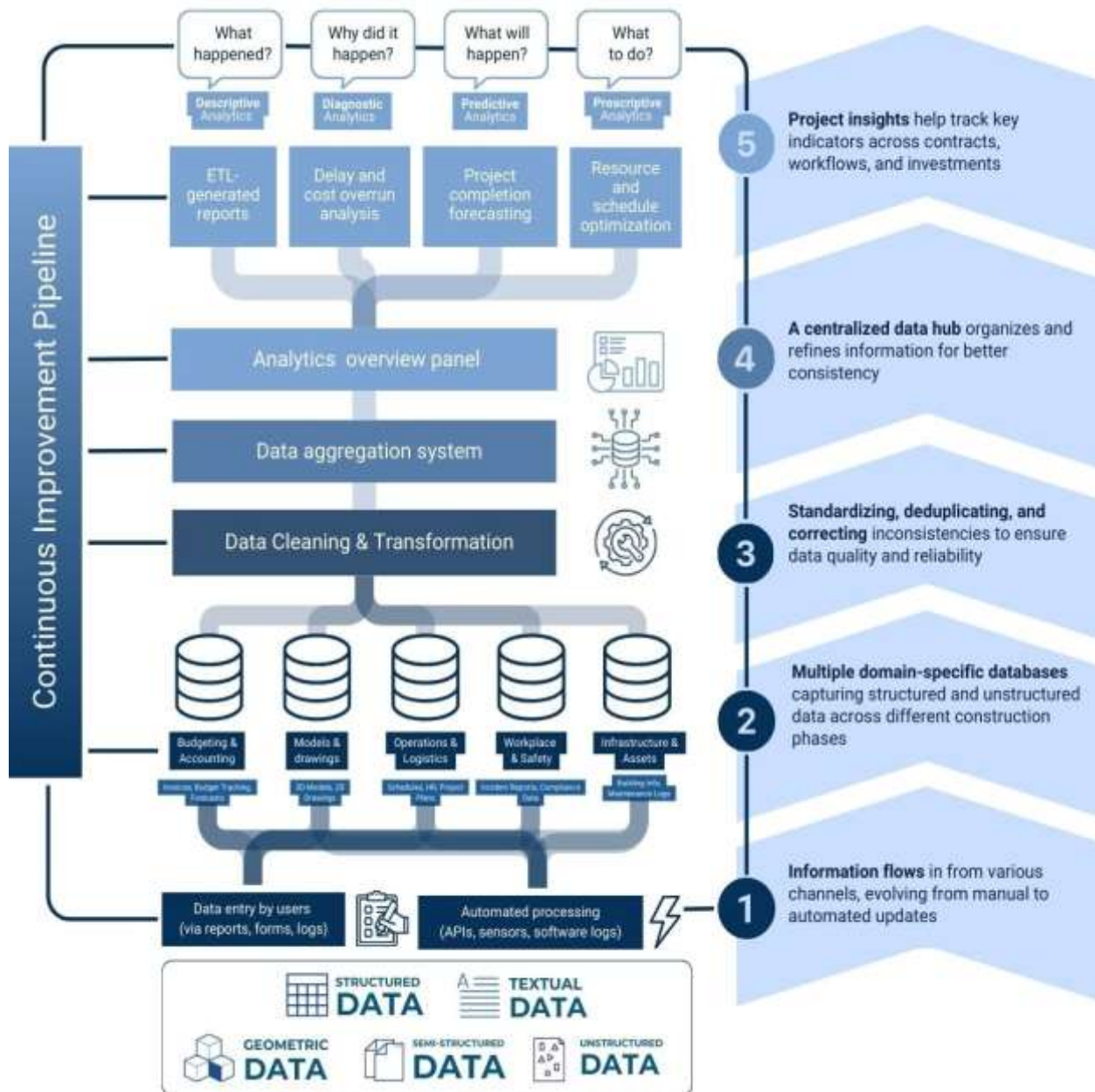
LLM. הפעולות העיקריות של קיבוץ, סינון ומיון עם יישום פונקציות יבוצעו על ידי צ'אט 2.24. ריש

העתיד של תעשיית הבינה יתבסס על שילוב של פתרונות מקומיים, יכולות ענן ומודלים אינטליגנטיים העובדים יחד כדי מאפשר למשתמשים ללא ידע טכני מעמיק לתקשר עם בסיסי נתונים LLM. ליצור מערכות ניהול נתונים יעילות ובטוחות וסוכני בינה מלאכותית וכיצד הם LLM ואחסון נתונים, על ידי ניסוח הבקשות שלהם בשפה טבעית. פרטים נוספים על "ופורמטים מובנים של נתונים LLM פועלים נדון בפרק "סוכני

לא רק יקל על העבודה עם המידע LLM-נתונים מאורגנים כראוי וכלים פשוטים ונוחים לניתוח הנתונים עם תמיכה באלא גם יסייעו במזעור טעויות, יגבירו את היעילות ויאוטומטו תהליכים

מאסוף נתונים לקבלת החלטות: הדרך לאוטומציה

בחלקים הבאים של הספר נבחן בפירוט כיצד המומחים מתקשרים זה עם זה וכיצד הנתונים הופכים לבסיס לקבלת מציגה דוגמה של תרשים המראה את רצף שלבי עיבוד 2.25. החלטות, אוטומציה והגברת היעילות בעבודה. ריש (Continuous Improvement Pipeline) הנתונים בגישה מונחת נתונים. תרשים זה ממחיש את המתווה של שיפור מתמשך של תהליכים חלקים ממנו ייבחנו בפירוט בהמשך הספר, (Improvement Pipeline)



דוגמה של צינור שיפור מתמשך של נתונים: זרם עיבוד וניתוח נתונים בפרויקט בנייה 2.25. ריש.

המערכת המתארת את תהליכי העסקים של חברה בגודל בינוני בנויה על עיקרון רב-שכבתי. היא כוללת: איסוף נתונים ניקוי, אגירה, עיבוד אנליטי וקבלת החלטות על סמך התוצאות שהושגו. כל השלבים הללו נלמד בהמשך הספר - הן בהקשר תיאורטי והן דרך דוגמאות מעשיות

- המידע מגיע הן במצב ידני (דרך דוחות, טפסים, יומנים) והן (3.11. ריש) בשלב הראשון מתבצע הזנת נתונים, הנתונים יכולים להיות במבנים שונים: גיאומטריים (חישנים, מערכות תוכנה, API-מ) בצורה אוטומטית. טקסטואליים, לא מובנים. בשלב זה עולה הצורך בסטנדרטיזציה, סטרוקטורזציה ואחידות של זרמי המידע.
- השלב הבא הוא עיבוד והמרת נתונים. זה כולל תהליכים של ניקוי, הסרת כפילויות, תיקון שגיאות והכנת מידע לניתוח נוסף. שלב זה הוא קריטי, שכן איכות האנליטיקה תלויה ישירות בניקיון ובדיוק הנתונים

■ לאחר מכן הנתונים נכנסים לטבלאות מיוחדות, מסגרות נתונים או בסיסי נתונים, המופרדים לפי תחומים פונקציונליים: תקצוב ורישום, מודלים ושרטוטים, לוגיסטיקה, ביטחון ותשתיות. חלוקה זו מאפשרת לארגן גישה נוחה ומספקת אפשרות לניתוח חוצה תחומים של המידע

■ לאחר מכן הנתונים מאוגדים ומוצגים בלוח בקרה אנליטי. כאן מיושמות שיטות של אנליטיקה תיאורית אבחנתית, חיזוי והמלצה. זה מאפשר לענות על שאלות מפתח: מה קרה, מדוע זה קרה, מה יקרה בעתיד ואילו - פעולות יש לנקוט. לדוגמה, המערכת יכולה לחשוף עיכובים, לחזות סיום פרויקטים או לייעל משאבים לבסוף, בשלב האחרון נוצרים מסקנות אנליטיות ומדדים מרכזיים, המסייעים במעקב אחר ביצועי חוזים, ניהול השקעות ושיפור תהליכי עסקים. מידע זה הופך לבסיס לקבלת החלטות ואסטרטגיית פיתוח החברה.

באופן דומה, הנתונים עוברים מסלול מאיסוף ועד שימוש בניהול אסטרטגי. בחלקים הבאים של הספר נבחן כל שלב בפירוט, תוך מתן תשומת לב מיוחדת לסוגי נתונים, שיטות עיבוד נתונים, כלים אנליטיים ומקרים אמיתיים של שימוש בגישות אלו בתעשיית הבנייה

צעדים נוספים: הפיכת הכאוס למערכת ניהולית

בחלק זה חקרנו את בעיות הסילו המידע ובחנו את השפעת המורכבות המוגזמת של מערכות על היעילות העסקית, ניתחנו את המעבר מהמהפכה התעשייתית הרביעית לחמישית, שבה התפקיד המרכזי הוא לנתונים ולא ליישומים. ראינו כיצד מפחית את הפרודוקטיביות IT-מערכות מידע מפוזרות יוצרות מחסומים להחלפת ידע, וכיצד ההמשך של סיבוך הנוף הומעכב חדשנות בתעשיית הבנייה

לסיכום חלק זה, כדאי להדגיש את הצעדים המעשיים העיקריים שיכולים לעזור ליישם את הגישות שנדונו במשימות היומיומיות שלך

■ ויזואליזציה של הנוף המידע שלך

☐ עימם אתה עובד באופן קבוע (Miro, Figma, Canva), צור מפה ויזואלית של מקורות הנתונים

☐ הוסף למפה זו את המערכות והיישומים שבהם אתה משתמש בעבודתך

☐ גלה פונקציות כפולות פוטנציאליות ופתרונות מיותרים

☐ זיהוי נקודות קריטיות, שבהן עשויה להתרחש אובדן או עיוות נתונים בעת העברה בין מערכות

■ יישם פרקטיקות ניהול נתונים אישיות

☐ העבר את המוקד מהיישומים לנתונים כנכס מרכזי בתהליכים

☐ תעד את מקורות הנתונים ואת המתודולוגיה לעיבוד כדי להבטיח שקיפות

☐ פתח מנגנונים להערכה ושיפור איכות הנתונים

☐ שאף לכך שהנתונים יוזנו פעם אחת, וישמשו מספר פעמים - זו הבסיס לארגון תהליכים יעיל

■ בצוות שלך (data-driven) קדם גישה מונעת נתונים

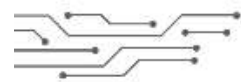
☐ הצע להשתמש בפורמטים סטנדרטיים ואחידים להחלפת נתונים בין עמיתים

☐ העלו באופן קבוע שאלות הקשורות לאיכות וזמינות הנתונים בפגישות הצוות

□ הכירו את האלטרנטיבות בקוד פתוח לכלים שאתם משתמשים בהם לפתרון בעיותיכם

התחילו בקטן - בחרו תהליך או קבוצת נתונים ספציפיים שהם קריטיים לעבודה שלכם, והחילו עליהם גישה ממוקדת נתונים, תוך העברת הדגש מהכלים לנתונים. לאחר שתשיגו הצלחה במשימה פיילוט אחת, תקבלו לא רק ניסיון מעשי, אלא גם הדגמה מוחשית של היתרונות של המתודולוגיה החדשה עבור הצוות שלכם. במהלך ביצוע רוב הצעדים הללו מודרני LLM אם יתעוררו שאלות, תוכלו לפנות להבהרות ועזרה מכל מודל.

בחלקים הבאים של הספר נעבור לבחינה מעמיקה יותר של שיטות לארגון ואחידות נתונים ונחקור גישות מעשיות לשילוב מידע מגוון. תשומת לב מיוחדת תינתן למעבר מאחסונים מבודדים לאקוסיסטמות נתונים מאוחדות, המשחקות תפקיד מרכזי בהמרה הדיגיטלית של תעשיית הבנייה.





III חלק מסגרת הנתונים בתהליכי עסקי הבנייה

בחלק השלישי נבנה תמונה מקיפה של טיפולוגיית הנתונים בבנייה ושיטות לארגוןם היעיל. ניתוחים ייערכו על מאפיינים וספציפיות העבודה עם נתונים מובנים, לא מובנים, חצי מובנים, טקסטואליים וגיאומטריים בהקשר של פרויקטים בנייתיים. ייבחנו פורמטים מודרניים לאחסון ופרוטוקולי החלפת מידע בין מערכות שונות בשימוש בתעשייה. יוסברו כלים מעשיים ושיטות להמיר נתונים בפורמטים יומלצו גישות להבטחת CAD (BIM) שונים לסביבה מאורגנת אחת, כולל דרכים לשילוב נתוני איכות הנתונים דרך סטנדרטיזציה ואימות, קריטריות לדיוקנות החישובים הבנייתיים. ינותחו היבטים עם דוגמאות קוד (LLM מודלים, Python Pandas) מעשיים של שימוש בטכנולוגיות מודרניות (CoE) לפתרון בעיות טיפוסיות בתעשיית הבנייה. תוצג הערכת ערך של הקמת מרכז מצוינות כמבנה ארגוני לתיאום וסטנדרטיזציה של גישות לניהול מידע.

פרק 3.1

סוגי נתונים בבנייה

סוגי הנתונים החשובים ביותר בתעשיית הבנייה

בתעשיית הבנייה המודרנית, מערכות, יישומים ואחסוני נתונים של חברות מתמלאים באופן פעיל במידע ובנתונים מסוגים ובפורמטים שונים. נבחן ביתר פירוט את סוגי הנתונים העיקריים המעצבים את הנוף המידע של חברה מודרנית הפועלת בתעשיית הבנייה

■ ומסדי נתונים Excel נתונים מובנים: נתונים אלו בעלי מבנה ארגוני ברור, לדוגמה, גיליונות אלקטרוניים של רלציוניים.

■ נתונים לא מובנים: מידע שאינו מאורגן בהתאם לכללים נוקשים. דוגמאות לנתונים כאלה כוללות טקסט וידאו, תמונות והקלטות אודיו.

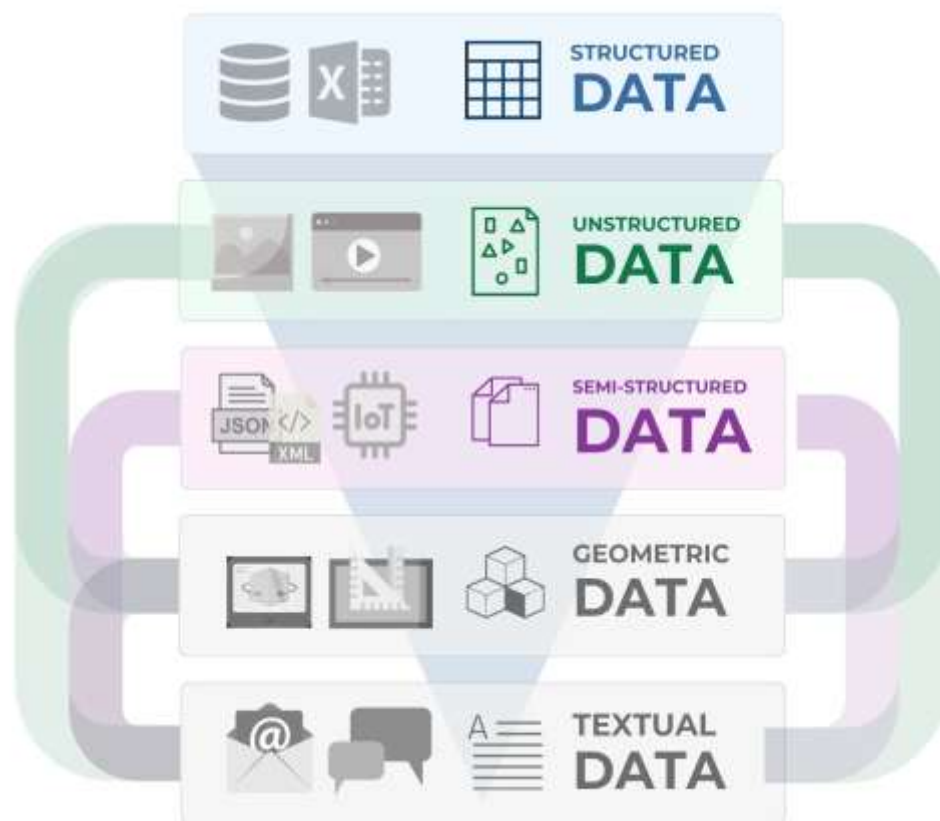
■ נתונים חצי מובנים: נתונים אלו תופסים מקום ביניים בין נתונים מובנים לנתונים לא מובנים. הם מכילים אלמנטים של מבנה, אך המבנה הזה לא תמיד ברור או מתואר לעיתים קרובות דרך סכמות שונות. דוגמאות לנתונים חצי מובנים בבנייה כוללות: מפרטים טכניים, מסמכי פרויקט או דוחות על עבודה שבוצעה.

■ נתונים טקסטואליים: כוללים את כל מה שהתקבל כתוצאה מהתקשורת בעל פה ובכתב, לדוגמה, מיילים. תמלילים של ישיבות ופגישות.

■ שבהן מומחים יוצרים נתונים גיאומטריים של רכיבי CAD נתונים גיאומטריים: נתונים אלו מגיעים מתוכנות הפרויקט לצורך ויזואליזציה, אישור ערכי נפחים או בדיקת התנגשויות.

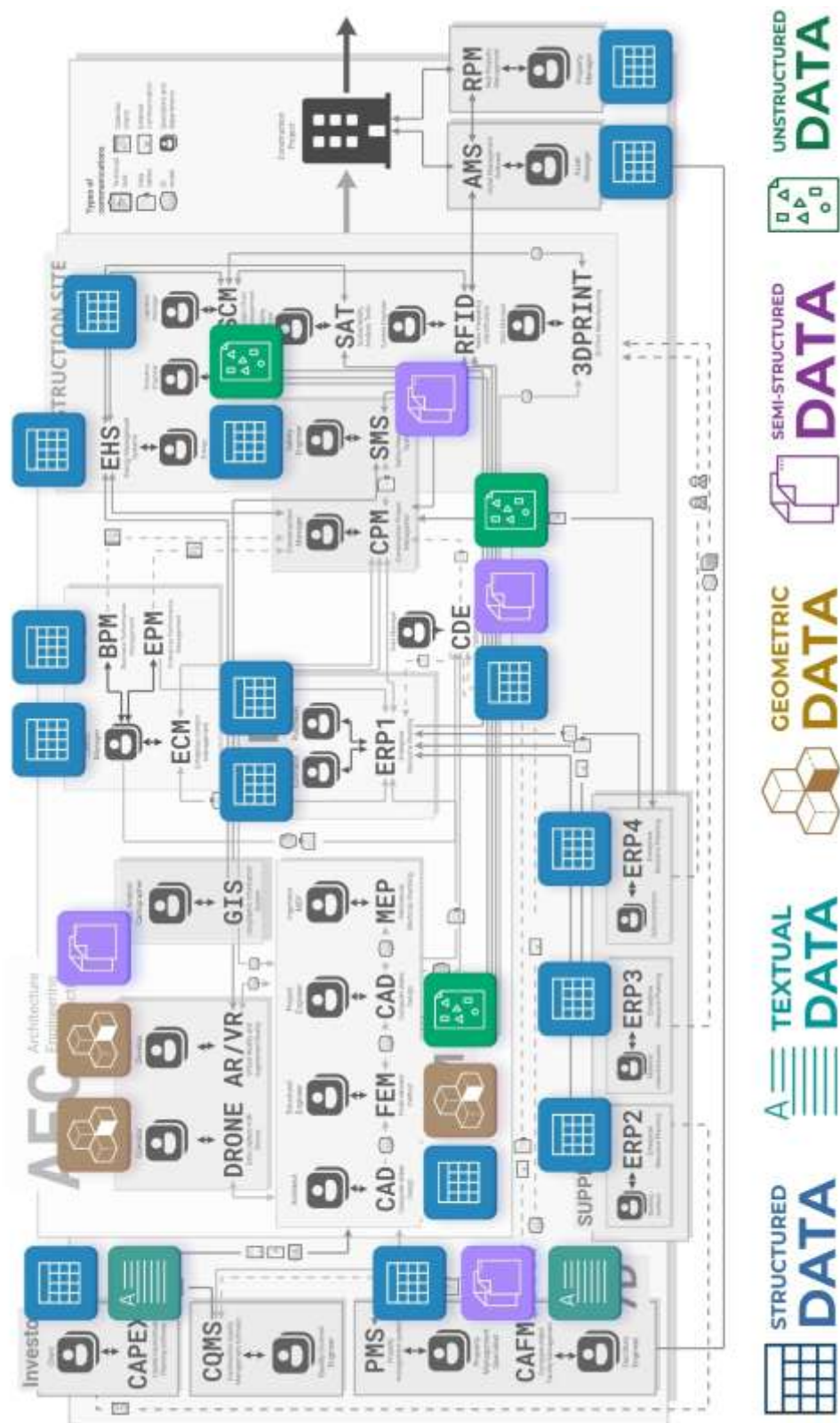
חשוב לציין כי נתונים גיאומטריים וטקסטואליים (אלפאנומריים) אינם קטגוריה נפרדת, אלא יכולים להופיע בכל שלושת סוגי הנתונים. נתונים גיאומטריים, לדוגמה, יכולים להיות חלק מנתונים מובנים (פורמטים פרמטריים של וגם מנתונים לא מובנים) שרטוטים סרוקים. (נתונים טקסטואליים יכולים גם הם להיות מאורגנים בבסיסי CAD נתונים) נתונים מובנים (או להתקיים במסמכים ללא מבנה ברור.

כל סוג נתונים בחברת בנייה הוא אלמנט ייחודי בפאזל הנכסים המידעיים של החברה. מנתונים לא מובנים, כמו תמונות מאתרי בנייה והקלטות אודיו של פגישות, ועד לרשומות מובנות, כולל טבלאות ובסיסי נתונים - כל אלמנט משחק תפקיד חשוב בעיצוב הנוף המידע של החברה.



מהנדסים ומנהלי נתונים צריכים ללמוד לעבוד עם כל סוגי הנתונים המשמשים בתעשיית הבנייה 3.11. ריש.

- המשמשות בבנייה, (3.12. ריש) הנה דוגמה לרשימה של כמה מערכות וסוגי נתונים הקשורים אליהן
 - (תכנון משאבים ארגוניים (- מעבד בדרך כלל נתונים מובנים, מסייע בניהול משאבי הארגון ומשלב ERP תהליכים עסקיים שונים
 - (מודל מידע בנייתי (- משתמש בנתונים גיאומטריים ונתונים BIM) עיצוב בעזרת מחשב (בשילוב עם CAD חצי-מובנים לתכנון ומידול פרויקטי בנייה, ומבטיח דיוק ועקביות של מידע בשלב התכנון
 - (מערכות מידע גיאוגרפיות (- פועלת עם נתונים גיאומטריים ומובנים ליצירה וניתוח של נתוני מפה GIS ויחסים מרחביים
 - (זיהוי באמצעות תדר רדיו (- משתמש בנתונים חצי-מובנים למעקב יעיל אחר חומרים וציוד באתר RFID הבנייה באמצעות זיהוי בתדר רדיו
 - (ניהול תוכן הנדסי (- זו מערכת לניהול נתוני הנדסה ודוקומנטציה, כולל נתונים חצי-מובנים ולא ECM מובנים, כמו שרטוטים טכניים ודוקומנטציה פרויקטית



פורמטים שונים ונתונים ממלאים מערכות שונות, ודורשים תרגום לצורה המתאימה לשילוב כולל 3.12. ריש.

מערכות אלו ורבות אחרות בחברה מנהלות מגוון רחב של נתונים, מנתונים טבלאיים מובנים ועד מודלים גיאומטריים מורכבים, ומבטיחות אינטראקציה משולבת בתהליכי תכנון, תכנון וניהול הבנייה.

- בין אנשי מקצוע בפרויקט הבנייה מתבצע חילופי נתונים שונים (3.13 ריש) בדוגמת דיאלוג מפושט

אדריכל: בהתחשב ברצונות הלקוח, הוספתי אזור מנוחה על הגג. אנא עיינו בעיצוב החדש ("נתונים גיאומטריים מודל -").

"מהנדס קונסטרוקציה": הפרויקט התקבל. אני מחשב את יכולת הנשיאה של הגג עבור אזור המנוחה החדש (נתונים מובנים וחצי-מובנים - טבלאות חישוב).

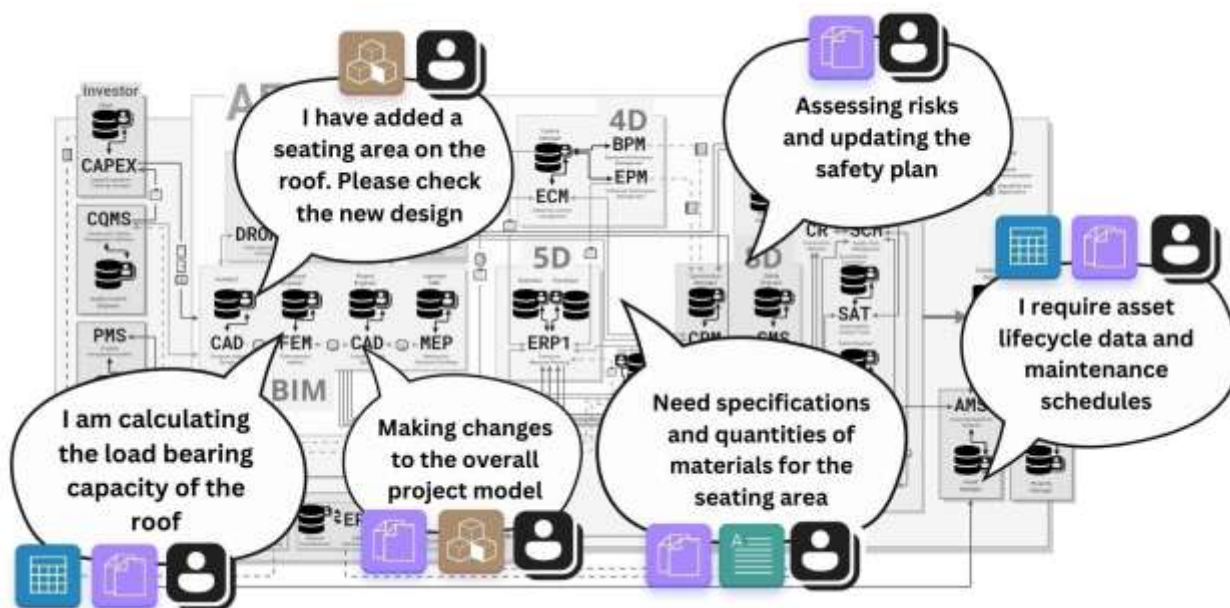
מנהל רכש: נדרשות מפרטים וכמות חומרים עבור אזור המנוחה, כדי לארגן את הרכישה ("נתונים טקסטואליים וחצי-מובנים - רשימות ומפרטים").

מהנדס בטיחות בעבודה: קיבלתי נתונים על אזור חדש. אני מעריך את הסיכונים ומעדכן את תוכנית הבטיחות (נתונים חצי-מובנים - מסמכים ותוכניות).

אני מבצע שינויים במודל הכללי של הפרויקט כדי לתקן את המסמכים העבודה (נתונים BIM: מומחה לדגם גיאומטריים ונתונים חצי-מובנים).

מנהל פרויקט: אני כולל את אזור המנוחה החדש בלוח הזמנים של העבודה. אני מעדכן את לוחות הזמנים והמשאבים במערכת ניהול הפרויקטים (נתונים מובנים וחצי-מובנים - לוחות זמנים ותוכניות).

אני מכין נתונים עבור תחזוקה עתידית של אזור המנוחה ומזין אותם במערכת (FM) מומחה לתחזוקת נכסים ניהול הנכסים (נתונים מובנים וחצי-מובנים - הוראות ותוכניות תחזוקה טכנית).



התקשורת בין המומחים מתבצעת הן ברמת הטקסט והן ברמת הנתונים 3.13 ריש.

כל מומחה עובד עם סוגים שונים של נתונים, מה שמבטיח אינטראקציה יעילה של הצוות והצלחה בביצוע הפרויקט. הבנת ההבדלים בין נתונים מובנים, חצי-מובנים ולא מובנים מאפשרת להבין את התפקיד הייחודי של כל סוג בתהליכי עסקים דיגיטליים. חשוב לא רק לדעת שיש צורות שונות של נתונים, אלא גם להבין כיצד, היכן ולמה הם מיועדים.

עוד לפני זמן לא רב, הרעיון של שילוב נתונים מסוגים שונים נראה שאפתני, אך קשה ליישום. היום - זה כבר חלק מהפרקטיקה היומיומית. אינטגרציית נתונים עם מבנים שונים הפכה לחלק בלתי נפרד מהארכיטקטורה המודרנית של מערכות מידע.

בפרקים הבאים נבחן בפירוט את הסטנדרטים והגישות המרכזיים המאפשרים לשלב נתונים מובנים, חצי-מובנים ולא מובנים לייצוג אחיד ומסודר. תשומת לב מיוחדת תינתן לנתונים מובנים ולמאגרי נתונים רלציוניים - כמנגנונים העיקריים לאחסון, עיבוד וניתוח מידע בתעשיית הבנייה.

נתונים מובנים

בתעשיית הבנייה, מידע מגיע ממקורות רבים - שרטוטים, מפרטים, לוחות זמנים ודוחות. כדי לנהל את הזרם הזה ביעילות יש צורך במבנה שלו. נתונים מובנים מאפשרים לארגן את המידע בצורה נוחה, קריאה ונגישה.

מהמומחים בניהול [17], 67% JB Knowledge לפי הדו"ח השנתי החמישי על טכנולוגיות בנייה של פרויקטים בבנייה עוקבים ומעריכים את ביצועי העבודה ידנית או באמצעות גיליונות אלקטרוניים.

הם בשימוש נרחב לאחסון, עיבוד וניתוח מידע. CSV ו-XLSX אחד מהפורמטים הנפוצים ביותר של נתונים מובנים הוא בגיליונות אלקטרוניים. בגיליונות אלה הנתונים מוצגים בשורות ובטורים, מה שהופך אותם לנוחים לקריאה, עריכה וניתוח.

ZIP ומאוחסן באמצעות אלגוריתם XML שנוצר על ידי מיקרוסופט, מבוסס על שימוש במבני XLSX פורמט תכונות עיקריות של הפורמט

■ תמיכה בנוסחאות מורכבות, דיאגרמות ומאקרו.

■ אפשרות לאחסן נתונים בלוחות שונים, כמו גם לעצב את המידע.

■ אך תואם גם עם חבילות משרדיות אחרות, Microsoft Excel אופטימלי לעבודה בסביבת

הוא קובץ טקסט פשוט שבו הערכים מופרדים בפסיקים, נקודותיים או סמלים מפרידים אחרים CSV פורמט היתרונות העיקריים

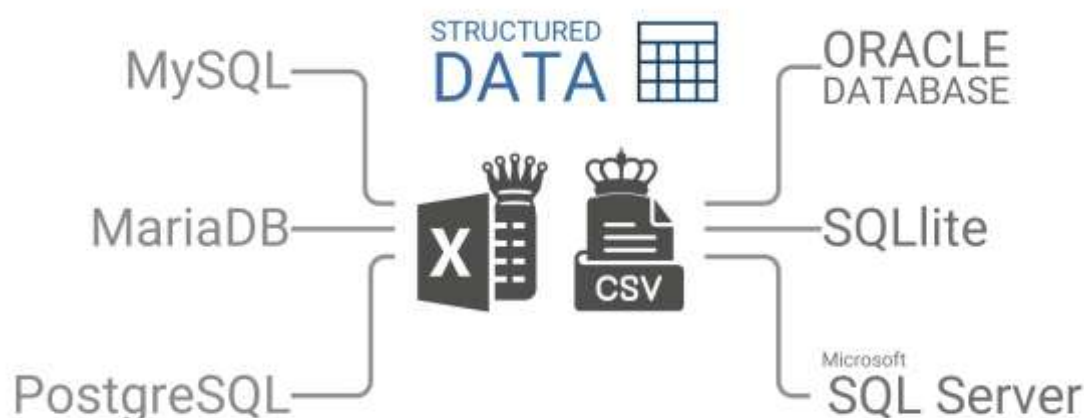
■ תאימות אוניברסלית עם תוכניות ומערכות הפעלה שונות.

■ נוחות בייבוא/ייצוא לבסיסי נתונים ולמערכות ניתוח.

■ קלות עיבוד גם בעורכי טקסט.

אינו תומך בנוסחאות ובפורמט, ולכן השימוש העיקרי שלו הוא בהחלפת נתונים בין מערכות ועדכון CSV, עם זאת הפך לכלי פופולרי להעברת נתונים בסביבות CSV, המידע בהמוניו. בזכות האוניברסליות והעצמאות שלו מפלטפורמות הטרונגניות IT.

משמשים כקשר בין מערכות שונות העובדות עם נתונים מובנים. הם במיוחד מועילים CSV ו-XLSX שני הפורמטים - במשימות שבהן קריאות, עריכה ידנית ותאימות בסיסית הם חשובים.



הם הקשר בין מערכות שונות העובדות עם נתונים מובנים **CSV** ו-**XLSX** פורמטים.

הטרוגניות IT לפורמט הפופולרי ביותר להעברת נתונים בסביבות **CSV** עצמאות מפלטפורמה עושה את ובמערכות.

אינם מיועדים לחישובים בעלי ביצועים גבוהים או לאחסון ארוך טווח של כמויות גדולות של **CSV** ו-**XLSX**, עם זאת פורמטים אלה נבחרו בפירוט בפרק "אחסון נתונים גדולים: ניתוח פורמטים פופולריים ויעילותם" בחלק התשיעי. **HDF5**, **Apache Parquet**, **Apache ORC**, **Feather**, **Apache Parquet**, **Apache ORC**, **Feather**, **HDF5** של הספר.

משמש לרוב למשימות קטנות ואוטומציה של תהליכים שגרתיים. תרחישים מורכבים **XLSX** עם פורמט **Excel**, למעשה ועוד. במערכות אלה מאוחסנים **ERP**, **PMIS**, **CAFM**, **CPM**, **SCM** יותר דורשים שימוש במערכות ניהול נתונים, כמו -נתונים מובנים שעליהם מתבססת הארגון וניהול זרמי המידע של החברה.

מערכות מידע מודרניות לניהול נתונים, המיועדות לתעשיית הבנייה, מתבססות על נתונים מובנים, מאורגנים בטבלאות לניהול אמין, נרחב ושלם של כמויות גדולות של מידע, מפתחים פונים למערכות ניהול בסיסי נתונים רלציוניות (**RDBMS**).

SQL ושפת השאילתות **RDBMS** מסדי נתונים רלציוניים

אלו הן מערכות - (**RDBMS**) לצורך אחסון, עיבוד וניתוח נתונים בצורה יעילה, נעשה שימוש בבסיסי נתונים רלציוניים. אחסון נתונים המארגנות מידע בטבלאות עם קשרים מוגדרים ביניהן.

אינם רק מידע דיגיטלי; הם מהווים את הבסיס לעסקאות (**RDBMS**) נתונים המיועדים לבסיסי נתונים ואינטראקציה בין מערכות שונות.

:- (3.15. ריש) הנה כמה מהמערכות הניהוליות הרלציוניות הנפוצות ביותר

MySQL (Linux, Apache, MySQL, PHP/Perl/Python) LAMP הפופולריות ביותר, חלק מהסטאק (RDBMS)-קוד פתוח (- אחת ממערכות ה MySQL בשימוש נרחב בפיתוח אתרים בזכות הפשטות והביצועים. הגבוהים שלה.

(קוד פתוח -) מערכת רלציונית אובייקטיבית חזקה, ידועה באמינותה וביכולות המתקדמות PostgreSQL שלה. מתאימה לפתרונות ארגוניים מורכבים.

מערכת מסחרית של מיקרוסופט, בשימוש נרחב בסביבות ארגוניות בזכות Microsoft SQL Server – האינטגרציה עם מוצרים אחרים של החברה ורמת אבטחה גבוהה.

אחת ממערכות ניהול בסיסי הנתונים החזקות והאמינות ביותר, בשימוש בעסקים – Oracle Database גדולים וביישומים קריטיים.

מיועדת לתאגידים גדולים, מספקת ביצועים גבוהים ועמידות בפני תקלות – IBM DB2.

(קוד פתוח -) מערכת בסיסי נתונים קלה, מתאימה במיוחד ליישומים ניידים ולמערכות עצמאיות SQLite CAD (BIM) כמו תוכנות לתכנון.

MySQL, PostgreSQL, Microsoft SQL Server, Oracle® Database, IBM® DB2 ו-SQLite - מערכות ניהול בסיסי נתונים פופולריות בעסקי הבנייה פועלות עם נתונים מובנים. כל המערכות הללו מציעות פתרונות - SQLite ו-IBM® DB2, Oracle® Database, הזקים וגמישים לניהול מגוון רחב של תהליכי עסקים ויישומים, מאתרי אינטרנט קטנים ועד מערכות ארגוניות רחבות (3.21. ריש) היקף.

היוו כ-72% (RDBMS) בשנת 2022 מערכות ניהול בסיסי נתונים רלציוניות, [48] Statista לפי נתוני מכלל מערכות ניהול בסיסי הנתונים בשימוש.

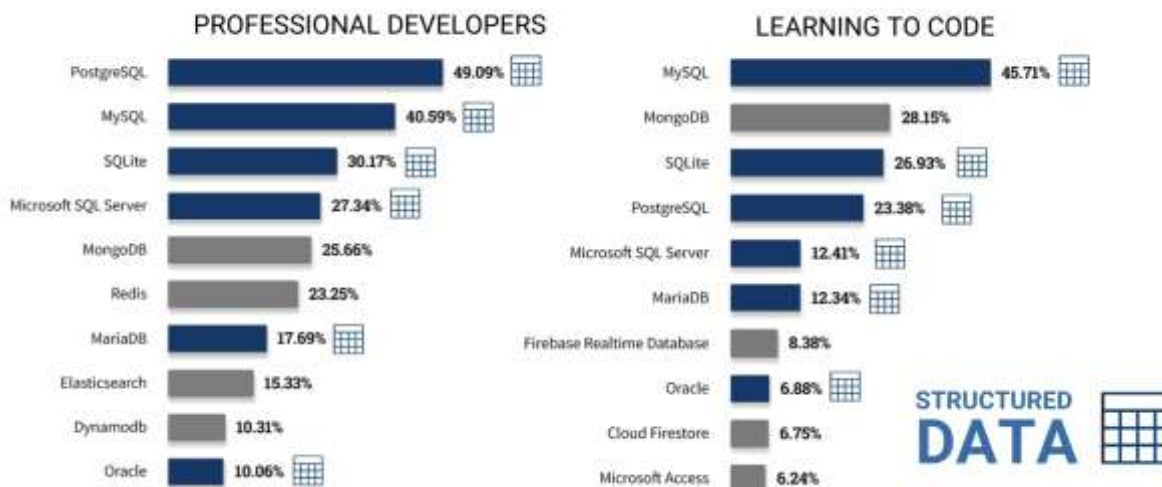
	Rank			DBMS	Database Model	Open Source vs Commercial
	Mar2025	Feb2025	Mar2024			
	1.	1.	1.	Oracle®	Relational, Multi-model	Commercial
	2.	2.	2.	MySQL	Relational, Multi-model	Open Source
	3.	3.	3.	Microsoft® SQL Server	Relational, Multi-model	Commercial
	4.	4.	4.	PostgreSQL	Relational, Multi-model	Open Source
	5.	5.	5.	MongoDB	Document, Multi-model	Open Source
	6.	7.	9.	Snowflake®	Relational	Commercial
	7.	6.	6.	Redis®	Key-value, Multi-model	Open Source
	8.	8.	7.	Elasticsearch®	Multi-model	Open Source
	9.	9.	8.	IBM Db2	Relational, Multi-model	Commercial
	10.	10.	10.	SQLite	Relational	Open Source
	11.	11.	12.	Apache Cassandra®	Multi-model	Open Source
	12.	12.	11.	Microsoft Access®	Relational	Open Source
	13.	13.	17.	Databricks®	Multi-model	Commercial
	14.	14.	13.	MariaDB	Relational, Multi-model	Open Source
	15.	15.	14.	Splunk	Search engine	Commercial
	16.	16.	16.	Amazon DynamoDB	Multi-model	Commercial
	17.	17.	15.	Microsoft Azure SQL	Relational, Multi-model	Commercial

פופולריות השימוש בסיסי נתונים מובנים) מסומן בכחול (בדירוג מערכות ניהול בסיסי נתונים) על פי חומרי [49] (3.15). ריש.

התקנת בסיסי נתונים עם קוד פתוח היא פשוטה למדי – אפילו ללא ידע טכני מעמיק. מערכות קוד פתוח כמו PostgreSQL, MySQL או SQLite זמינות בחינם ופועלות על רוב מערכות ההפעלה Windows, macOS ולינוקס. כל מה שצריך זה להיכנס לאתר הרשמי של הפרויקט, להוריד את ההתקנה ולפעול לפי ההוראות. ברוב המקרים, ההתקנה לא לוקחת יותר מ-10–15 דקות. אחת ממערכות אלו ניצור ונדגם בחלק הרביעי של הספר (4.38). ריש.

ניתן להקים בסיס נתונים בכמה קליקים – הפלטפורמה תציע לכם תבניות מוכנות להתקנה. בזכות פתיחות, Amazon Web Services, Google Cloud או Microsoft Azure (כגון) הקוד, בסיסי נתונים אלו קלים להתאמה לצרכים שלכם, וקהל המשתמשים הרחב תמיד יוכל לעזור במציאת פתרון לכל בעיה.

המאפשרות לחברות, (3.16). ריש) נשארות הבסיס למגוון רחב של יישומי עסקים ופלטפורמות אנליטיות RDBMS. לאחסן, לעבד ולנתח נתונים ביעילות – כלומר, לקבל החלטות מבוססות ובזמן.



על אילו בסיסי נתונים הם השתמשו בשנה שעברה (הגדול ביותר IT הפורום) **StackOverflow**-סקר מפתחים ב 3.16. ריש [50] על פי חומרי (מודגשות בכחול **RDBMS**) ואילו הם רוצים להשתמש בשנה הבאה.

SQL – מספקות אמינות, עקביות נתונים, תמיכה בעסקאות ומשתמשות בשפת שאילתות עוצמתית **RDBMS** (Structured Query Language) הנמצאת בשימוש נרחב באנליטיקה ומאפשרת לקבל, לשנות ולנתח בקלות מידע, היא הכלי המרכזי לעבודה עם נתונים במערכות רלציוניות **SQL**. המאוחסן בבסיסי נתונים

במסדי נתונים ומגמות חדשות **SQL** שאילתות

הנמצאת בשימוש בבסיסי נתונים רלציוניים, בהשוואה לשיטות ניהול מידע אחרות (כגון **SQL**, היתרון המרכזי של שפת הוא התמיכה בנפחים גדולים מאוד של בסיסי נתונים עם מהירות (Excel באמצעות גיליונות אלקטרוניים קלאסיים כמו גבוהה בעיבוד שאילתות

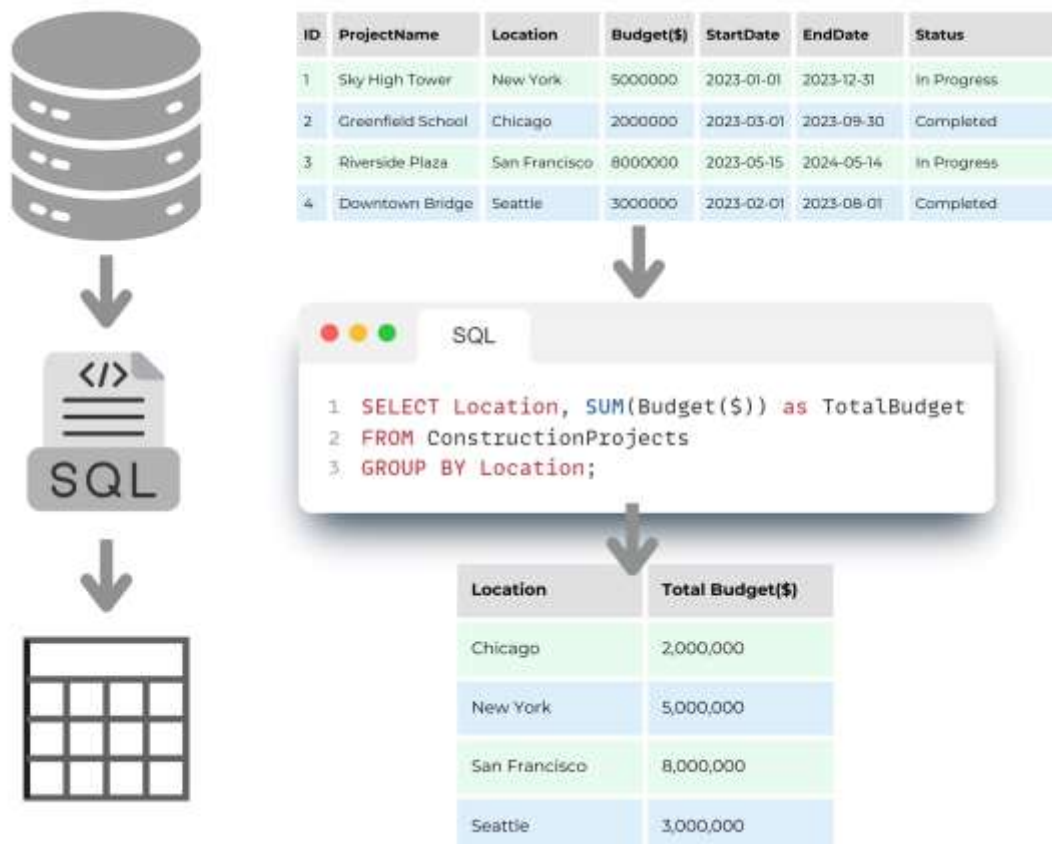
היא שפת תכנות מיוחדת, המיועדת לאחסון, עיבוד וניתוח מידע בבסיסי (**SQL**) שפת שאילתות מובנות משמשת ליצירה, ניהול וגישה לנתונים, ומאפשרת למצוא, לסנן, לאחד ולאגד **SQL**. נתונים רלציוניים מידע בצורה יעילה. היא משמשת ככלי מרכזי לגישה לנתונים, ומספקת דרך נוחה ומסודרת לתקשורת עם מאגרי מידע

Oracle, IBM DB2, עברה דרך מוצרים וחברות משמעותיות כמו **SQL-SEQUEL** האבולוציה של מערכות **SQL** ו-MariaDB ו-SQLite ומסתיימת בהופעת **MySQL** ו-PostgreSQL, SAP, Microsoft SQL Server. **SQL** מה שהופך את העבודה עם נתונים ליותר ניתנת להרחבה, בטוחה, Excel-מספקת אפשרויות עבודה עם טבלאות שאין ב ונוחה לאוטומציה

ניתן ליצור, לשנות ולמחוק טבלאות בבסיס הנתונים, לקבוע **SQL**-ב: (DDL) יצירה וניהול מבני נתונים לעומת זאת, העבודה מתבצעת עם גיליונות ותאים, Excel-קשרים ביניהן ולהגדיר מבני אחסון נתונים. ב. קבועים, ללא קשרים ברורים בין גיליונות וסטי נתונים

מאפשרת להוסיף, לשנות, למחוק ולהוציא נתונים בקצב גבוה, תוך ביצוע **SQL**: (DML) פעולות עם נתונים

עיבוד כמויות גדולות של מידע דורש פעולות, Excel-שאלות מורכבות עם סינון, מיון ואיחוד טבלאות. ב. ידניות או מאקרו מיוחדים, מה שמאט את התהליך ומגביר את הסיכון לטעויות. מאפשרת להפריד זכויות גישה לנתונים עבור משתמשים שונים, ומגבילה את SQL: (DCL) בקרת גישה לעומת זאת, הגישה היא או כללית) בעת העברת הקובץ (או, Excel-האפשרות לערוך או לצפות במידע. ב. דורשת הגדרות מורכבות עם חלוקת זכויות דרך שירותים בענן.



עיבוד מהיר, קיבוץ ואיגוד בעזרת מספר שורות קוד לעיבוד אוטומטי של נתונים: SQL-ב DML-דוגמה ל

מקלה על העבודה עם נתונים בזכות המבנה הוויזואלי והאינטואיטיבי שלה. עם זאת, ככל שכמות הנתונים גדלה Excel, יש גם מגבלות על כמות הנתונים שניתן לאחסן – מקסימום מיליון שורות Excel-יורדים. ל Excel הביצועים של נראה מועדף עבור ויזואליזציה ומניפולציות עם Excel-והביצועים יורדים הרבה לפני הגעה לגבול הזה. לכן, למרות ש מתאימה יותר SQL, כמויות קטנות של נתונים, עבור עבודה עם כמויות גדולות של נתונים

המהווים (Columnar Databases), השלב הבא בהתפתחות נתונים מובנים היה הופעת בסיסי נתונים עמודתיים. חלופה לבסיסי נתונים רלציוניים מסורתיים, במיוחד כאשר מדובר בכמויות גדולות מאוד של נתונים וחישובים אנליטיים. בניגוד לבסיסי נתונים בשורות, שבהם הנתונים מאוחסנים בשורות, בבסיסי נתונים עמודתיים המידע נרשם בעמודות: בהשוואה לבסיסי נתונים קלאסיים, זה מאפשר

צמצום נפח האחסון בזכות דחיסה יעילה של נתונים דומים בעמודות. האצת שאלות אנליטיות, מכיוון שנקראים רק העמודות הנדרשות, ולא כל הטבלה.

■ **Data Lakehouse** ומאגרים של נתונים, לדוגמה, בארכיטקטורת **Big Data** אופטימיזציה של העבודה עם

Big Data כמו גם על יצירת מאגרי **HDF5**, **Apache Parquet**, **Pandas DataFrame**, על בסיסי נתונים עמודתיים פורמט: **DataFrame** – על בסיסם למטרות ניתוח ועיבוד נתונים נדבר בפירוט בפרקים הבאים של ספר זה **Big Data** וארכיטקטורת **DWH** מאגרי **Apache Parquet** אוניברסלי לנתונים טבלאיים "ו" פורמטי אחסון נתונים ועבודה עם **Data Lakehouse**".

נתונים לא מובנים

למרות שברוב המקרים הנתונים המשמשים באפליקציות ובמערכות מידע הם בצורה מובנית, רוב המידע המיוצר בתחום הבנייה מוצג בצורה של נתונים לא מובנים - תמונות, סרטונים, מסמכים טקסטואליים, הקלטות אודיו וצורות תוכן אחרות. זה במיוחד רלוונטי בשלב הבנייה, התפעול והפיקוח הטכני, שם המידע הווזואלי והטקסטואלי שולט.

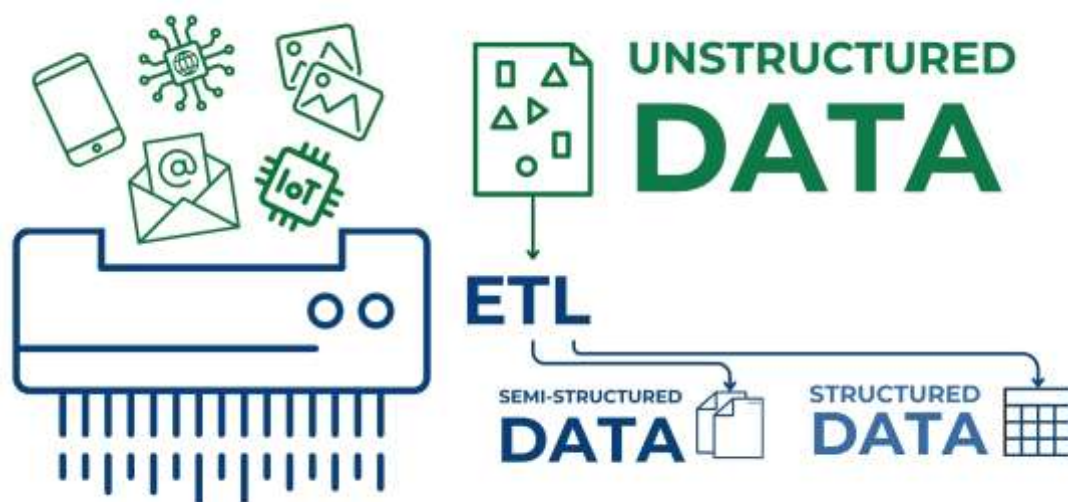
נתונים לא מובנים הם מידע שאין לו מודל או מבנה שנקבע מראש, ואינם מאורגנים בשורות ועמודות מסורתיות, כמו בבסיסי נתונים או בטבלאות.

באופן כללי, ניתן לסווג נתונים לא מובנים לשתי קטגוריות:

■ **נתונים לא מובנים המיוצרים על ידי בני אדם**, הכוללים סוגים שונים של תוכן שנוצר על ידי בני אדם מסמכים טקסטואליים, מיילים, תמונות, סרטונים ועוד.

■ **נתונים לא מובנים המיוצרים על ידי מכונות נוצרות על ידי מכשירים וחיישנים**: אלו הם קבצי יומן, נתוני ודוגמאות אחרות של מידע טלמטרי מאתר הבנייה (**IoT**) תוצאות עבודה של האינטרנט של הדברים, **GPS**.

בניגוד לנתונים מובנים, המאורגנים בנוחות בטבלאות ובבסיסי נתונים, נתונים לא מובנים דורשים שלבים נוספים של עיבוד לפני שהם משתלבים במערכות מידע. השימוש בטכנולוגיות לאיסוף, ניתוח והמרת נתונים כאלה פותח הזדמנויות חדשות לשיפור היעילות בבנייה, צמצום טעויות ומזעור השפעת הגורם האנושי.



עיבוד נתונים לא מובנים מתחיל בהמרתם לנתונים חצי-מובנים ומובנים.

נתונים לא מובנים מהווים עד 80% מכל המידע, שבו נתקלים אנשי מקצוע בחברות, ולכן נבחן בפירוט עם דוגמאות את סוגיהם ועיבודם בפרקים הבאים של הספר.

לשם נוחות הדיון, נפריד את הנתונים הטקסטואליים לקטגוריה נפרדת. למרות שהם סוג של נתונים לא מובנים, החשיבות והנפיצות שלהם בתחום הבנייה דורשות תשומת לב מיוחדת.

נתונים טקסטואליים: בין כאוס לא מובנה למבנה

נתונים טקסטואליים בתחום הבנייה מכסים מגוון רחב של פורמטים וסוגי מידע, החל ממסמכים מודפסים ועד שיטות תקשורת לא פורמליות, כמו מכתבים, שיחות, התכתבויות עבודה ופגישות בעל פה באתר הבנייה. כל הנתונים הטקסטואליים הללו נושאים מידע חשוב לניהול פרויקטי בנייה - מפרטי החלטות פרויקט ושינויים בתוכניות ועד דיונים על בעיות בטיחות ומו"מ עם קבלנים ולקוחות.



נתונים טקסטואליים, אחד מסוגי המידע הפופולריים ביותר בשימוש בתקשורת בין משתתפי הפרויקט.

Word מידע טקסטואלי יכול להיות גם פורמלי וגם לא מובנה. נתונים פורמליים כוללים מסמכים פורמט נתונים לא פורמליים כוללים התכתבויות (.txt) וכן קבצי טקסט של פרוטוקולי ישיבות, PDF, (.doc, .docx).

כמו גם הקלטות אודיו של דיונים (Teams, Zoom, Google Meet) במסג'רים ובמיילים, תמלילים של ישיבות אשר דורשים המרה לטקסט (.mp3,.wav).

אך אם מסמכים כתובים, כגון בקשות רשמיות, תנאי חוזים והודעות דוא"ל, בדרך כלל כבר יש להם מבנה מסוים, אז הודעות בעל פה והתכתבות עבודה לעיתים קרובות נשארות לא מובנות, מה שמקשה על ניתוחן ואינטגרציה שלהן במערכות ניהול פרויקטים.

המפתח לניהול יעיל של נתוני טקסט הוא המרתם לפורמט מובנה. זה מאפשר אינטגרציה אוטומטית של המידע המעובד במערכות קיימות שכבר פועלות עם נתונים מובנים.



המרת תוכן טקסטואלי לנתונים מובנים 3.110 איור.

תהליך זה כולל בדרך. (3.110 איור) כדי להשתמש ביעילות במידע טקסטואלי, יש להמיר אותו אוטומטית לצורת מבנה - כלל מספר שלבים:

- המרת תמונות של מסמכים ושרטוטים לפורמט קריא על ידי מכונה - (OCR) זיהוי טקסט
- זיהוי אוטומטי של פרמטרים מרכזיים (תאריכים, סכומים ומספרים הקשורים - (NLP) ניתוח טקסט (לפרויקט).
- סיווג נתונים - חלוקת מידע לקטגוריות (פיננסים, לוגיסטיקה, ניהול סיכונים)

לאחר זיהוי וסיווג, הנתונים המובנים כבר יכולים להיות משולבים בבסיסי נתונים ולהיות בשימוש במערכות אוטומטיות לדיווח וניהול.

נתונים חצי-מובנים ונתונים עם מבנה חלש

נתונים חצי-מובנים מכילים רמה מסוימת של ארגון, אך אין להם מבנה או תבנית מחמירים. למרות שמידע כזה כולל אלמנטים מובנים (כגון: תאריכים, שמות עובדים ורשימות משימות שבוצעו, (פורמט ההצגה יכול להשתנות באופן משמעותי בין פרויקטים שונים או אפילו בין עובדים שונים. דוגמאות לנתונים כאלה כוללות יומני עבודה, דוחות על עבודה שבוצעה ולוחות זמנים, שיכולים להיות מוצגים בפורמטים שונים.

נתונים חצי-מובנים קלים יותר לניתוח מאשר נתונים לא מובנים, אך הם דורשים עיבוד נוסף כדי להשתלב במערכות ניהול פרויקטים סטנדרטיות.

עבודה עם נתונים חצי-מובנים, המאפיינים את המבנה המשתנה באופן קבוע, מציבה קשיים משמעותיים. זה נובע מכך שהשינוי במבנה הנתונים דורש גישות פרטניות נפרדות לעיבוד וניתוח של כל מקור נתונים חצי-מובנים.

אך אם העבודה עם נתונים לא מובנים דורשת מאמצים רבים, אז עיבוד נתונים חצי-מובנים יכול להתבצע באמצעות שיטות וכלים יחסית פשוטים.

נתונים חלשים-מובנים הם מונח כללי יותר שמתאר נתונים עם מבנה מינימלי או חסר. לרוב מדובר במסמכים טקסטואליים, צ'אטים, הודעות דוא"ל, שבהם מופיעים נתוני מטא בודדים (כגון תאריך, שולח) אך רוב המידע מוצג בצורה כאוטית.

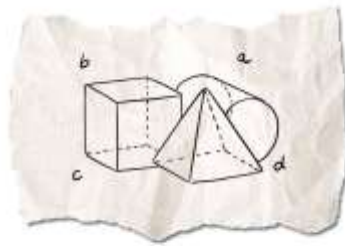
בענף הבנייה נתונים חלשים-מובנים מופיעים בתהליכים שונים. לדוגמה, הם עשויים לכלול:

- חישובי עלויות והצעות מסחריות - טבלאות עם נתונים על חומרים, כמויות ועלויות, אך ללא פורמט אחיד
 - המכילים הערות טקסטואליות ונתוני מטא, אך DWG או PDF שרשומים וסכמות הנדסיות - קבצים בפורמט ללא מבנה קבוע
 - או מערכות אחרות, שיכולים MS Project, Primavera P6-לוחות זמנים לביצוע עבודות - נתונים מ להיות בעלי מבנה ייצוא שונה
 - מכילים אלמנטים של מבנה, אך הצגת הנתונים תלויה בתוכנה ובתקן של – (BIM מודלים של) CAD הפרויקט
- CAD ניתן לסווג גם כנתונים חצי-מובנים. עם זאת, נ את נתוני CAD נתונים גיאומטריים המיוצרים על ידי מערכות הגיאומטריים כסוג נפרד, מכיוון שהם, כמו נתונים טקסטואליים, יכולים לעיתים קרובות להיחשב בתהליכים של (BIM) החברה כסוג נפרד של נתונים.

נתונים גיאומטריים והשימושים שלהם

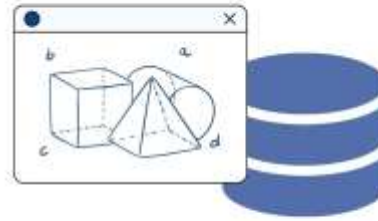
אם מטא-נתונים על אלמנטים של הפרויקט כמעט תמיד נשמרים בצורה של טבלאות, בפורמטים מובנים או חצי-מובנים, אז נתונים גיאומטריים של אלמנטים של הפרויקט ברוב המקרים נוצרים באמצעות כלים מיוחדים של או גופים גיאומטריים (2D) המאפשרים להציג בצורה מפורטת את אלמנטים של הפרויקט כסט של קווים, (3.111, 3D).

3000 BCE - 1960s



physical medium
(artefact)

1960s to present day



digital project data
(CAD data)

סייעו להעביר מידע גיאומטרי ממקורות פיזיים לצורת מסדי נתונים CAD כלים של 3.111 איור

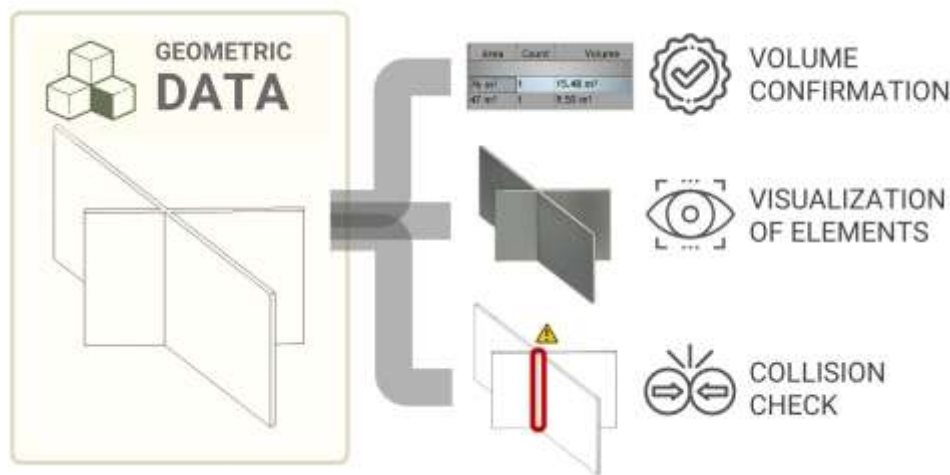
בעבודה עם נתונים גיאומטריים בבנייה ובאדריכלות ניתן להבחין בשלוש תחומי יישום עיקריים של נתונים גיאומטריים (3.112 איור):-

■ בעזרת ליבות גיאומטריות (BIM) CAD אישור נפחים: נתונים גיאומטריים, המיוצרים בתוך תוכנות מיוחדות, נחוצים לקביעת נפחים ומידות של אלמנטים של הפרויקט באופן אוטומטי ומדויק. נתונים אלה כוללים שטחים, נפחים, אורכים ואטריבוטים חשובים אחרים, הנדרשים לתכנון, הכנת תקציב והזמנת משאבים וחומרים.

■ ויזואליזציה של הפרויקט: במקרה של שינויים כלשהם בפרויקט, ויזואליזציה של אלמנטים מאפשרת ליצור באופן אוטומטי שרטוטים מעודכנים במישורים שונים. ויזואליזציה של הפרויקט בשלביו הראשוניים מאפשרת להאיץ את ההבנה ההדדית בין כל המשתתפים, כדי לחסוך זמן ומשאבים במהלך הבנייה.

■ בדיקת התנגשויות: בפרויקטים בנייה והנדסה מורכבים, שבהם האינטראקציה של מספר קטגוריות של אלמנטים (כגון צינורות וקירות) (ללא "קונפליקטים גיאומטריים" היא קריטית, בדיקת התנגשויות משחקת תפקיד מרכזי. השימוש בתוכנה לגילוי התנגשויות מאפשר לזהות מראש קונפליקטים גיאומטריים פוטנציאליים בין אלמנטים של הפרויקט, ובכך למנוע טעויות יקרות במהלך הבנייה.

מאז ומעולם, עם הקמת משרדי ההנדסה והעיצוב, מאז בניית המבנים המורכבים הראשונים, מהנדסי העיצוב סיפקו או פורמטים "A0" מידע גיאומטרי בצורה של שרטוטים, קווים ואלמנטים גיאומטריים שטוחים) על פפירוס, נייר במשך אלפי שנים, בעזרת סרגלים, (3.111 איור) שעל בסיסם מנהלי עבודה ומעריכים, (DWG, PDF, PLT) - ומעגלים, אספו את הנפחים האטריבוטיביים או את כמות האלמנטים וקבוצות האלמנטים



הגיאומטריה היא הבסיס לקבלת פרמטרים נפחיים של אלמנטים, אשר לאחר מכן משמשים לחישוב עלויות ולוחות **3.112** איור זמנים של הפרויקט.

היום, משימה ידנית ומייגעת זו נפתרת על ידי אוטומציה מלאה הודות להופעת המודלים התלת-ממדיים בכלים המודרניים המאפשרת באופן אוטומטי, בעזרת ליבת גיאומטריה מיוחדת, לקבל אטריבוטים נפחיים של כל אלמנט, CAD (BIM) של מבלי הצורך לחשב פרמטרים נפחיים ידנית.

מאפשרים גם לסווג ולקטלג את רכיבי הפרויקט, כך שניתן להוציא מהמאגר של הפרויקט CAD הכלים המודרניים של CO2 (אזור 3.113) טבלאות מפרט לשימוש במערכות שונות, לדוגמה, להערכת עלויות, הכנת לוחות זמנים או חישוב "-ונפחים, כמו גם על דוגמאות מעשיות נדבר בפרק "קבלת נפחים וחישוב כמותי QTO על קבלת המפרטים, טבלאות



שומרים נתונים במאגרי נתונים, המיועדים לאינטגרציה ואינטראקציה עם מערכות אחרות **CAD (BIM)** איור 3.113 כלים של

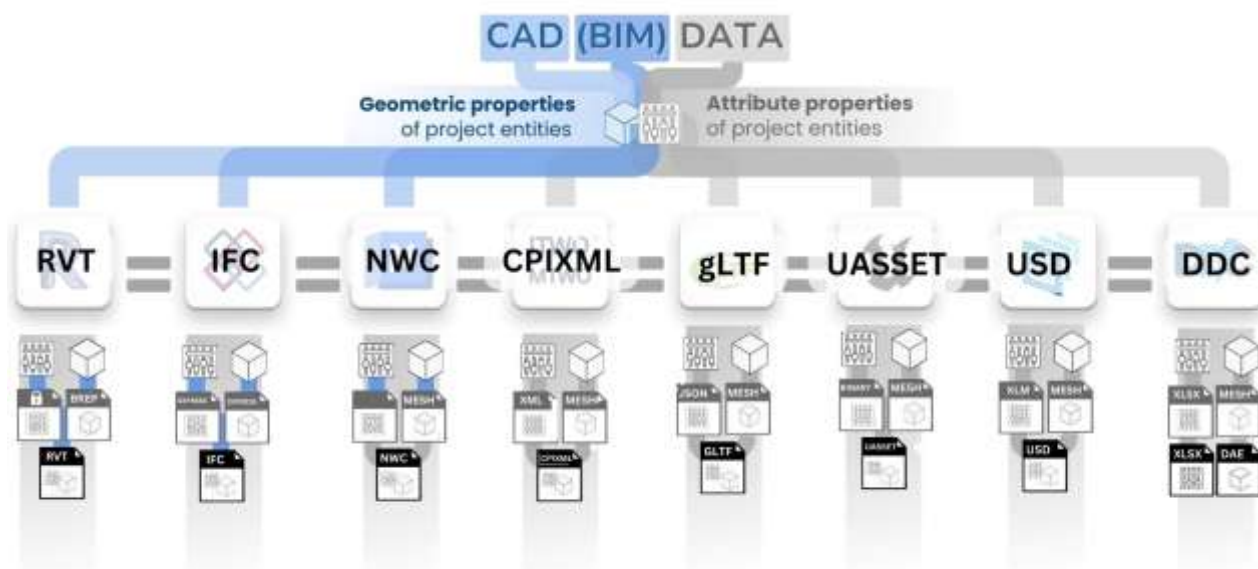
הנתונים הגיאומטריים שנוצרים בפתרונות CAD בשל סגירות המאגרי נתונים והפורמטים שבהם משתמשים בסביבת למעשה הפכו לסוג מידע נפרד. הוא משלב הן את הגיאומטריה של הרכיבים והן את המידע המטה (מובנה או חצי-CAD מובנה), הכלול בקבצים ובפורמטים מיוחדים

מעיצוב ועד אחסון נתונים: CAD נתוני

DWG, DXF, RVT, מודרניות, הנתונים נשמרים בפורמטים ייחודיים, לעיתים קרובות קנייניים CAD-BIM במערכות ועוד. פורמטים אלו תומכים הן בייצוגים דו-ממדיים והן בתלת-ממדיים של אובייקטים, תוך שמירה לא רק DGN, PLN על הגיאומטריה אלא גם על תכונות הקשורות לאובייקטים. הנה הפורמטים הנפוצים ביותר

- פורמט קובץ בינארי, המשמש לאחסון נתוני פרויקט דו-ממדיים (ולעיתים נדירות תלת-ממדיים) (ומטא- DWG נתונים).
- מכיל גיאומטריה, שכבות ונתוני CAD בין מערכות D-3 פורמט טקסט להחלפת שרטוטים 2 - DXF ובינארי ASCII תכונה, תומך בייצוג.
- כולל גיאומטריה תלת-ממדית, תכונות רכיבים, קשרים, פורמט בינארי לאחסון מודלים של - RVT ופרמטרים של הפרויקט.
- כולל גיאומטריה, תכונות CAD (BIM). פורמט טקסט פתוח להחלפת נתוני בנייה בין מערכות - IFC אובייקטים ומידע על הקשרים ביניהם.

PLN, DB1, SVF, NWC, CPIXML, BLEND, BX3, USD, XLSX, למרות שהם שונים במטרתם וברמת הפתיחות) איור 3.114, (כולם יכולים לייצג את אותה מודל מידע של DAE. הפרויקט בצורות שונות. בפרויקטים מורכבים פורמטים אלו משמשים לעיתים במקביל - מהשרטוט ועד לתיאום מודלי הפרויקטים.



MESH, או BREP מתארים את הגיאומטריה באמצעות פרמטרים CAD-איור 3.114 פורמטים פופולריים לאחסון מידע מ ומוסיפים להם נתוני תכונה.

כל הפורמטים המפורטים לעיל מאפשרים לאחסון נתונים על כל רכיב בפרויקט הבנייה ובכל הפורמטים המוזכרים קיימים שני סוגי נתונים מרכזיים

- פרמטרים גיאומטריים - מתארים את הצורה, המיקום והמידות של האובייקט. הגיאומטריה ושימושיה יידונו

CAD (BIM) בפירוט בחלק השישי של הספר, המוקדש לפתרונות

■ תכונות תכונה – מכילות מידע מגוון: על חומרים, סוגי רכיבים, מאפיינים טכניים, מזהים ייחודיים ותכונות אחרות שעשויות להיות לרכיבי הפרויקט

בפרויקטים מודרניים, נתוני התכונה מקבלים חשיבות רבה, שכן הם קובעים את המאפיינים התפעוליים של האובייקטים מאפשרים לבצע חישובים הנדסיים, חישובי עלויות ומספקים אינטראקציה רציפה בין המשתתפים בתכנון, בבנייה ובתפעול. לדוגמה

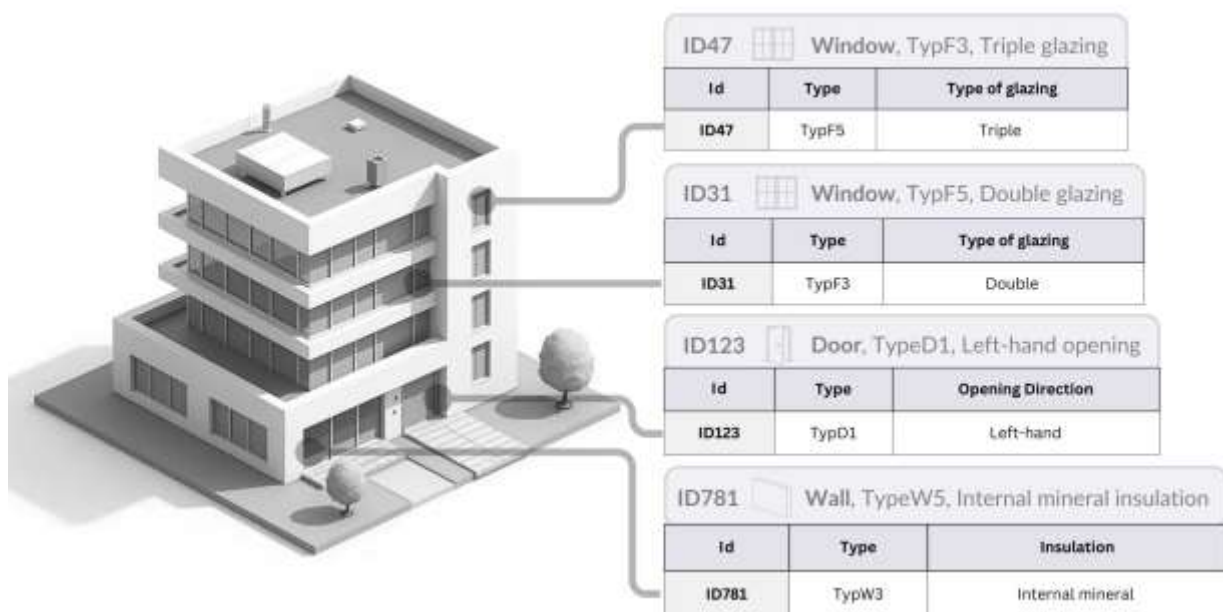
■ (3.21 איור) עבור חלונות ודלתות מצוינים: סוג הבנייה, סוג ההזכוכית, כיוון הפתיחה.

■ עבור קירות נרשמת מידע על החומרים, בידוד תרמי ותכונות אקוסטיות

■ עבור מערכות הנדסיות נשמרים פרמטרים של צינורות, ערוצי אוויר, מסלולי כבלים וחיבורים שלהם

עצמם והן בבסיסי נתונים חיצוניים - כתוצאה מייצוא, המרה (BIM)-CAD פרמטרים אלו יכולים להישמר הן בתוך קבצי באמצעות כלים של הנדסה הפוכה. גישה זו מקלה על אינטגרציה של מידע CAD או גישה ישירה למבנים הפנימיים של פרויקט עם מערכות ופלטפורמות ארגוניות אחרות

היא תהליך של חילוף וניתוח המבנה הפנימי של מודל דיגיטלי (BIM) CAD הנדסה הפוכה בהקשר של במטרה לשחזר את הלוגיקה שלו, מבנה הנתונים ותלותם ללא גישה לאלגוריתמים המקוריים או לתיעוד

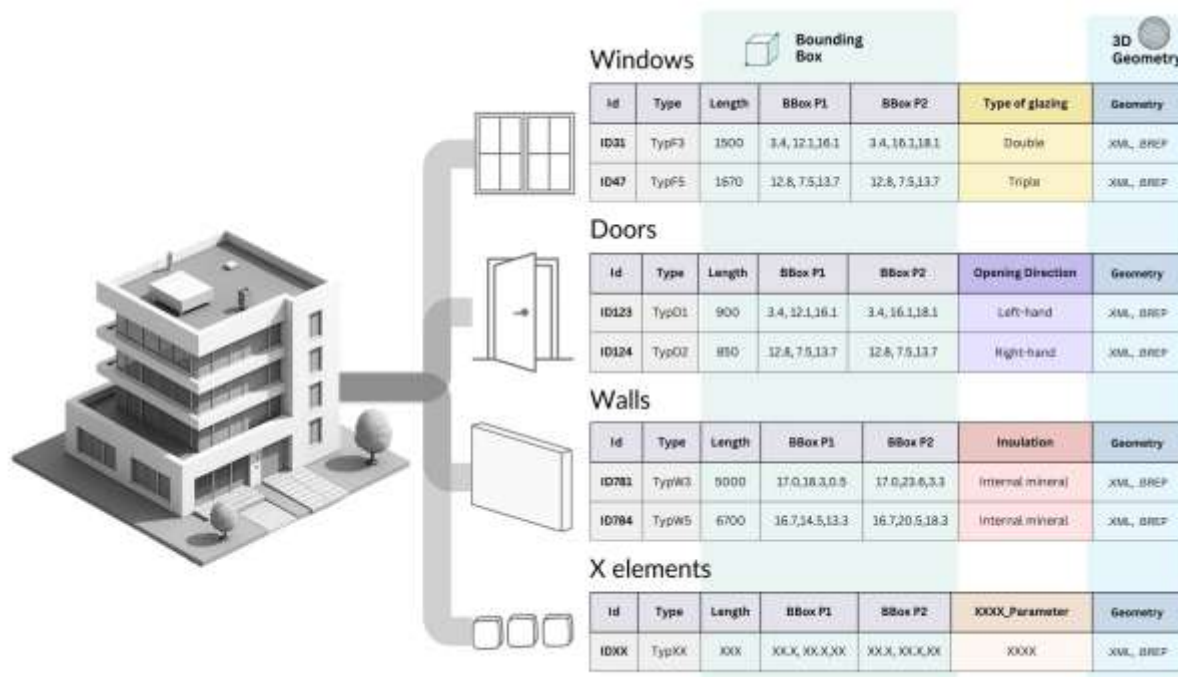


אלמנט הפרויקט, בנוסף לתיאור הגיאומטריה הפרמטרית או הפוליגונית, מכיל מידע על פרמטרים ותכונות של 3.115 איור אלמנטים

בסופו של דבר, סביב כל אלמנט מתהווה סט ייחודי של פרמטרים ותכונות, הכולל הן תכונות ייחודיות של כל אובייקט (למשל, מזהה ומידות) והן תכונות כלליות עבור קבוצות אלמנטים. זה מאפשר לא רק לנתח אלמנטים-ישויות בודדים של הפרויקט, אלא גם לאחד אותם לקבוצות לוגיות, שיכולות לשמש אנשי מקצוע אחרים למשימותיהם וחישוביהם במערכות

ובבסיסי נתונים.

היא אובייקט קונקרטי או מופשט מהעולם האמיתי, שניתן להבחין בו, לתאר (entity: באנגלית) ישות אותו ולהציגו כנתונים.

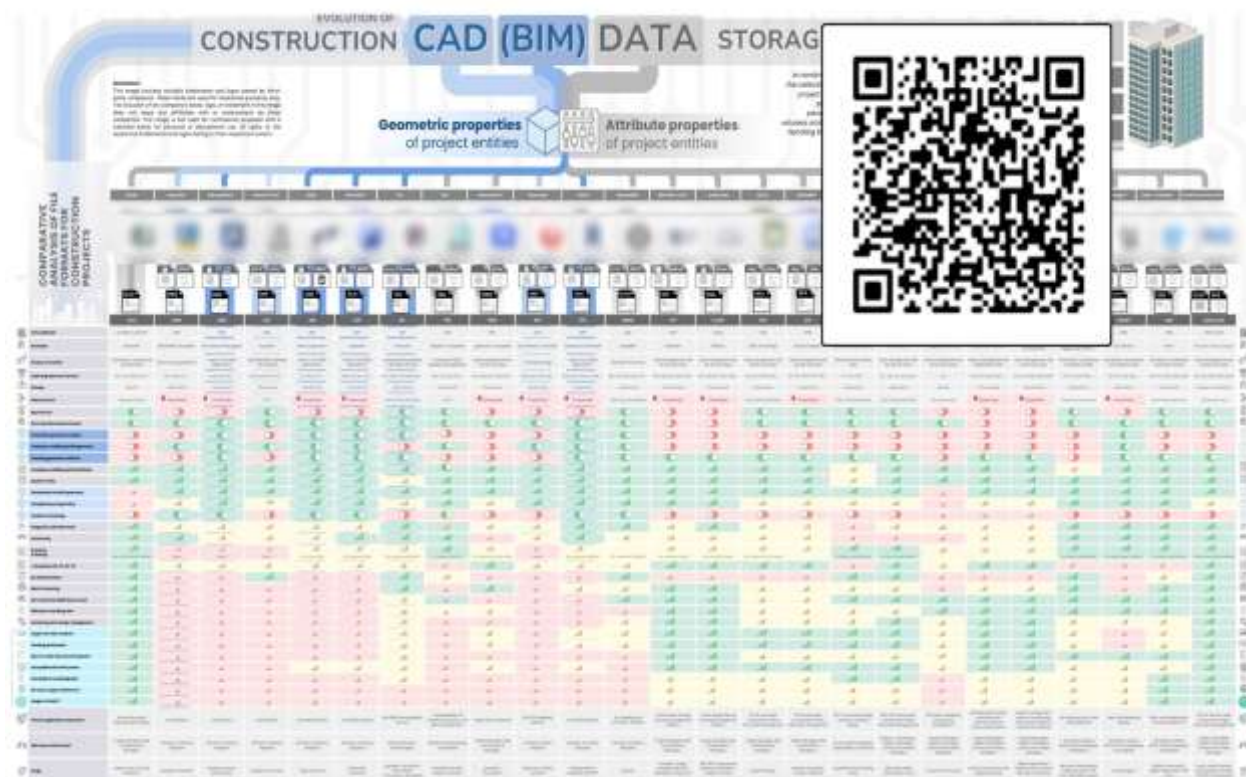


CAD. כל אלמנט בפרויקט מכיל תכונות, אשר או נרשמות על ידי המהנדס או מחושבות בתוך תוכנת **3.116** איור

המפשטים את יצירת **CAD (BIM)** בעשורים האחרונים, בתחום הבנייה הופיעו מספר רב של פורמטים חדשים של אחסון והעברת נתונים. פורמטים אלו עשויים להיות סגורים ופתוחים, טבלאיים, פרמטריים או גרפיים. עם זאת, המגוון והפיצול שלהם מקשים באופן משמעותי על ניהול הנתונים בכל שלבי מחזור חיי הפרויקט. טבלת השוואה של הפורמטים **QR** הגרסה המלאה זמינה באמצעות קוד **3.117** העיקריים בשימוש להחלפת מידע בבנייה מוצגת באיור

ומנחים, שתפקידם הוא לפקח **(BIM)** נכנסים לתמונה מנהלי **CAD**, כדי לפתור בעיות של אינטרופרביליות וגישה לנתוני במערכות אחרות **CAD (BIM)** על הייצוא, לבדוק את איכות הנתונים ולשלב חלקי נתוני

עם זאת, בשל סגירות ומורכבות הפורמטים, קשה לאוטומט את התהליך הזה, מה שמאלץ את המומחים לבצע רבות **(pipeline)** מהפעולות ידנית, ללא אפשרות לבנות תהליכי זרימה מלאים לעיבוד נתונים



[53] טבלת השוואה של הפורמטים העיקריים של נתונים, שבהם נשמרת מידע על אלמנטים של הפרויקט 3.117 איור.

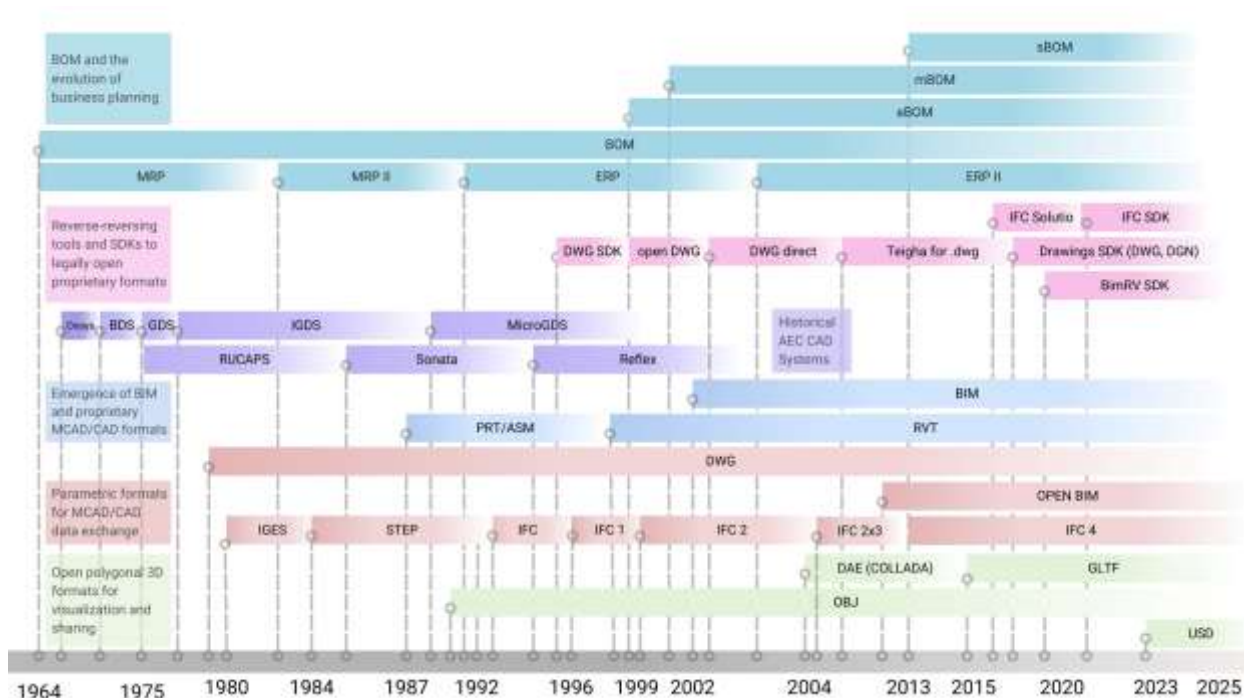
כדי להבין מדוע קיימים כל כך הרבה פורמטים שונים של נתונים, ומדוע רובם סגורים, חשוב להעמיק בתהליכים מה שיתואר בפירוט בחלק השישי של הספר, CAD (BIM), המתרחשים בתוך תוכנות

(מודל מידע בנייתי) BIM כקונספט של CAD שכבת מידע נוספת, המתווספת לגיאומטריה, הוצגה על ידי מפתחי מערכות [54] מונה שיווקי, המתקדם באופן פעיל בתחום הבנייה מאז 2002 -

בתהליכים CAD-והשימוש ב BIM (BOM) הופעת המושג

משנת 2002, נוצר הודות Whitepaper BIM שהוצג לראשונה במסמך, (BIM) מושג המידול המידע של בניינים והפך לניסיון להתאים CAD הוא נבע מיוזמות שיווקיות של מפתחי תוכנות CAD. ליוזמות שיווקיות של יצרני תוכנות עקרונית שכבר הוכיחו את עצמם בהנדסת מכונות לצרכים של תעשיית הבנייה

מפרט הרכב המוצר, שהיה בשימוש פעיל - BOM (Bill of Materials) היה המושג BIM מקור ההשראה עבור ניהול) PDM עם מערכות CAD אפשר לקשר נתונים ממערכות BOM, בתעשייה מאז סוף שנות ה-80. בהנדסת מכונות מה שאפשר ניהול כולל של מידע הנדסי לאורך כל מחזור חיי (ERP)-ניהול מחזור חיי מוצר (PLM, נתוני מוצר -המוצר



ופורמטים דיגיטליים בתעשיית הבנייה ההנדסית (BIM) מידול מידע, (BOM) ריש 3.118. אבולוציה של מפרטים

הכולל לא רק, XBOM (Extended BOM) - הביאה להופעת מבנה מורחב BOM-ההתפתחות המודרנית של מושג ה XBOM ממלא את הרכב המוצר אלא גם תרחישי התנהגות, דרישות תפעול, פרמטרים של עמידות ונתונים לניתוח חיזוי (Single Source of Truth) בבנייה: שני הגישות שואפות להפוך את המודל הדיגיטלי למקור מידע מהימן BIM למעשה את אותה תפקיד כמו עבור כל המשתתפים בפרויקט לאורך כל מחזור חיי האובייקט

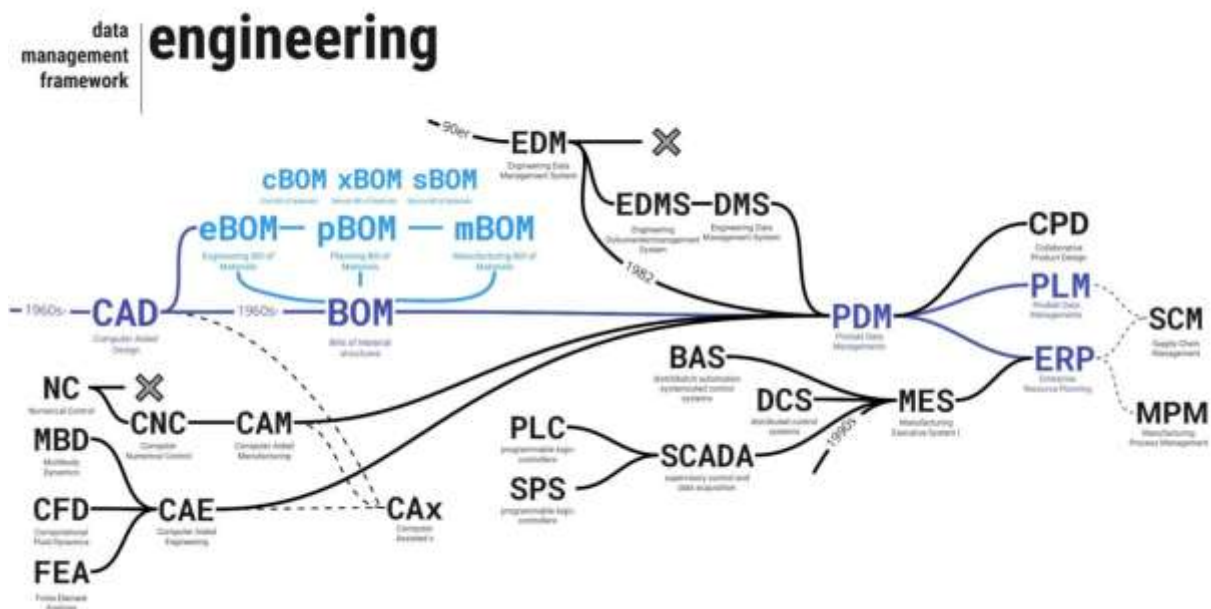
שהותאם במיוחד, (MCAD) הפרמטרי הראשון בשנת 2002 CAD-בבנייה היה הופעת ה BOM-שלב מפתח בהופעת ה מהפכנית להנדסת מכונות MCAD מערכת - Pro-E® לתעשיית הבנייה. הוא פותח על ידי צוות, שיצר קודם לכן את שהופיעה כבר בסוף שנות ה-80 והפכה לסטנדרט בתעשייה

של אז. המשימה העיקרית CAD כבר בסוף שנות ה-80, המטרה הייתה להסיר את המגבלות שהיו קיימות בתוכנות הייתה להפחית את עלויות העבודה הנדרשות לשינויים בפרמטרים של רכיבי הפרויקט ולאפשר עדכון המודל על בסיס דרך מסד נתונים. תפקיד מרכזי בכך היה הפרמטריזציה: קבלת אוטומטית של מאפיינים CAD נתונים מחוץ לתוכנות CAD ממסד הנתונים ושימוש בהם לעדכון המודל בתוך מערכות

העומד בבסיסו, השפיעו רבות על התפתחות שוק, BOM והמושג של מידול פרמטרי של רכיבים עם Pro-E במשך 25 שנה, המודל הזה היה בתעשייה, ורבות מהמערכות המודרניות הפכו. MCAD-וה CAD-ה. ליורשות הקונספטואליות שלה

המטרה היא ליצור מערכת שתהיה גמישה מספיק כדי לעודד מהנדס לשקול בקלות מבנים שונים. ועלויות השינויים המסורתיות מגבילות באופן לא מציאותי CAD/CAM בפרויקט צריכות להיות קרובות ככל האפשר לאפס. תוכנות Parametric את השינויים הזולים רק בשלב הראשוני ביותר של תהליך התכנון. - סמואל גיזנברג, מייסד חברת RVT. המשתמש בפורמט CAD, ומורה לממציא המוצר MCAD Pro-E מפתח המוצר, Technology Corporation®.

הן משחקות תפקיד. ERP, MRP, PLM, PDM בתעשיית ההנדסה המכאנית הפלטפורמות המרכזיות הפכו למערכות ומארגנות את הפעילות CAX (CAD, CAM, CAE) מרכזי בניהול נתונים ותהליכים, אוספות מידע ממערכות אינטגרציה כזו מאפשרת. (3.118 איור) (BOM: eBOM, pBOM, mBOM) הפרויקטית על בסיס מבנה המוצר להפחית את מספר השגיאות, להימנע מהכפלת נתונים ולספק מעקב רציף בכל השלבים - מעיצוב ועד ייצור.



והעברתם למערכות ניהול CAX התפתח היסטורית בשנות ה-60 כדרך לארגון נתונים ממערכות BOM-ה-3.119 איור.

BOM, לשעבר ומבוסס על גישת Pro-E על ידי אחד מהספקים המובילים, שפותח על ידי צוות CAD רכישת פתרון (בשנים 2002–2003) [60][61]. [כבר מאמצע Whitepaper BIM ציינה פרסום כמעט מידי של סדרת מסמכי בתעשיית הבנייה, מה שהגביר באופן ניכר את העניין בתוכנה BIM-שנות ה-2000 החלה קידום פעיל של מושג ה - פרמטרי, שהוקדם על ידי ספק זה CAD - Pro-E פרמטרית. הפופולריות עלתה כל כך מהר, שהפורק הבנייתי של עד תחילת שנות ה-2020 הוא הפך. (3.120 איור) למעשה הדיח מתחרים בתחום התכנון הארכיטקטוני והבנייה -BIM (CAD) [62]-לדומיננטי בשוק ה



לשעבר עם תמיכה ב **Pro-E** פרמטרי, שנוצר על ידי צוות **CAD**: (מול **IFC** ו **RVT**) הפופולריות של חיפוש בגוגל **3.120** איור
צבר פופולריות כמעט בכל המדינות בעולם, **BOM-BIM**.

צבר מספר רב של פרשנויות, שהרב-משמעות שלהן נובעת מהקונספטים **BIM** במהלך 20 השנים האחרונות, המונח השיחק תפקיד חשוב בפופולריזציה של **ISO 19650** השיווקיים הראשוניים שהופיעו בתחילת שנות ה-2000. התקן כגישה "מבוססת מדע" לניהול מידע. עם זאת, בטקסט של התקן, המוקדש לניהול נתונים **BIM**-המונח, למעשה קיבע את ה. אמנם מוזכר, אך לא מקבל הגדרה ברורה **BIM** המונח, **BIM** במהלך מחזור חיי האובייקטים באמצעות

[61] [בשנת 2002] 60 [ובשנת **BIM 2003** על **Whitepaper** באתר המקורי של הספק, שפרסם את סדרת המסמכים (ניהול מחזור חיי המוצר) **PLM**-(חשבונות חומרים) **BOM** שוחזרו למעשה חומרים שיווקיים על הקונספטים של [63]. כבר בשנות ה-90 **Pro-E** שהיו בשימוש בתוכנת ההנדסה המכאנית

מודל מידע בניין - גישה חדשנית חדשה לתכנון, בנייה וניהול בניינים, שהוצגה על ידי חברת [.....שם חברת ספק
בשנת 2002, שינתה את התפיסה של אנשי מקצוע בתעשייה ברחבי העולם לגבי כיצד ניתן ליישם **CAD**
[61] **Whitepaper BIM, 2003** טכנולוגיות בתכנון, בנייה וניהול בניינים. מסמך

נקשר ישירות למושג של בסיס נתונים אינטגרטיבי מרכזי. כפי שצוין במסמך **BIM**-בפרסומים המוקדמים הללו, ה. הוא ניהול המידע על הבניין, שבו כל העדכונים מתבצעים באחסון אחד, מה שמבטיח **BIM**-של 2003, ה **Whitepaper** (חשבונות חומרים – **BOM**) סנכרון של כל השרטוטים, החתכים והמפרטים

מתואר כמערכת ניהול מידע על בניינים, שבה כל העדכונים וכל השינויים מתבצעים בבסיס הנתונים. כך, בין **BIM**
CAD אם מדובר בתוכניות, חתכים או שרטוטים על דפים, הכל תמיד מתואם, מוסכם ועדכני. – אתר החברה של ספק
[54] **BIM, 2003** עם מסמך לבן על

הרעיון של ניהול תכנון באמצעות בסיס נתונים אינטגרטיבי אחד היה נושא לדיון רחב כבר במחקרים של שנות ה-80. עד 2004. (6.12 איור) "של צ'אק איסטמן" [57] כלל 43 אזכורים למונח "בסיס נתונים **BDS** לדוגמה, הקונספט של מספר זה צנח כמעט בחצי – ל-23 במסמך הלבן של 2002 [64]. [וכבר באמצע שנות ה-2000 **BIM**, במסמכים על נושא בסיסי הנתונים כמעט ונעלם מחומרי השיווק של הספקים ומכלל סדר היום של הדיגיטציה

עם הזמן הדגש עבר לגיאומטריה, BIM-אף על פי שבסיס הנתונים גישה אליו תוכנו במקור כמרכז מערכת ה – בשנת 2002 BIM בשנת 1994, שפרסם את המסמך הלבן על IFC עם זאת, המפרסם של תקן D.ויזואליזציה ו-3 IGES, STEP אותו ספק – ציין במסמך הלבן של תחילת שנות ה-2000 את המגבלות של פורמטים נייטרליים, כמו CAD: ואת הצורך בגישה ישירה לבסיסי הנתונים של IFC,

יישומים שונים עשויים להיות לא תואמים, ונתונים שהוזנו מחדש עשויים להיות לא מדויקים. [...] התוצאה של תכנון עלייה בעלויות, הגדלת זמן היציאה לשוק והפחתת איכות המוצר. כיום, כל היישומים: [CAD] ממוחשב מסורתי המרכזיים משתמשים בממשקי תעשייה סטנדרטיים להחלפת נתונים ברמה נמוכה. באמצעות הסטנדרטים הישנים להחלפת נתונים בין [STEP/IGES] הוא דה-פקטו ודה-יורה העתק של פורמט STEP [IFC] או החדשים IGES יישומים של ספקים שונים, משתמשים יכולים להשיג תאימות מסוימת של נתונים בין המוצרים הטובים ביותר פועלים רק ברמות נמוכות, והם אינם יכולים להחליף נתונים עשירים כמו המידע STEP ו-IGES בתחומם. אך שנוצר על ידי יישומים מובילים מודרניים. [...] ואילו סטנדרטים אלה ואחרים משתפרים כמעט מדי יום, הם תמיד יישארו מאחור ביחס למוצרים של ספקים מודרניים מבחינת עושר הנתונים [...]. תוכניות במסגרת היישום צריכות IGES, STEP [IFC] להיות מסוגלות להחליף נתונים ולשמור על עושרם, מבלי להיעזר במתרגמים נייטרליים, כמו במקום זאת, יישומי המסגרת צריכים להיות מסוגלים לגשת ישירות לבסיס הנתונים המרכזי של PATRAN. או תכנון וייצור "CAD (IFC, BIM) כדי לא לאבד את הפרטים והדיוק של המידע. – מסמך לבן של ספק, CAD, [65] אינטגרטיביים: יתרונות והצדקה, "2000"

נחשב לבסיס הנתונים, ולא CAD כך, כבר בשנות ה-80 ותחילת שנות ה-2000, רכיב מרכזי בתכנון דיגיטלי בסביבת הוצע לוותר על מתרגמים ולספק גישה ישירה של יישומים לנתונים. עם זאת, IFC לפורמט קובץ או לפורמט נייטרלי הפך להיות דומה לאסטרטגיית "חלקו ושלטו", שבה העדיפות BIM במציאות, עד אמצע שנות ה-2020, הקונספט של נותרה באינטרסים של ספקי התוכנה, המשתמשים בליבות גיאומטריות סגורות, ולא בפיתוח של חילופי מידע פתוחים.

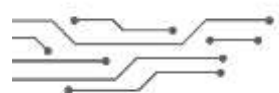
נתפס כחלק בלתי נפרד מתעשיית הבנייה. אך בעשרים השנים האחרונות ההבטחות לאינטראקציה פשוטה BIM היום ואינטגרציה של נתונים נותרו במידה רבה לא ממומשות. רוב הפתרונות עדיין תלויים בפורמטים סגורים או פורמטים כמו גם IFC, open BIM, BIM נייטרליים, כמו גם בכלים מיוחדים. נבחן בפירוט את שאלות ההיסטוריה של הופעת שיווק, מציאות ועתיד: "BIM ו-CAD" את בעיות האינטרופרוביליות והגרעינים הגיאומטריים בחלק השישי של הספר "נתוני התכנון בבנייה".

ככלי מודלינג CAD (BIM) היום התעשייה ניצבת בפני אתגר מרכזי - לעבור מהבנה המסורתית של לשימוש בו כבסיס נתונים מלא. זה דורש גישות חדשות לעבודה עם מידע, ונטישה של תלות באקוסיסטמות סגורות והטמעה של פתרונות פתוחים.

ובזכות הפצת טכנולוגיות קוד פתוח ו, CAD עם התפתחות הכלים להנדסה הפוכה, המאפשרים גישה לבסיסי נתונים של משתמשים ומפתחים בתעשיית הבנייה מתרחקים יותר ויותר מהמונחים המעורפלים של ספקי התוכנה. במקום, LLM, זאת, תשומת הלב מתמקדת במה שחשוב באמת: נתונים) בסיסי נתונים (ותהליכים

מאחורי האותיות המודרניות והוויזואליזציות מסתתרות פרקטיקות סטנדרטיות לניהול נתונים: אחסון, כמו בתעשיות ETL (Extract, Transform, Load) העברה והמרה - כלומר, התהליך הקלאסי אחרות, הדיגיטציה של הבנייה דורשת לא רק סטנדרטים להחלפת נתונים, אלא גם עבודה מסודרת עם מידע מגוון.

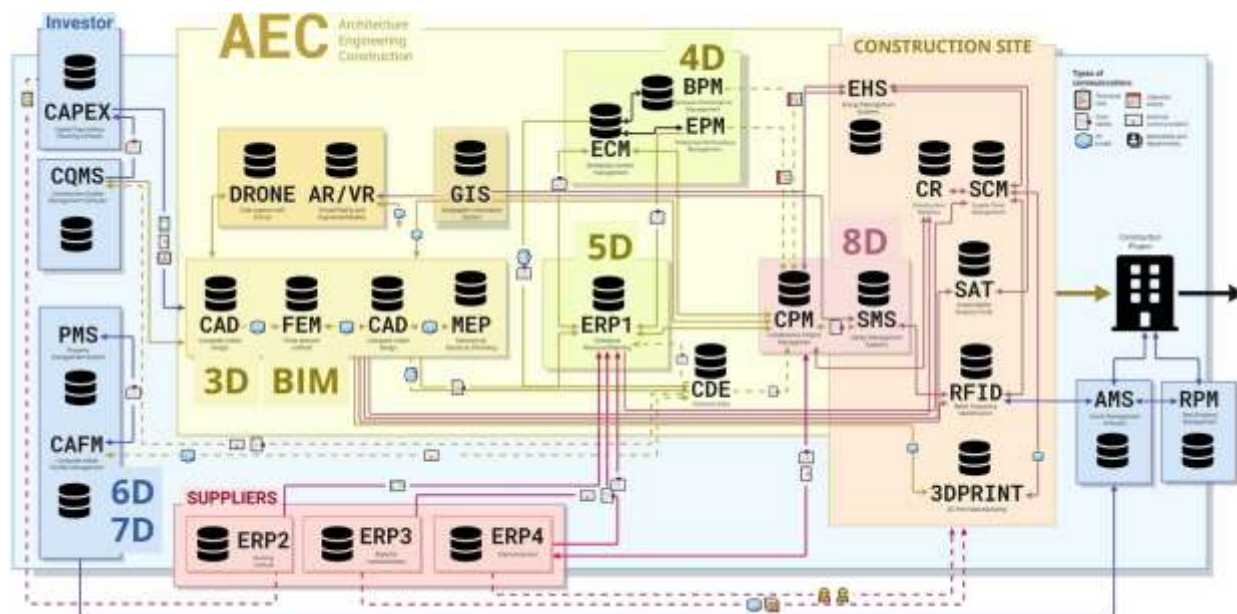
חברות צריכות לחשוב מחדש על הגישה שלהן לניהול מידע. זה, CAD (BIM) כדי למצות את הפוטנציאל של נתוני יוביל בהכרח לרכיב מרכזי של טרנספורמציה דיגיטלית - אחדות, סטנדרטיזציה ומבנה משמעותי של נתונים, אשר עימם עובדים אנשי מקצוע בתעשיית הבנייה מדי יום.



פרק 3.2. אחידות ומבנה של נתונים

מילוי מערכות בנתונים בתעשיית הבנייה

בין אם מדובר בתאגידים גדולים או חברות בינוניות, אנשי מקצוע עוסקים מדי יום במילוי מערכות תוכנה ובסיסי נתונים עם ממשקים שונים במידע מגוון (איור 3.21), אשר באמצעות מנהלים צריכים לפעול בהרמוניה זה עם זה. בדיוק המורכבות של מערכות ותהליכים אינטראקטיביים יוצרת בסופו של דבר הכנסה ורווח לחברה.



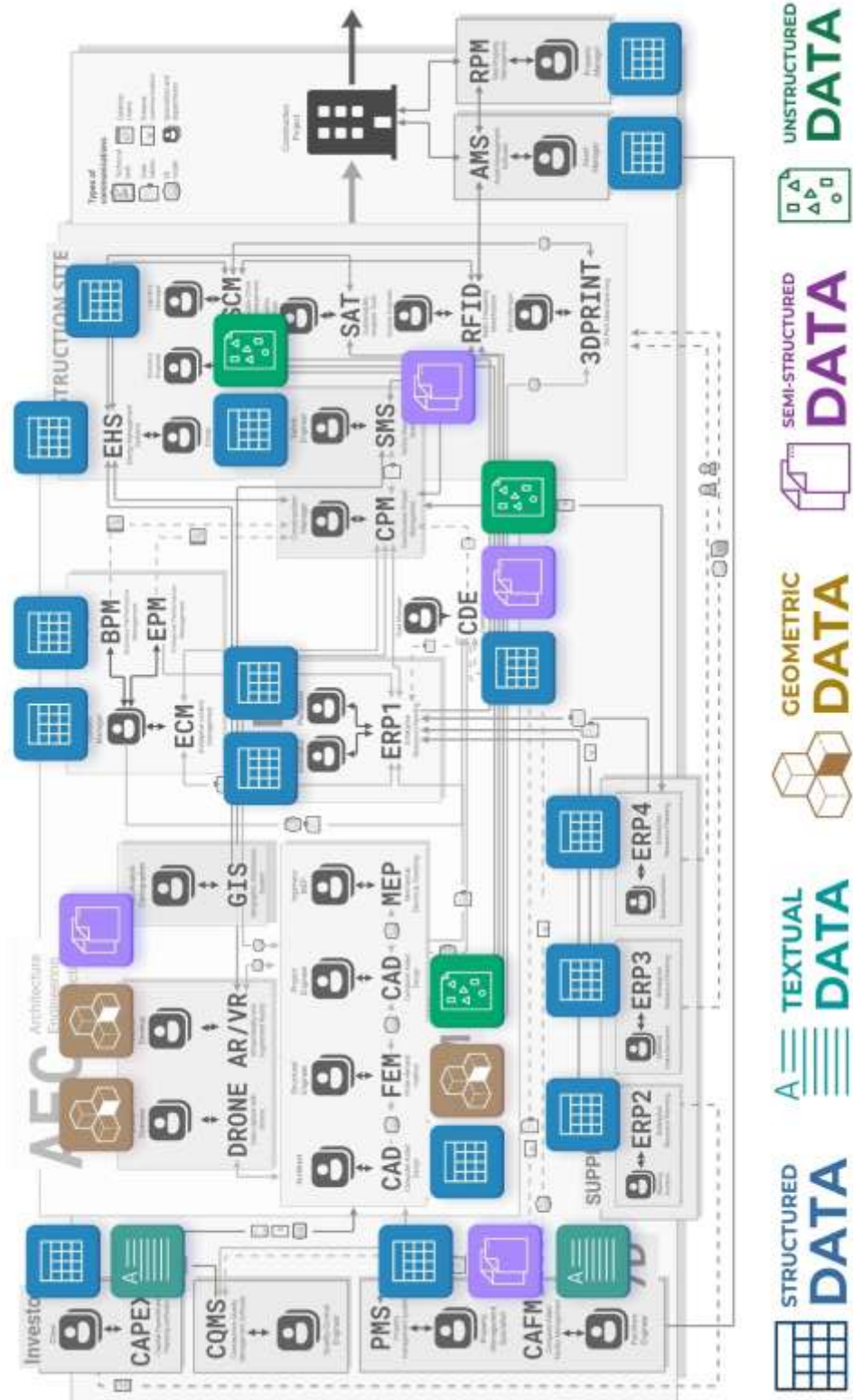
איור 3.21 כמעט כל מערכת או אפליקציה בשימוש בעסקי הבנייה מבוססת על אחת ממערכות ניהול בסיסי הנתונים הפופולריות (RDBMS).

כל אחת מהקטגוריות של המערכות שהוזכרו קודם לכן, המיועדות לתעשיית הבנייה, פועלת עם סוגי נתונים המתאימים לתפקיד הפונקציונלי של אותן מערכות. כדי לעבור מרמה אבסטרקטית לרמה קונקרטית, אנו עוברים מסוגי נתונים לייצוגם בפורמטים ובמסמכים.

לרשימה שהוצגה קודם לכן (איור 1.24) נוסף כעת סוגי פורמטים ומסמכים ספציפיים, עימם הן פועלות לעיתים קרובות

משקיע (CAPEX) ■

- נתונים פיננסיים: תקציבים, תחזיות הוצאות (נתונים מובנים)
- נתונים על מגמות השוק: ניתוח שוק (נתונים מובנים ולא מובנים)
- נתונים משפטיים וחוזיים: חוזים (נתונים טקסטואליים)



איור 3.22 בתעשיית הבנייה נעשה שימוש במגוון מערכות עם ממשקים שונים, הפועלות עם סוגי נתונים שונים.

■ (PMS, CAFM, CQMS) מערכות ניהול

- נתוני פרויקט: לוחות זמנים, משימות (נתונים מובנים)
- נתוני תחזוקה של המתקן: תוכניות תחזוקה (נתונים טקסטואליים וחצי-מובנים)
- נתוני בקרת איכות: תקנים, דוחות בדיקה (נתונים טקסטואליים ולא-מובנים)

■ CAD, FEM ו-BIM

- שרטוטים טכניים: תוכניות ארכיטקטוניות, מבניות (נתונים גיאומטריים ולא-מובנים)
- מודלים של מבנים: מודלים תלת-ממדיים, נתוני חומרים (נתונים גיאומטריים וחצי-מובנים)
- חישובים הנדסיים: ניתוח עומסים (נתונים מובנים)

■ (EHS, SCM) מערכות ניהול אתר הבנייה

- נתוני בטיחות ובריאות: פרוטוקולי בטיחות (נתונים טקסטואליים ומובנים)
- נתוני שרשרת אספקה: מלאים, הזמנות (נתונים מובנים)
- דוחות יומיים: שעות עבודה, פרודוקטיביות (נתונים מובנים)

■ הדפסה תלת-ממדית, AR/VR, GIS, רחפנים

- נתוני גיאו: מפות טופוגרפיות (נתונים גיאומטריים ומובנים)
- נתונים בזמן אמת: סרטונים ותמונות (נתונים לא-מובנים)
- מודלים להדפסה תלת-ממדית: שרטוטים דיגיטליים (נתונים גיאומטריים)

■ (4D BPM, 5D ERP) מערכות ניהול נוספות

- נתוני זמן ועלויות: לוחות זמנים, אומדנים (נתונים מובנים)
- ניהול שינויים: רישומים על שינויים בפרויקט (נתונים טקסטואליים ומובנים)
- דיווח על תוצאות פעילות: מדדי הצלחה (נתונים מובנים)

■ (CDE, RFID, AMS, RPM) אינטגרציה של נתונים ותקשורת

- חילופי נתונים: חילופי מסמכים, מודלים של נתונים (נתונים מובנים וטקסטואליים)
- ונתוני מעקב: לוגיסטיקה, ניהול נכסים (נתונים מובנים) RFID
- ניטור ובקרה: חיישנים במתקנים (נתונים מובנים ולא-מובנים)

כך, כל מערכת בתעשיית הבנייה – ממערכות ניהול אתר הבנייה ועד בסיסי נתונים של תפעול – פועלת עם סוגי מידע שונים: מובנים, טקסטואליים, גיאומטריים ועוד. נוף הנתונים "שבו מתמודדים אנשי מקצוע מדי יום הוא מגוון מאוד. עם זאת, פשוט למנות את הפורמטים אינו חושף את כל המורכבות של העבודה האמיתית עם מידע

בפועל, חברות מתמודדות עם כך שדאטה, גם כאשר היא מתקבלת ממערכות, אינה מוכנה לשימוש "כפי שהיא". במיוחד ופורמטים אחרים, אשר קשה לנתח באמצעות כלים סטנדרטיים CAD קבצי PDF מדובר בטקסטים, תמונות, מסמכי לכן, הצעד הקריטי הבא הוא טרנספורמציה של נתונים – תהליך שללאו לא ניתן לאוטומט את העיבוד, הניתוח. הוויזואליזציה וקבלת ההחלטות בצורה יעילה

טרנספורמציה של נתונים: יסוד קריטי לניתוח עסקי מודרני

היום רוב החברות מוצאות את עצמן בפני פרדוקס: כ-80% מהתהליכים היומיומיים שלהן עדיין מתבססים על נתונים עם זאת, 80% מהמידע החדש (RDBMS) מוכרות ומסדי נתונים רלציוניים Excel מסודרים קלאסיים - טבלאות

שמגיע לאקוסיסטמה הדיגיטלית של החברות הוא בעל אופי לא מסודר או מסודר באופן חלקי. זהו טקסט, גרפיקה, הקלטות אודיו ווידאו, התכתבויות אלקטרוניות ועוד, PDF, מסמכים בפורמט, CAD, גיאומטריה, תמונות, מודלים

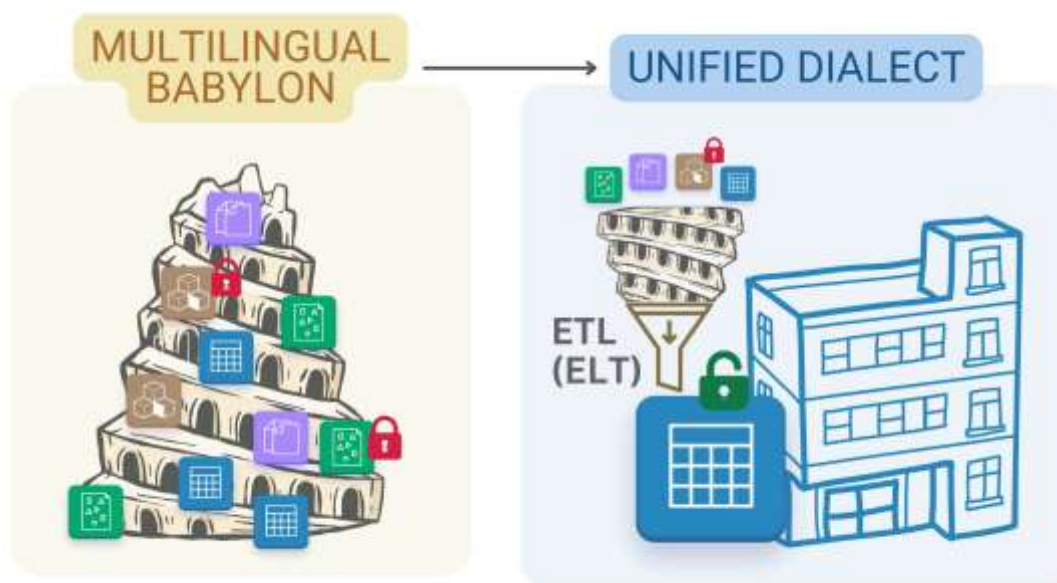
יתרה מכך, נפח הנתונים הלא מסודרים ממשיך לגדול במהירות - הגידול השנתי מוערך ב-55%–65%. דינמיקה כזו יוצרת קשיים משמעותיים בשילוב המידע החדש בתהליכי העסקים הקיימים. התעלמות מזרם הנתונים המגוון הזה מובילה להיווצרות פערים מידע ולירידה בניהוליות של כל הסביבה הדיגיטלית של החברה



- הגידול השנתי בנפח הנתונים הלא מסודרים יוצר בעיות בשילוב המידע הזורם בתהליכי העסקים

התעלמות מנתונים לא מסודרים ומסודרים באופן חלקי בתהליכי האוטומציה עלולה להוביל לפערים משמעותיים בנוף המידע של החברה. בעולם המודרני שבו המידע זורם ללא שליטה ובכמויות עצומות, חברות נדרשות לאמץ גישה היברידית לניהול נתונים, הכוללת שיטות עבודה יעילות עם כל סוגי הנתונים.

המפתח לניהול נתונים יעיל טמון בארגון, סידור ומיון של סוגי הנתונים השונים. תהליך זה ממיר את הכאוס של הנתונים - למבנים מאורגנים לשילוב במערכות, ובכך מאפשר קבלת החלטות על בסיסם



- המשימה המרכזית של מחלקות ניהול הנתונים היא להמיר את "בבל" של נתונים מגוונים ולא אחידים למערכת מסודרת וממינת

"אחד מהמכשולים המרכזיים בדרך לאיחוד כזה הוא רמת התאימות הנמוכה בין פלטפורמות דיגיטליות שונות - "סילונים עליהם דיברנו בפרקים הקודמים

רמת התאימות הנמוכה בין פלטפורמות בנייה (NIST, ארה"ב) לפי דוח המכון הלאומי לתקנים וטכנולוגיה שונות מובילה לאובדן מידע ולעלויות נוספות משמעותיות. רק בשנת 2002, בעיות התאמה בתוכנה גרמו לאובדן של 15.8 מיליארד דולר בשוק הבנייה בארה"ב, כאשר שני שלישים מהאובדן נופלים על בעלי הנכסים ומפעילי הבניינים, במיוחד במהלך השימוש והתחזוקה. במחקר גם מצוין כי סטנדרטיזציה של פורמטי נתונים עשויה לצמצם את האובדנים הללו ולהגביר את היעילות בכל שלבי מחזור חיי הפרויקט

משנת 2016, אשר כלל 16 אלף אנשי מקצוע בתחום הנתונים ברחבי העולם, הבעיה CrowdFlower לפי מחקר של המרכזית נותרה נתונים "מלוכלכים" ולא אחידים. לפי מחקר זה, המשאב היקר ביותר הוא לא מסדי הנתונים הסופיים ולא מודלים של למידת מכונה, אלא הזמן של המומחים המוקדש להכנת המידע

לניקוי, עיצוב וארגון נתונים נדרש עד 60 אחוז מזמן העבודה של אנליסט ומנהל נתונים. כמעט חמישית מהזמן מוקדשת לחיפוש ואיסוף מערכי נתונים נדרשים, אשר לעיתים קרובות מוסתרים במחסנים סגורים ("סילונים") ואינם נגישים, לניתוח. ורק כ-9 אחוז מהזמן מוקדש ישירות למידול, אנליזה, בניית תחזיות ובדיקת השערות. כל השאר כולל תקשורת ויזואליזציה, דיווחים וחקר מקורות מידע נוספים

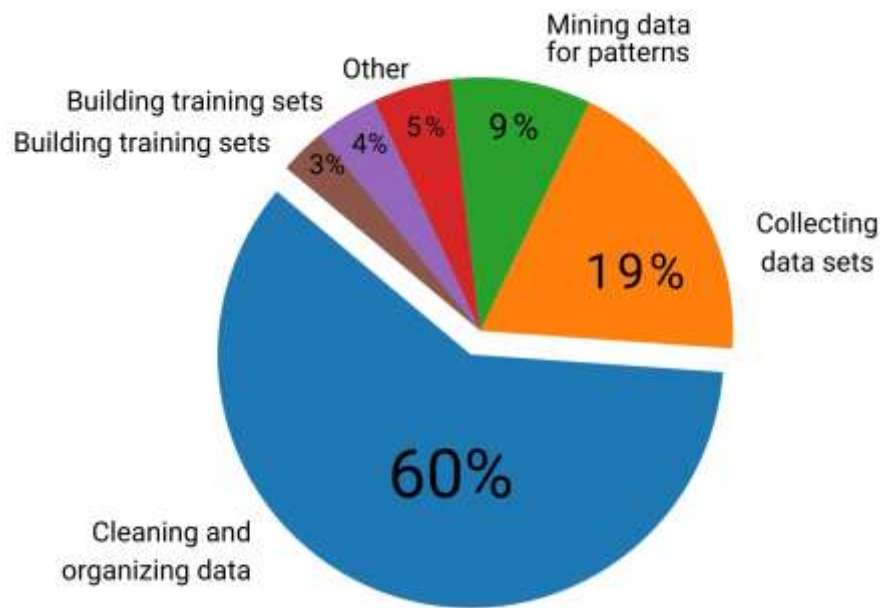
- (3.25 איור) בממוצע, העבודה של מנהל הנתונים מחולקת כך

■ ניקוי וארגון נתונים (60%): (קיום נתונים נקיים ומסודרים יכול לקצר משמעותית את זמן העבודה של האנליסט ולהאיץ את תהליך ביצוע המשימות

■ איסוף נתונים (19%): (הקושי העיקרי עבור אנשי מקצוע בתחום מדעי הנתונים הוא חיפוש מערכי נתונים רלוונטיים. לעיתים קרובות, נתוני חברות נערכים בסילונים מאורגנים באופן כאוטי, מה שמקשה על הגישה למידע הנדרש

■ מידול/למידת מכונה (9%): (לעיתים קרובות, זה מסובך עקב חוסר בהירות במטרות העסקיות מצד הלקוחות. חוסר בהגדרת משימה ברורה יכול לבטל את הפוטנציאל של אפילו המודל האיכותי ביותר

■ משימות נוספות (5%): (מעבר לעיבוד נתונים, אנליסטים נדרשים לעסוק במחקר, לחקור את הנתונים מכיוונים שונים, לתקשר תוצאות באמצעות ויזואליזציות ודיווחים, כמו גם להמליץ על אופטימיזציה של תהליכים ואסטרטגיות



על מה מנהלי נתונים, העובדים עם נתונים, מבזבזים את רוב זמנם (על פי חומרי [70] 3.25 איור).

BizReport-שפורסם ב, Xplenty הערכות אלו מאושרות גם על ידי מחקרים אחרים. לפי מחקר של (BI) בשנת 2015 [71], בין 50% ל-90% מזמן העבודה של אנשי מקצוע בתחום האנליזה העסקית מוקדש להכנת נתונים לניתוח.

ניקוי, בדיקה וארגון נתונים מהווים בסיס קריטי לכל התהליכים הבאים של עיבוד נתונים ואנליזות, תופסים עד 90% מזמן העבודה של אנשי מקצוע בתחום הנתונים.

עבודה זו, שהיא קשה ולא נראית למשתמש הסופי, היא בעלת חשיבות מכרעת. טעויות בנתונים המקוריים מעוותות בהכרח את תוצאות הניתוח, מבלבלות ויכולות להוביל לטעויות ניהול יקרות. לכן, תהליכי ניקוי וסטנדרטיזציה של נתונים החל מהסרת כפילויות ומילוי חוסרים ועד התאמת יחידות מידה והבאתן למודל משותף - הופכים לאבן הפינה של - האסטרטגיה הדיגיטלית המודרנית.

כך, טרנספורמציה, ניקוי וסטנדרטיזציה קפדניים של נתונים לא רק תופסים חלק גדול מזמן העבודה של אנשי מקצוע (עד 80% מעבודת הנתונים, (אלא גם קובעים את האפשרות לשימוש יעיל בהם במסגרת תהליכים עסקיים מודרניים עם זאת, ארגון וניקוי נתונים לבדם אינם מספקים את המשימה של ניהול אופטימלי של זרמי מידע בחברה. במהלך שלב הארגון והמבנה, יש לבחור את מודל הנתונים המתאים, אשר משפיע ישירות על הנוחות והיעילות של העבודה עם המידע בשלבים הבאים של העיבוד.

מכיוון שהנתונים והמטרות העסקיות שונים, חשוב להבין את המאפיינים של מודלי הנתונים ולדעת לבחור או ליצור את המבנה הנדרש. בהתאם לרמת הסדר והאופן שבו מתוארים הקשרים בין המרכיבים, ניתן להבחין בשלוש מודלים עיקריים: מודלים מסודרים, מודלים חצי-מסודרים ומודלים גרפיים. כל אחד מהם מתאים למשימות שונות ויש לו יתרונות וחסרונות משלו.

מודלים של נתונים: יחסים בנתונים וקשרים בין רכיבים

הנתונים במערכות מידע מאורגנים בדרכים שונות – בהתאם למשימות ולדרישות לאחסון, עיבוד והעברת מידע. ההבדל המרכזי בין סוגי מודלי הנתונים, בצורה שבה מאוחסן המידע, טמון ברמת הסדר ובאופן תיאור הקשרים בין המרכיבים.

נתונים מסודרים כוללים תבנית ברורה וחוזרת: הם מאורגנים בטבלאות עם עמודות קבועות. פורמט זה סינון ואגרגציה. דוגמאות – מסדי נתונים, SQL, מבטיח ניבוי, פשטות בעיבוד ויעילות בביצוע שאילתות (RDBMS), Excel, CSV.

נתונים חצי-מסודרים מאפשרים מבנה גמיש: מרכיבים שונים יכולים להכיל סט שונה של תכונות ולהיות מאוחסנים או פורמטים דוקומנטריים אחרים. נתונים אלה נוחים כאשר יש צורך לדגם JSON, XML – בצורה היררכית. דוגמאות: -אובייקטים מקוננים וקשרים ביניהם, אך מצד שני, זה מקשה על ניתוח וסטנדרטיזציה של הנתונים.

	Data Model	Storage Format	Example
	Relational	CSV, SQL	A table of doors in Excel
	Hierarchical	JSON, XML	Nested door objects inside a room
	Graph-based	RDF, GraphDB	Relationships between building elements

מודל הנתונים הוא מבנה לוגי, המתאר כיצד הנתונים מאורגנים, מאוחסנים ומעובדים במערכת.

הבחירה בפורמט המתאים תלויה במשימות:

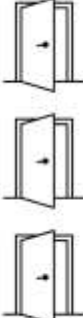
■ (מסדי נתונים עמודתיים, RDBMS, CSV, SQL) אם מהירות הסינון והאנליזה חשובה – טבלאות רלציוניות יתאימו.

■ XML או JSON-אם יש צורך בגמישות במבנה – עדיף להשתמש ב

■. אם לנתונים יש קשרים מורכבים – מסדי נתונים גרפיים מספקים בהירות ויכולת הרחבה

כל ישות (למשל, דלת) (מיוצגת על ידי שורה, ותכונותיה – על ידי (RDBMS) במסדי נתונים רלציוניים קלאסיים Room-גובה, רוחב, עמידות באש ו ID, עמודות הטבלה. לדוגמה, טבלת פריטים מקטגוריית "דלתות" עשויה לכלול שדות המצביע על החדר, ID.

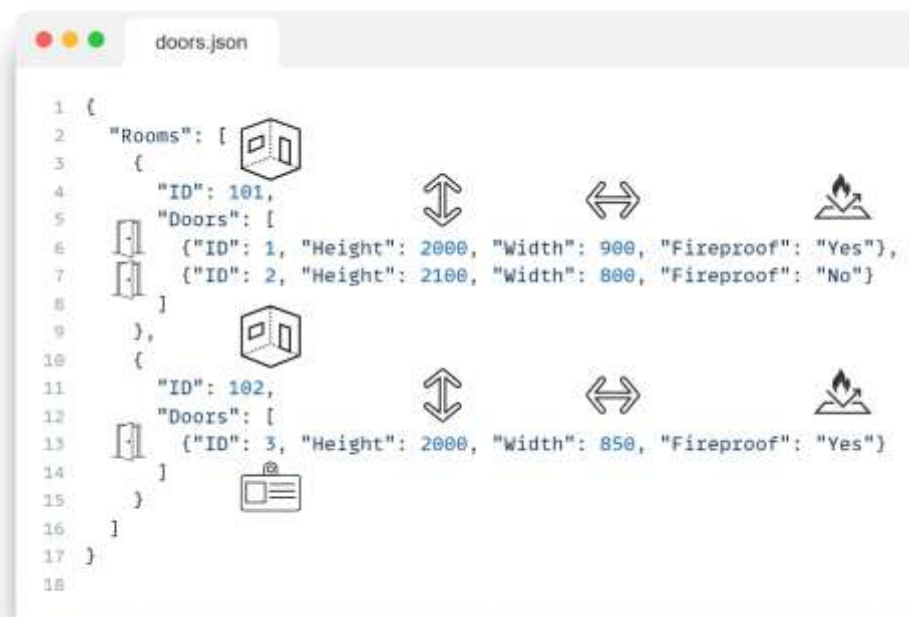
הקשרים נוצרים בצורה של טבלאות, כאשר כל רשומה מייצגת אובייקט (RDBMS) במסדי נתונים רלציוניים קלאסיים והעמודות – את הפרמטרים שלו. בפורמט הטבלאי, הנתונים על דלתות בפרויקט נראים כך, כאשר כל שורה מייצגת "Room ID" פריט נפרד - דלת עם מזהה ייחודי ואטריבוטים, והקשר עם החדר מתבצע דרך הפרמטר



					
	Door ID	Room ID	Height (mm)	Width (mm)	Fireproof
	ID1001	101	2000	900	Yes
	ID1002	101	2100	800	No
	ID1003	102	2000	850	Yes

מידע על שלושה פריטים מקטגוריית "דלתות" בפרויקט בצורה טבלאית ומסודרת.

הנתונים מאוחסנים בצורה היררכית או מקוננת, שבה המרכיבים יכולים, XML או JSON בפורמטים חצי-מסודרים, כמו להכיל אובייקטים אחרים, ומבנה שלהם יכול להשתנות. זה מאפשר לדגם קשרים מורכבים בין המרכיבים. מידע דומה מוצג כך, שהם הופכים לאובייקטים מקוננים (JSON) על דלתות בפרויקט, שנרשם בצורה מסודרת, בפורמט חצי-מסודר. --בתוך החדרים, מה שמשקף בצורה הגיונית את ההיררכיה



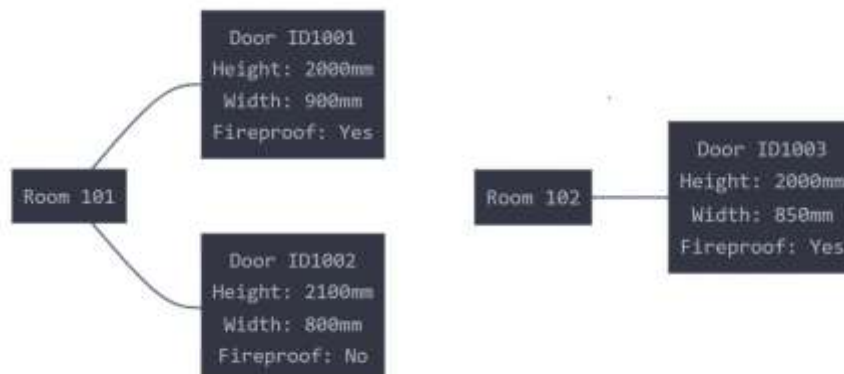
JSON. מידע על פריטי קטגוריית "דלתות" בפרויקט בפורמט

במודל גרפי, הנתונים מוצגים בצורת צמתים (קודקודים) (וקשרים) קצוות (ביניהם). זה מאפשר להציג בצורה ברורה את

הקשרים המורכבים בין אובייקטים ואטריבוטים שלהם. במקרה של נתונים על דלתות וחדרים בפרויקט, הייצוג הגרפי נראה כך:

(ID1001, ID1002, ID1003) (הצמתים) קודקודים (מייצגים ישויות עיקריות: חדרים) 101, חדר 102 (ודלתות)

■ (הקצוות) קשרים (מראים את הקשרים בין ישויות אלו, לדוגמה, שייכות דלת לחדר מסוים)
 ■ האטריבוטים מחוברים לצמתים ומכילים תכונות של ישויות (גובה, רוחב, עמידות באש לדלתות)

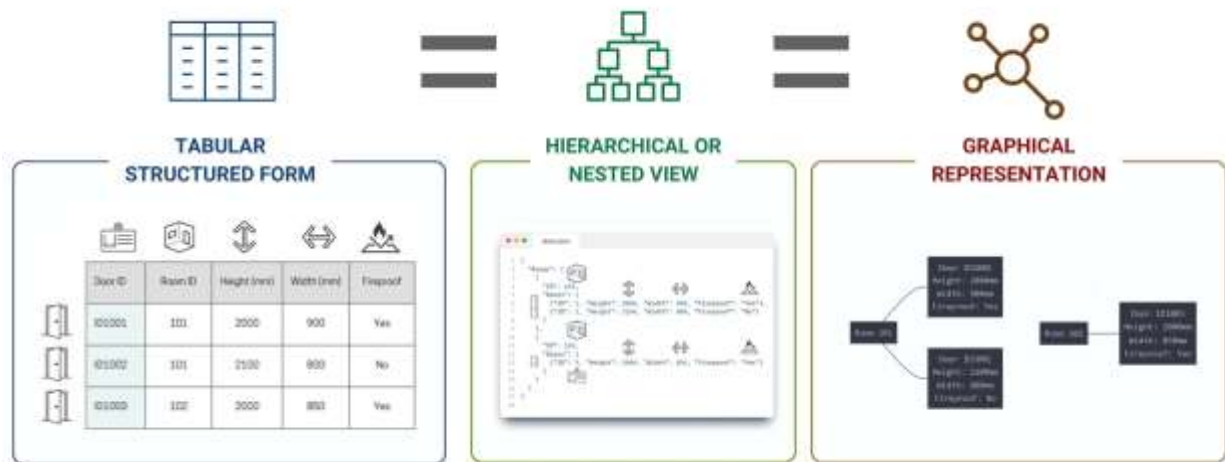


מידע על ישויות דלתות הפרויקט בייצוג גרפי 3.29. ריש.

במודל הגרפי של נתוני תיאורי דלתות, כל חדר וכל דלת הם צמתים נפרדים. הדלתות מחוברות לחדרים דרך קצוות, המצביעים על שייכות הדלת לחדר מסוים. בעת זאת, האטריבוטים של הדלתות (גובה, רוחב, עמידות באש (נשמרים כתכונות של הצמתים המתאימים. נדון יותר על פורמטים גרפיים וכיצד סמנטיקה גרפית הופיעה בתעשיית הבנייה בפרק "הופעת הסמנטיקה והאונטולוגיה בבנייה".

מסדי נתונים גרפיים יעילים במקרים שבהם חשובים לא רק הנתונים עצמם, אלא גם הקשרים ביניהם, לדוגמה, במערכות המלצה, מערכות ניתוב או בעת מודל קשרים מורכבים בפרויקטים לניהול נכסים. הפורמט הגרפי מפשט את יצירת הקשרים החדשים, ומאפשר להוסיף לגרף סוגי נתונים חדשים מבלי לשנות את מבנה האחסון. עם זאת, בהשוואה לטבלאות יחסיות ולפורמטים מובנים, בגרף אין קשר נוסף בין הנתונים - המרת נתונים מבסיסים דו-ממדיים לגרף אינה מגדילה את מספר הקשרים ואינה מאפשרת לקבל מידע חדש

הצורה והסכימה של הנתונים צריכות להתאים למקרה השימוש הספציפי ולבעיות הנפתרות. כדי לעבוד ביעילות בתהליכי עסקים, חשוב להשתמש בכלים ובמודלים של נתונים המסייעים להשיג תוצאה במהירות ובפשטות המרבית.



אותה מידע על רכיבי הפרויקט יכול להישמר בפורמטים שונים באמצעות מודלים שונים של נתונים **3.210**. ריש.

כיום, רוב החברות הגדולות מתמודדות עם בעיית מורכבות יתר של נתונים. כל אחד ממאות או אלפי היישומים משתמש במודל נתונים משלו, מה שיוצר מורכבות מיותרת - מודל נפרד לעיתים קרובות מורכב בעשרות מונים מהנדרש, והצבר של כל המודלים - באלפי מונים. מורכבות מיותרת זו מקשה מאוד על העבודה הן על המפתחים והן על המשתמשים הסופיים.

מורכבות זו מטילה מגבלות חמורות על הפיתוח והתחזוקה של מערכות החברה. כל אלמנט חדש במודל דורש קוד נוסף, הטמעת לוגיקות חדשות, בדיקות קפדניות והתאמה לפתרונות הקיימים. כל זה מגדיל את העלויות ומאט את עבודת הצוות האוטומטי בחברה, והופך אפילו משימות פשוטות לתהליכים יקרים ומורכבים.

ההסתבכות נוגעת לכל הרמות של ארכיטקטורת הנתונים. בבסיסי נתונים רלציוניים זה מתבטא בהתרבות מספר הטבלאות והעמודות, לעיתים קרובות בצורה מיותרת. במערכות מונחות אובייקטים, המורכבות גוברת בזכות מספר רב של מחלקות הגודש מתבטא דרך מבנים מקוננים מסובכים, מפתחות ייחודיים, JSON או XML ותכונות מקושרות. בפורמטים כמו וסכמות לא עקביות.

המורכבות המיותרת של מודלי הנתונים מקשה על המערכות לא רק להפוך לפחות יעילות, אלא גם עלולה לLLM. להקשות על ההבנה של המשתמשים הסופיים ובסופו של דבר על מודלים לשוניים גדולים וסוכנים בעיה זו של הבנה ומורכבות מודלי הנתונים ועיבודיהם מעלה את השאלה: כיצד ניתן להפוך את הנתונים לפשוטים מספיק לשימוש, כך שהם באמת יתחילו להניב תועלת במהירות.

אפילו בבחירה נכונה של מודלי נתונים, התועלת המעשית שלהם פוחתת באופן דרסטי אם הגישה לנתונים מוגבלת פורמטים קנייניים ופלטפורמות סגורות מקשות על האינטגרציה, מסבכות את האוטומציה ומונעות מהחברה שליטה על המידע שלה, מה שיוצר לא רק סילו של נתונים חדשים, אלא סילו נעול, שהמפתח אליו ניתן להשיג רק באישור של הספק. כדי להבין את היקף הבעיה, חשוב לבחון כיצד סגירות המערכות משפיעה על תהליכים דיגיטליים בתחום הבנייה.

פורמטים קנייניים והשפעתם על תהליכים דיגיטליים

אחת הבעיות המרכזיות בהן נתקלות חברות הבנייה במהלך הדיגיטציה היא הגישה המוגבלת לנתונים. זה מקשה על אינטגרציה של מערכות, מפחית את איכות המידע ומסבך את הארגון של תהליכים יעילים. בבסיס הקשיים הללו לעיתים קרובות עומד השימוש בפורמטים קנייניים ובפתרונות תוכנה סגורים.

למרבה הצער, עד כה רבות מהתוכנות בשימוש בתעשיית הבנייה מאפשרות למשתמש לשמור נתונים אך ורק בפורמטים קנייניים או באחסון בענן, שהגישה אליו אפשרית רק דרך ממשקים מוגבלים מאוד. במקרים רבים, פתרונות כאלה נבנים בהתאם למערכות סגורות עוד יותר של ספקים גדולים יותר. כתוצאה מכך, אפילו המפתחים שרוצים להציע ארכיטקטורות פתוחות יותר, נאלצים לעמוד בכללים שמכתיבים הספקים הגדולים.

בעוד שמערכות ניהול נתוני הבנייה המודרניות תומכות יותר ויותר בפורמטים ובסטנדרטים פתוחים (איור 3.15), הקשורות אליהם נשארות CAD-ERP ו-CAFM-ERP וכן מערכות CAD- (BIM) בסיסי הנתונים של כלי, (3.211).-



האופי הסגור והקנייני של הנתונים יוצר מחסומים לאינטגרציה ולגישה לנתונים 3.211 איור.

הסגירות והמונופוליציה של הפורמטים והפרוטוקולים אינן בעיה רק של תעשיית הבנייה. במגוון תחומים כלכליים, קיומם של (3.212) איור) המאבק נגד הסטנדרטים הסגורים והגישה המוגבלת לנתונים החל עם האטת החדשנות מחסומים מלאכותיים לכניסת שחקנים חדשים והעמקת התלות בספקים הגדולים. על רקע הצמיחה המהירה של חשיבות הנתונים, רשויות ההגבלים העסקיים פשוט לא מצליחות להגיב לאתגרים הקשורים לשווקים הדיגיטליים החדשים, ובסופו של דבר הפורמטים הסגורים והגישה הסגורה לנתונים הופכים ל"גבולות" דיגיטליים, המעכבים את תנועת המידע ואת הצמיחה.

אם המכונות מייצרות את כל מה שאנחנו צריכים, אז המצב שלנו יהיה תלוי כיצד המשאבים הללו מחולקים. כל אחד יוכל ליהנות מחיים בשפע רק אם העושר המיוצר על ידי המכונות יהפוך לרכוש ציבורי. אחרת, רוב האנשים בסופו של דבר יחיו בעוני נורא, אם בעלי המכונות יצליחו לוביסטית למנוע את חלוקת העושר. עד כה, נראה כי הכל מתנהל לפי האפשרות השנייה, הטכנולוגיות מובילות לעלייה באי שוויון

סטיבן הוקינג, אסטרופיזיקאי, 2015 –

Monopolies or tight control over critical data formats

Telecommunications:
Proprietary Protocols

1970s-1980s

Computing Industry:
Open Source Movement

1980s

Document Formats:
PDFs and DOCs

Late 1980s to 1990s

Web Browsing:
Browser Wars

Mid-1990s to early 2000s

Media:
Audio and Video Codecs

1990s-2000s



אי.פ. 3.212 בעלות מונופוליסטית על פורמטים מרכזיים ופרוטוקולי נתונים אינה בעיה ייחודית לתעשיית הבנייה.

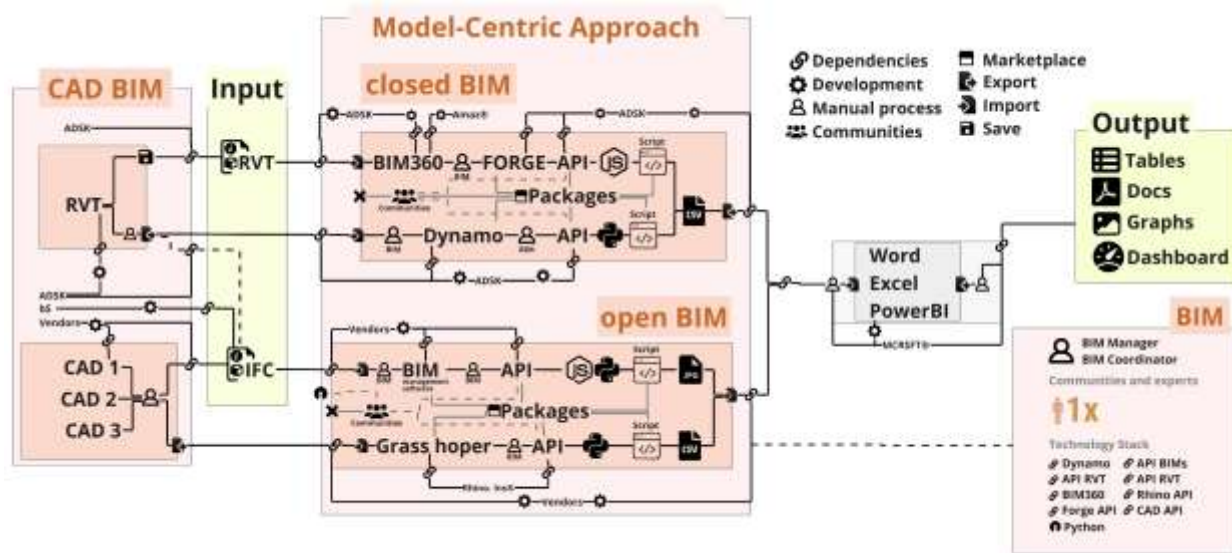
ומפתחים, המייצרים IT בסופו של דבר, בשל הגישה הסגורה לבסיסי נתונים של תוכנה, מנהלי נתונים, אנליסטים, אנשי בתעשיית הבנייה אפליקציות לגישה לנתונים, עיבוד ואוטומציה, מתמודדים כיום עם תלות רבה בספקי תוכנה. התלות - מיוחדים וכלים ותוכנה מיוחדים API הזו, בצורת רמות גישה נוספות, דורשת יצירת פתרונות עם חיבורים

(,ממשק תכנות יישומים) הוא ממשק פורמלי, באמצעותו תוכנה אחת יכולה לתקשר עם תוכנה אחרת API מתאר אילו בקשות יכולה API-נתונים ופונקציות מבלי צורך לגשת לקוד המקור. ה. exchanging מערכת חיצונית לבצע, באיזה פורמט הן צריכות להיות, ואילו תגובות היא תקבל. זהו "חווה" סטנדרטי בין מודולים תוכנותיים

כמות גדולה של תלות בפתרונות סגורים גורמת לכך שכל הארכיטקטורה של הקוד והלוגיקה של תהליכי העסקים בחברה. הופכת ל"ארכיטקטורת ספגטי" של כלים התלויים במדיניות הספק של תוכנה לגבי מתן גישה איכותית לנתונים

התלות בפתרונות ופלטפורמות סגורות לא רק מפסידה גמישות, אלא גם מביאה לסיכונים עסקיים ממשיים. שינוי בתנאי כל אלה יכולים לחסום תהליכים קריטיים. פתאום – API-הרישוי, סגירת גישה לנתונים, שינוי פורמטים או מבנה ה מתברר כי עדכון של טבלה אחת דורש שיפוט של בלוק שלם של אינטגרציות וחיבורים, וכל עדכון רחב היקף של תוכנה

- של הספק הופך לאיום פוטנציאלי על יציבות כל המערכת של החברה API-או ה



יוצרת מכשולים לאינטגרציה של נתונים במערכת האקולוגית CAD אי.פ. 3.213. דוגמה לכמות גדולה של תלות בעיבוד נתוני של חברות הבנייה

מפתחים ואדריכלי מערכות בתנאים כאלה נאלצים לעבוד לא על מנת להקדים את זמנם, אלא על מנת לשרוד. במקום ליישם פתרונות חדשים – הם מתאימים את עצמם. במקום להתפתח – הם מנסים לשמור על תאימות. במקום לאוטומט ולהאיץ תהליכים, הם מבזבזים זמן בלימוד ממשקים סגורים נוספים, תיעוד ובנייה מחדש אינסופית של הקוד API

עבודה עם פורמטים וסיסטמות סגורות אינה רק בעיה טכנית - היא מגבלה אסטרטגית. למרות ההזדמנויות הברורות וניתוח חיזוי, רבות מהחברות אינן LLM, שמספקות הטכנולוגיות המודרניות של אוטומציה, אינטליגנציה מלאכותית מצליחות לממש את הפוטנציאל שלהן במלואו. והמחסומים שהוקמו על ידי פורמטים קנייניים מונעים מהעסקים גישה לנתונים שלהם. ובזה, אולי, טמונה האירוניה המרכזית של טרנספורמציה דיגיטלית בתחום הבנייה

שקיפות נתונים ופתיחות מערכות אינן מותרות, אלא תנאי הכרחי למהירות וליעילות. בהיעדר פתיחות תהליכי העסקים מתמלאים בבירוקרטיה מיותרת, בשרשרות אישור מרובות ובתלות גוברת בעקרון קבלת החלטות על בסיס דעתו של האדם המרוויח ביותר - HiPPO

עם זאת, באופן מתהווה שינוי פרדיגמה. למרות הדומיננטיות של פתרונות קנייניים, יותר ויותר חברות מודעות למגבלות הארכיטקטורות שנוצרו ברוח המהפכה התעשייתית הרביעית. כיום, הווקטור זז לעבר עקרונות המהפכה החמישית ואינטרופראביליות אמיתית בין מערכות (API) שבהם הנתונים נמצאים במרכז כמשאב אסטרטגי, ממשקים פתוחים המעבר הזה מסמן את היציאה מהאקוסיסטמות הסגורות לעבר ארכיטקטורות דיגיטליות גמישות ומודולריות, שבהן פורמטים פתוחים, סטנדרטים ושקיפות בהחלפת נתונים ממשקים תפקיד מרכזי

פורמטים פתוחים משנים את הגישה לדיגיטציה

תחום הבנייה התמודד עם בעיית הסגירות והקנייניות של הנתונים אחת האחרונות. בניגוד לתחומים אחרים בכלכלה הדיגיטליזציה כאן התפתחה לאט. הסיבות לכך כוללות את השמרנות המסורתית של התחום, את הדומיננטיות של פתרונות מקומיים מבודדים ואת השתרשות העמוקה של ניהול מסמכים פיזיים. במשך עשורים, תהליכים מרכזיים בבנייה התבססו על שרטוטים פיזיים, שיחות טלפון ומסדי נתונים שאינם מסונכרנים. בהקשר זה, פורמטים סגורים נתפסו במשך זמן רב כנורמה ולא כמכשול.

הניסיון מתחומים אחרים מראה: הסרת מחסומים הקשורים לנתונים סגורים מובילה לעלייה בחדשנות, להאצת ההתפתחות - ולצמיחה בתחרות. במדע, החלפת נתונים פתוחים מאפשרת להאיץ גילויים ולפתח שיתוף פעולה בינלאומי. ברפואה לשפר את היעילות של אבחון וטיפול. בהנדסת תוכנה - ליצור אקוסטמות של יצירה משותפת ושיפור מהיר של מוצרים.

לפי דוח מקינזי "נתונים פתוחים: שחרור חדשנות ופרודוקטיביות באמצעות זרמי מידע" משנת 2013, נתונים פתוחים יכולים לשחרר בין 3 ל-5 טריליון דולר בשנה בשבעה תחומים מרכזיים בכלכלה, כולל בנייה, תחבורה, בריאות ואנרגיה. לפי אותו מחקר, אקוסטמות נתונים מבוזרות מאפשרות לחברות בנייה גדולות ולקבלנים לצמצם את העלויות של פיתוח ותחזוקת תוכנה, ולהאיץ את אימוץ הטכנולוגיות הדיגיטליות.

המעבר לארכיטקטורות פתוחות, שהחל מזמן בתחומים אחרים בכלכלה, הולך ומקיף גם את ענף הבנייה. חברות גדולות ולקוחות ממשלתיים, ובמיוחד ארגונים פיננסיים, השולטים בהשקעות בפרויקטים בנייה, מציבים יותר ויותר דרישות לשימוש בנתונים פתוחים ולהבטחת גישה לקוד המקור של חישובים, חישובים ויישומים. המפתחים כבר לא צריכים רק ליצור פתרונות דיגיטליים ולהציג את המספרים הסופיים של הפרויקט - מהם מצפים שקיפות, שחזור ועצמאות מספקי יישומים חיצוניים.

השימוש בפתרונות עם קוד פתוח מספק ללקוחות ביטחון בכך שאפילו אם מפתחים חיצוניים מפסיקים לשתף פעולה או עוזבים את הפרויקט, זה לא ישפיע על האפשרות להמשיך פיתוח הכלים והמערכות. אחד היתרונות המרכזיים של נתונים פתוחים הוא היכולת שלהם להפסיק את התלות של מפתחי היישומים בפלטפורמות ספציפיות לגישה לנתונים.

אם חברה לא יכולה להפסיק לחלוטין להשתמש בפתרונות קנייניים, פשרה אפשרית היא השימוש בשיטות הנדסה הפוכה שיטות אלו, שהן חוקיות ומבוססות טכנית, מאפשרות להמיר פורמטים סגורים לפורמטים נגשים, מובנים ומתאימים לשילוב. זה חשוב במיוחד במקרים שבהם יש צורך להתחבר למערכות מיושנות או להעביר מידע מנוף תוכנה אחד לאחר.

אחד הדוגמאות הבולטות בהיסטוריה של המעבר לפורמטים פתוחים ושימוש בהנדסה הפוכה (פריצה חוקית הנמצא בשימוש נרחב, DWG של מערכות קנייניות (בבנייה הוא הסיפור של המאבק לפתיחת פורמט בשנת 1998, בתגובה למונופול של אחד מספקי התוכנה, 15 ספקי. (CAD) במערכות תכנון ממוחשבות שמטרתה הייתה לספק למפתחים כלים חינוניים, "Open DWG" אחרים הקימו ברית חדשה בשם CAD (הסטנדרט דה-פקטו להעברת שרטוטים) (מבלי הצורך להשתמש DWG ועצמאיים לעבודה עם פורמט סגורים. אירוע זה היה נקודת מפנה, שאפשרה לעשרות אלפי חברות לקבל גישה API-בתוכנה קניינית או ב פופולרי, מאז סוף שנות ה-80 ועד היום, וליצור פתרונות תואמים CAD חופשית לפורמט הסגור של פתרון שנוצר לראשונה כבר בשנת 1996, SDK "Open DWG"-כיום, ה. CAD-שהקנו יתרון תחרותי בשוק ה מחוץ ליישום הרשמי, DWG נמצא בשימוש כמעט בכל הפתרונות שבהם ניתן לייבא, לערוך ולייצא פורמט DWG של מפתח פורמט

טרנספורמציות דומות מתרחשות גם אצל ענקיות טכנולוגיה אחרות. מיקרוסופט, שהייתה פעם סמל לגישה קניינית ורכשה את Azure החלה להשתמש בלינוקס בתשתית שירותי הענן, NET Framework. פתחה את קוד המקור של (שהייתה פייסבוק) שחררה מודלים של אינטליגנציה Meta כדי לחזק את מעמדה בקהילת הקוד הפתוח. חברת GitHub המנכ"ל מארק AI. כדי לקדם חדשנות ושיתוף פעולה בתחום פיתוח סוכני Llama, מלאכותית עם קוד פתוח, כמו סדרת צוקרברג סבור כי פלטפורמות עם קוד פתוח יובילו את ההתקדמות הטכנולוגית בעשור הקרוב

קוד פתוח הוא מודל לפיתוח והפצת תוכנה, שבו קוד המקור פתוח לשימוש חופשי, ללימוד, לשינוי ולהפצה.

נתונים פתוחים ופתרונות בקוד פתוח הופכים לא רק לטרנד, אלא לבסיס של עמידות דיגיטלית. הם מעניקים לחברות גמישות, עמידות לשינויים, שליטה על הפתרונות שלהן ואפשרות להרחיב תהליכים דיגיטליים מבלי להיות תלויים במדיניות הספקים. ומה שחשוב לא פחות, הם מחזירים לעסקים שליטה על המשאב היקר ביותר של המאה ה-21: הנתונים שלהם.

שינוי פרדיגמה: קוד פתוח כסוף עידן הדומיננטיות של ספקי התוכנה

תעשיית הבנייה עומדת בפני שינוי שלא ניתן למנף בדרכים המסורתיות. הקונספט של המעבר לשימוש בנתונים, גישה ממוקדת נתונים ושימוש בכלים בקוד פתוח מביאים לרה-חשיבה של כללי המשחק שעליהם מתבססים ענקי תוכנת השוק.

בניגוד לשינויים טכנולוגיים קודמים, המעבר הזה לא יוקדם על ידי ספקים. שינוי הפרדיגמה מאיים על מודלים עסקיים מסורתיים המבוססים על רישיונות, מנויים וייעוץ. המציאות החדשה לא מציעה מוצר. מוכן "בקופסה" - או מנוי בתשלום - היא דורשת שינוי בתהליכים ובמחשבה.

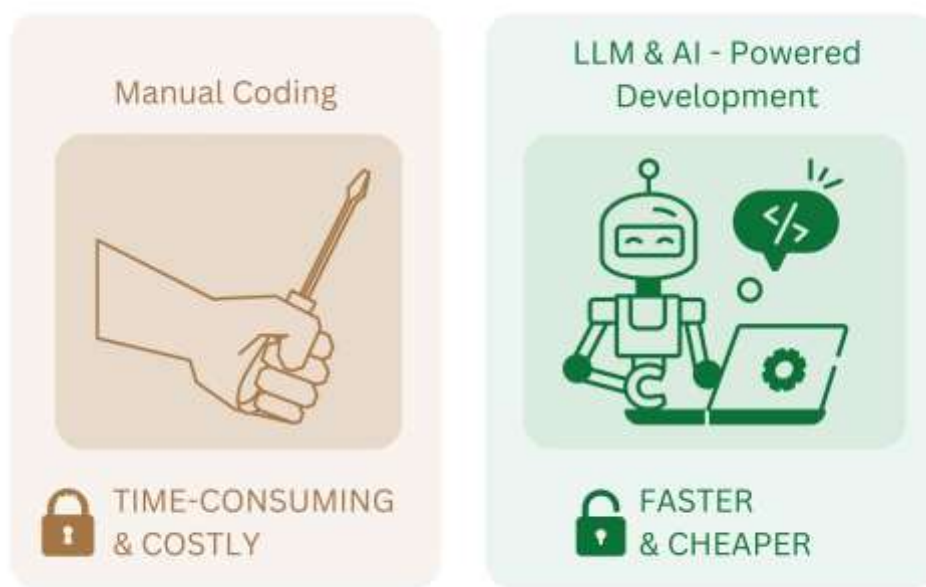
לניהול ופיתוח פתרונות ממוקדי נתונים, המבוססים על טכנולוגיות פתוחות, חברות ידרשו לבחון מחדש את התהליכים הפנימיים. אנשי מקצוע ממחלקות שונות לא רק יצטרכו לשתף פעולה, אלא גם לרה-חשוב את הגישות לעבודה משותפת.

נתפזרם הפרדיגמה החדשה , , (LLM). 2024 25% Google . LLM 80%
20%

McKinsey 2020 , CPU GPU – Open Source- .
 , McKinsey, PwC Deloitte , Open Source .

PwC Open Source Monitor 2019, 69% 100 Open Source . OSS : 71%
 200–499 , 78% 500–1999 86% 2000 . Synopsys OSSRA 2023 , 96% .

העתיד של תפקיד המפתח הוא לא כתיבת קוד ידנית, אלא תכנון מודלים של נתונים, ארכיטקטורת זרמים וניהול סוכני בינה מלאכותית המייצרים חישובים נדרשים לפי בקשה. ממשיקי המשתמש יהפכו למינימליסטיים, והאינטראקציה תהיה מגמות מודרניות. (3.214 איור) דיאלוגית. תכנות קלאסי יפנה מקום לתכנון ברמה גבוהה ולתיאום פתרונות דיגיטליים יאפשרו להפחית – (7.44 איור) LLM ואקוסיסטמות הנתמכות על ידי (7.46 איור) low-code כמו פלטפורמות – IT. באופן משמעותי את העלויות של פיתוח ותחזוקה של מערכות



אם היום אפליקציות נבנות ידנית על ידי מפתחים, בעתיד חלק משמעותי מהקוד ייווצר על ידי פתרונות מבוססי בינה 3.214 איור LLM-מלאכותית ו

המעבר הזה לא יהיה דומה לקודמים, וספקי תוכנה גדולים ככל הנראה לא יהיו הקטליזטורים שלו.

מחקר של בית הספר לעסקים של הרווארד "ערך התוכנה עם קוד פתוח" לשנת 2024 [40], [מערך את העלות הכוללת של תוכנה עם קוד פתוח משתי נקודות מבט. מצד אחד, אם נחשב כמה כסף היה דרוש כדי ליצור את כל הפתרונות הקיימים עם קוד פתוח מאפס, הסכום הזה היה כ-4.15 מיליארד דולר. מצד שני אם נניח שכל חברה מפתחת את המקבילות שלה לפתרונות קוד פתוח בעצמה) (מה שקורה בכל מקום), (מבלי גישה לכלים הקיימים, העלויות הכוללות של העסקים לכך היו מגיעות לסכום עצום של 8.8 טריליון דולר – זו עלות הביקוש –

לא קשה לנחש שאין ספקי תוכנה גדולים המעוניינים לצמצם את שוק התוכנה עם פוטנציאל עלות של 8.8 טריליון דולר ל-4.15 מיליארד בלבד. זה היה אומר צמצום הביקוש ביותר מ-2,000 פעמים. שינוי כזה פשוט לא יהיה רווחי עבור ספקים, שהמודלים העסקיים שלהם מבוססים על שמירה על תלות לקוחות בפתרונות סגורים במשך שנים רבות. לכן חברות הממתינות לכך שמישהו יציע להן פתרון נוח ופתוח מוכן לשימוש, "עשויות להתאכזב – ספקים כאלה פשוט לא יופיעו"

המעבר לארכיטקטורה דיגיטלית פתוחה לא אומר צמצום משרות או הכנסות. להפך, הוא יוצר תנאים למודלים עסקיים גמישים ואדפטיביים, שעם הזמן עשויים להדיח את השוק המסורתי של רישיונות ותוכנה באריזות

במקום מכירת רישיונות – שירותים, במקום פורמטים סגורים – פלטפורמות פתוחות, במקום תלות בספק – עצמאות. ואפשרות לבנות פתרונות לפי הצרכים האמיתיים. אלה שבעבר פשוט השתמשו בכלים, יוכלו להפוך לשותפים שלהם ואילו אלה שמסוגלים לעבוד עם נתונים, מודלים, תרחישים ולוגיקה – ימצאו את עצמם במרכז הכלכלה הדיגיטלית החדשה של הענף. נדון בשינויים אלה ובתפקידים החדשים, המודלים העסקיים ופורמטי שיתוף הפעולה המתהווים סביב נתונים פתוחים בחלק העשירי והאחרון של הספר

מיושנים API פתרונות המבוססים על נתונים פתוחים וקוד פתוח יאפשרו לחברות להתמקד לא במאבק עם ואינטגרציה של מערכות סגורות, אלא ביעילות של תהליכי העסקים. המעבר המודע לארכיטקטורה פתוחה מעניק אפשרות להעלות באופן משמעותי את הפרודוקטיביות ולהפחית את התלות בספקים

המעבר למציאות חדשה אינו רק שינוי בגישות לפיתוח תוכנה, אלא גם רהור מחדש של העיקרון עצמו של עבודה עם נתונים. במרכז השינוי הזה לא נמצא הקוד, אלא המידע: המבנה שלו, זמינותו ויכולת הפרשנות שלו. וכאן נתונים פתוחים חלק בלתי נפרד מהארכיטקטורה הדיגיטלית החדשה becoming, ומסודרים יוצאים לחזית

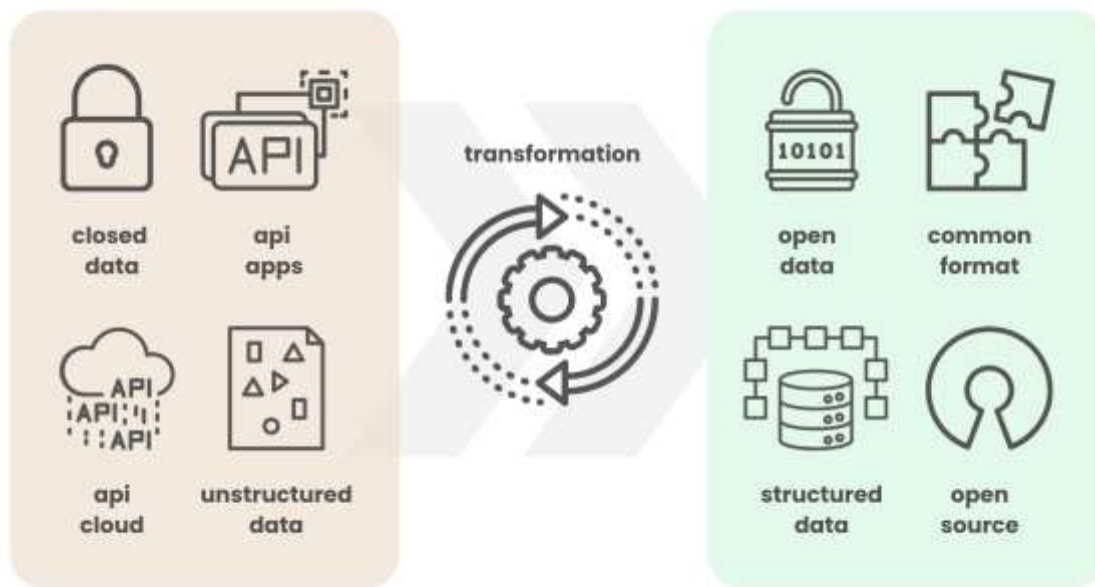
נתונים פתוחים מובנים: יסוד של טרנספורמציה דיגיטלית

אם בעשורים הקודמים יציבות העסק הוגדרה במידה רבה על ידי הבחירה בפתרונות תוכנה ותלות בספקים מסוימים, אז בעידן הכלכלה הדיגיטלית המודרנית, גורם המפתח הוא איכות הנתונים ויכולת העבודה היעילה איתם. קוד פתוח הוא חלק חשוב מהפרדיגמה הטכנולוגית החדשה, אך הפוטנציאל שלו מתגלה באמת רק כאשר יש נתונים ברורים, מאורגנים וניתנים לקריאה על ידי מכונה. בין כל סוגי מודלי הנתונים, נתונים פתוחים ומסודרים הופכים לאבן הפינה של טרנספורמציה דיגיטלית בת קיימא

היתרון המרכזי של נתונים פתוחים ומסודרים הוא הפרשנות החד משמעית והיכולת לעיבוד אוטומטי. זה מאפשר להעלות באופן משמעותי את היעילות הן ברמת פעולות בודדות והן ברמת הארגון כולו

לניהול העברת נתונים IT לפי דוח של דלויט "תהליך העברת נתונים במהלך שינויים ארגוניים", "עבודה עם מסודרים היא קריטית. לפי דוח ממשלת בריטניה "אנליטיקה של נתונים ובינה מלאכותית במימוש פרויקטים ממשלתיים" (2024), (הסרת מחסומים בהחלפת נתונים בין פרויקטים וארגונים שונים היא גורם מפתח להעלאת היעילות בניהול פרויקטים. במסמך מודגש כי סטנדרטיזציה של פורמטי נתונים והטמעת עקרונות של נתונים פתוחים מאפשרים להימנע מהכפלת מידע, למזער אובדן זמן ולהגביר את דיוק התחזיות

עבור תעשיית הבנייה, שבה קיימת מסורתית רמה גבוהה של פיצול ומגוון פורמטים, תהליך הסידור-אחידות ונתונים פתוחים ומסודרים משחקים תפקיד מכריע ביצירת תהליכים מסודרים ומנוהלים. הם מאפשרים למשתתפי הפרויקט להתמקד בהגברת הפרודוקטיביות ולא בפתרון בעיות טכניות הקשורות לחוסר תאימות של פלטפורמות סגורות, מודלי נתונים ופורמטים.



נתונים פתוחים ומסודרים מפחיתים את התלות בפתרונות תוכנה ופלטפורמות ומאיצים חדשנות.

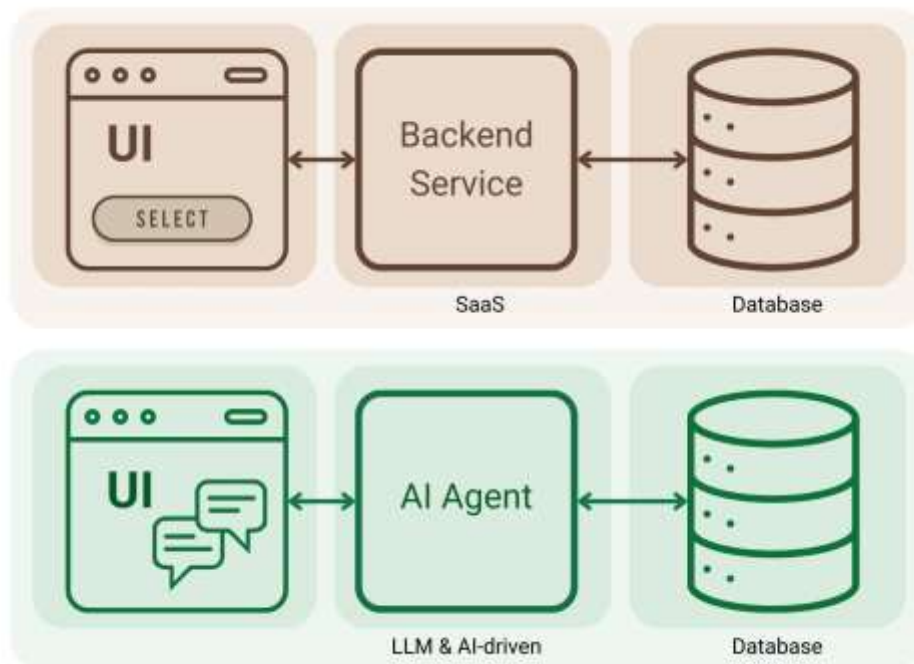
הטכנולוגיות המודרניות, הכלים שנבחנו בהמשך הספר, מאפשרים לא רק לאסוף מידע, אלא גם לנקות אותו אוטומטית, להסיר הכפלות, לתקן שגיאות, לנרמל ערכים. זה אומר שהאנליסטים והמהנדסים עובדים לא עם מסמכים מנותקים אלא עם בסיס ידע מאורגן, המתאים לניתוח, אוטומציה וקבלת החלטות.

עשה את זה פשוט ככל האפשר, אבל לא פשוט יותר.

אלברט איינשטיין, פיזיקאי תאורטי (השתייכות הציטוט שנויה במחלוקת) –

כיום, רוב הממשקים המשתמשים לעבודה עם נתונים יכולים להיווצר אוטומטית - מבלי הצורך לכתוב קוד ידנית עבור כל מקרה עסקי. לשם כך, יש צורך בשכבת תשתית, שמבינה את מבנה הנתונים, המודל והלוגיקה ללא הנחיות נוספות רק נתונים מובנים מאפשרים גישה כזו: טפסים, טבלאות, מסננים ותצוגות יכולים להיווצר אוטומטית עם מינימום עלויות - תכנות.

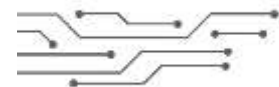
הממשקים החשובים ביותר, הקריטיים למשתמש, עדיין עשויים לדרוש שיפוט ידני. אך ברוב המקרים - שזה בין 50 ל-90 אחוזים מהתסריטים העבודה - די ביצירת יישומים וחשובים אוטומטיים מבלי להשתמש ביישומים מיוחדים לכך 90 מה שמפחית באופן משמעותי את העלויות לפיתוח ולתחזוקה, מצמצם את מספר השגיאות ומאיץ את יישום הפתרונות הדיגיטליים.



LLM. - עם **AI** דיאגרמה **3.216** מודלים ארכיטקטוניים לעבודה עם נתונים: ארכיטקטורת יישומים מסורתית ומודל מונחה

המעבר מארכיטקטורות המבוססות על יישומים נפרדים למערכות מנוהלות אינטלקטואלית, המתבססות על מודלים הוא הצעד הבא באבולוציה הדיגיטלית. בארכיטקטורה כזו, נתונים מובנים הופכים לא רק לאובייקט, (LLM) לשוניים המסוגלים לנתח, לפרש ולהמליץ על פעולות בהתאם להקשר, AI, אחסון, אלא גם לבסיס האינטראקציה עם כלים של

בפרקים הבאים נבחן דוגמאות אמיתיות ליישום ארכיטקטורה המבוססת על נתונים מובנים פתוחים, ונראה כיצד מודלים לשוניים מיועדים לפרש, לאמת ולעבד נתונים באופן אוטומטי. מקרים מעשיים אלה יעזרו להבין טוב יותר כיצד הלוגיקה הדיגיטלית החדשה פועלת בפועל - ואילו יתרונות היא מספקת לחברות המוכנות לשינוי



פרק 3.3 מודלים לשוניים גדולים ותפקידם בעיבוד נתונים ובתהליכי עסקים

לאוטומציה של תהליכי עיבוד נתונים: ChatGPT, LLaMa, Mistral, Claude, DeepSeek, QWEN, Grok

הייתה המשך טבעי של התנועה לעבר נתונים פתוחים מובנים ופילוסופיית (LLM) הופעת המודלים הלשוניים הגדולים הקוד הפתוח. כאשר הנתונים הופכים לארגוניים, זמינים וקריאים על ידי מכונה, הצעד הבא הוא כלי המסוגל לתקשר עם מידע זה מבלי הצורך לכתוב קוד מורכב או להחזיק בידע טכני מיוחד.

הם תוצר ישיר של פתיחות: מערכות נתונים פתוחים גדולות, פרסומים ותנועת הקוד הפתוח. ללא LLM ChatGPT מאמרים מדעיים פתוחים, נתוני טקסט זמינים ותרבות הפיתוח המשותף, לא היו קיימים לא הם במובן מסוים "דיסטילט" של הידע הדיגיטלי המצטבר של האנושות, שנאסף LLM. אחרים LLM ולא והוכשר בזכות עקרונות הפתיחות.

ChatGPT® (OpenAI), כמו (LLM - Large Language Models) מודלים לשוניים גדולים מודרניים מספקים LLaMa™ (Meta AI), Mistral DeepSeek™, Grok™ (xAI), Claude™ (Anthropic), QWEN™ למשתמשים אפשרות לנסח בקשות לנתונים בשפה טבעית. זה הופך את העבודה עם מידע לזמינה לא רק למפתחים, אלא גם לאנליסטים, מהנדסים, מעצבים, מנהלים ומומחים אחרים, שהיו בעבר רחוקים מתכנות.

הוא אינטליגנציה מלאכותית, שהוכשרה להבין וליצור טקסט על (LLM Large Language Model) בסיס כמות עצומה של נתונים שנאספו מכל רחבי האינטרנט. היא מסוגלת לנתח הקשר, לענות על שאלות לקיים שיחה, לכתוב טקסטים וליצור קוד תוכנה.

או סקאלה, כמו גם R, SQL, אם בעבר הוויזואליזציה, עיבוד או ניתוח נתונים דרשו ידע בשפת תכנות מיוחדת: פייתון. ורבות אחרות, אז החל משנת 2023 המצב השתנה באופן רדיקלי DB יכולת לעבוד עם ספריות כמו פנדס, פולרס או דאק עכשיו המשתמש יכול פשוט לתאר מה הוא רוצה לקבל - והמודל ייצור את הקוד בעצמו, יבצע אותו, יוציא טבלה או גרף ויסביר את התוצאה. לראשונה מזה עשורים, התפתחות הטכנולוגיות לא הלכה בדרך של סיבוכ, אלא בדרך של רדיקליזציה ופשטות ונגישות.

העיקרון הזה - "עיבוד נתונים במילים" (פרומפטים) - סימן שלב חדש בהתפתחות העבודה עם מידע, למעשה העלה את יצירת הפתרונות לרמה גבוהה יותר של אבסטרקציה. בדומה לכך, כמו שבעבר לא היה צורך להבין את היסודות הטכניים של האינטרנט כדי להפעיל חנויות מקוונות או ליצור אתרים באמצעות וורדפרס, ג'ומלה ומערכות מודולריות אחרות עם קוד פתוח) מחבר הספר עובד עם מערכות כאלה מאז 2005, כולל בתחום הפלטפורמות המקוונות החינוכיות וההנדסיות) מה שהוביל לצמיחה מרשימה של תוכן דיגיטלי ועסקים באינטרנט - היום מהנדסים, אנליסטים ומנהלים יכולים לאוטומט - LLaMA, החזקים - הן חינוכיים והן עם קוד פתוח, כמו LLM את תהליכי העבודה מבלי לדעת שפות תכנות. לכך תורמים ואחרים - שמביאים טכנולוגיות מתקדמות לקהל הרחב ביותר Mistral, Qwen, DeepSeek.

מודלים לשוניים גדולים: איך זה עובד

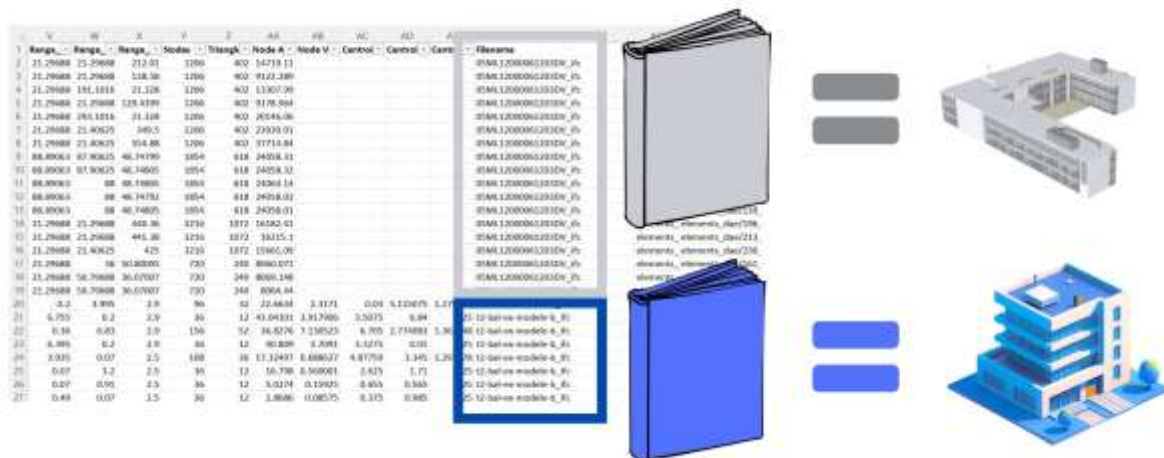
הם רשתות (ChatGPT, LLaMa, Mistral, Claude, DeepSeek, QWEN, Grok) מודלים לשוניים גדולים נזרונים שהוכשרו על כמויות עצומות של נתוני טקסט מהאינטרנט, ספרים, מאמרים ומקורות אחרים. המשימה העיקרית שלהם היא להבין את ההקשר של השפה האנושית וליצור תגובות משמעותיות

המודרניים עומדת ארכיטקטורת הטרנספורמר, שהוצעה על ידי חוקרים מגוגל בשנת 2017. רכיב מרכזי LLM-בבסיס ה המאפשר למודל לקחת בחשבון את הקשרים בין מילים ללא, (attention) בארכיטקטורה זו הוא מנגנון תשומת הלב תלות במיקומן בטקסט

מזכיר מרחוק את הדרך שבה אדם לומד שפה - רק במיליון פעמים יותר LLM תהליך ההכשרה של המודל מנתח מיליארדי דוגמאות לשימוש במילים וביטויים, ומגלה דפוסים במבנה השפה ובלוגיקה של המעברים המשמעותיים. במהלך זאת, כל הטקסט מפורק לטוקנים - יחידות משמעות מינימליות (מילים או חלקים מהן), (אשר לאחר מכן מומרות לווקטורים במרחב רב-ממדי. ייצוגים וקטוריים אלה מאפשרים -למכונה "להבין" את הקשרים הסמויים בין המונחים, ולא רק לפעול על הטקסט כסדרה של תווים

מודלים לשוניים גדולים אינם רק כלים ליצירת טקסט. הם יודעים לזהות משמעות, למצוא קשרים בין מושגים ולעבוד עם נתונים, גם אם הם מוצגים בפורמטים שונים. העיקר הוא שהמידע יהיה מפורק למודלים מובנים ויוצג בצורה של לעבוד LLM-טוקנים, עימם יכול ה

גישה זו יכולה להתבצע גם בפרויקטים בנייה. אם נחשוב על הפרויקט כעל טקסט ייחודי, שבו כל בניין, רכיב או מבנה הם טוקנים, נוכל להתחיל לעבד מידע זה בצורה דומה. פרויקטים בנייה יכולים להתפרק לקטגוריות, פרקים וקבוצות פסקאות, המורכבות מטוקנים מינימליים - רכיבי הפרויקט הבנייתי. על ידי המרת מודלים נתונים לפורמט מובנה, נוכל LLM- גם להמיר נתונים מובנים לבסיסי נתונים וקטוריים, שהם מקור אידיאלי ללמידת מכונה וטכנולוגיות כמו



רכיב בפרויקט בנייה הוא כמו טוקן בטקסט: יחידת המינימום שממנה נוצרים קבוצות (פסקאות) (פרקים) (קטגוריות) של כל הפרויקט -

אם פרויקט הבנייה דיגיטלי ורכיביו מיוצגים כטוקנים או וקטורים, מתאפשרת גישה אליהם לא באמצעות בקשות היכולת להבין את משמעות LLM - פורמליות נוקשות, אלא בשפה טבעית. כאן מתגלה אחד היתרונות המרכזיים של

הבקשה ולחבר אותה לנתונים המתאימים.

כדי לקבל את הנתונים הנדרשים - הוא יכול פשוט, תוך Python-או קוד ב SQL המהנדס כבר לא חייב לכתוב בקשת ומבנה הנתונים, לנסח את המשימה בדרך המוכרת: "מצא את כל המכנים מבטון מזוין עם דרגת בטון LLM הבנת עבודת, וחישב את הנפח הכולל שלהם". המודל מזהה את משמעות הבקשה, ממיר אותה לצורת קריאה על ידי מכונה B30 מעל מוצא את הנתונים) מקבץ וממיר (ומחזיר את התוצאה הסופית

מסמכים, טבלאות, מודלים של פרויקטים מומרות לייצוגים וקטוריים) הטמעה (ונשמרות בבסיס הנתונים. כאשר המשתמש שואל שאלה, הבקשה גם היא מומרת לוקטור, והמערכת מוצאת את הנתונים הקרובים ביותר במשמעותם. זה להסתמך לא רק על הידע שהוכשר, אלא גם על נתוני החברה העדכניים, גם אם הם הופיעו לאחר סיום LLM-מאפשר ל הכשרת המודל

בבנייה הוא היכולת לייצר קוד תוכנה. במקום להעביר את LLM אחד היתרונות החשובים ביותר של המפרט הטכני למתכנת, המומחים יכולים לתאר את המשימה בשפה טבעית, והמודל ייצור את הקוד הנדרש, שניתן להשתמש בו) על ידי העתקתו מהשיחה (ביצירת קוד לאוטומציה של תהליכים. מודלי מאפשרים למומחים ללא ידע מעמיק בתכנות לתרום לאוטומציה ולשיפור תהליכי העסק של LLM החברה.



- מספקים למשתמשים אפשרות לכתוב קוד ולקבל תוצאות מבלי צורך לשלוט בכישורי תכנות LLM

בשנת 2024, שבו השתתפו SAP 300 ומומן על ידי Wakefield Research לפי מחקר שנערך על ידי מנהיגים בכירים בחברות עם הכנסות שנתיות של לפחות 1 מיליארד דולר בארצות הברית: 52% בנוגע לניתוח נתונים והמלצות לקבלת החלטות. 48% נוספים AI מהמנכ"לים הבכירים סומכים על לזיהוי סיכונים שלא נלקחו בחשבון, ו-47% - להציע תוכניות חלופיות. בנוסף, 40% משתמשים ב לפיתוח מוצרים חדשים, תכנון תקציב וביצוע מחקרי שוק. המחקר גם הראה השפעה AI-משתמשים ב על החיים האישיים: 39% מהנשאלים ציינו שיפור באיזון בין עבודה לחיים אישיים, 38% AI חיובית של דיווחו על שיפור בבריאות הנפש, ו-31% - על ירידה ברמות הלחץ

נשארים כלי שחשוב להשתמש בהם בצורה מודעת. כמו כל (LLM) עם כל הכוח שלהן, מודלים לשוניים גדולים

טכנולוגיה, יש להם מגבלות. אחת הבעיות המוכרות ביותר היא מה שנקרא "הזיות" - מקרים שבהם המודל מספק תשובה שנראית סבירה, אך למעשה היא שגויה. לכן, חשוב להבין כיצד המודל פועל: אילו נתונים ודגמים הוא יכול לפרש ללא מוגבל לתאריך ההכשרה LLM שגיאיות, כיצד הוא מפרש בקשות ומאפה הוא שואב מידע. כמו כן, יש לזכור שהידע של שלה, וללא חיבור לנתונים חיצוניים, המודל עשוי לא להתחשב בתקנים, מחירים או טכנולוגיות עדכניות.

אוטונומיית AI הפתרון לבעיות אלו הוא עדכון קבוע של בסיסי נתונים וקטוריים, חיבור למקורות עדכניים ופיתוח סוכני שאינם רק עונים על שאלות, אלא גם משתמשים בנתונים באופן פרואקטיבי לצורך למידה, מנהלים משימות, מזהים סיכונים, מציעים אפשרויות אופטימיזציה ומבקרים את ביצועי הפרויקט.

בבנייה אינו רק חידוש טכנולוגי. זהו שינוי פרדיגמה, הסרת מחסומים בין אנשים LLM המעבר לממשקי לנתונים. זו הזדמנות לעבוד עם מידע בקלות כמו שאנחנו מדברים זה עם זה - ובו בזמן לקבל תוצאות מדויקות, מאומתות ומוכנות לפעולה.

החברות שיתחילו להשתמש בכלים כאלה לפני אחרים, יקבלו יתרון תחרותי משמעותי. זהו גם האצה בעבודה, גם הפחתת עלויות, וגם שיפור איכות הפתרונות הפרויקטיים בזכות גישה מהירה לניתוח נתונים ואפשרות למצוא תשובות לשאלות בענן עשוי להיות כרוך LLM מורכבות במהירות. אך יש לקחת בחשבון גם את סוגיות האבטחה. השימוש בשירותי LLM בסיכוני דליפת נתונים. לכן, יותר ויותר ארגונים מחפשים פתרונות חלופיים המאפשרים להפעיל כלים של בתשתית הפנימית שלהם - באופן מקומי, עם הגנה מלאה ושליטה על המידע.

שימוש במודלים לשוניים גדולים מקומיים עבור נתונים רגישים של החברה

הראשונים בשנת 2022 סימנה שלב חדש בהתפתחות האינטליגנציה המלאכותית. עם זאת LLM-הופעתם של צ'אט מיד לאחר הפצת המודלים הללו עלה השאלה המתבקשת: עד כמה בטוח להעביר נתונים ובקשות הנוגעות לחברה לענן? רוב המודלים הלשוניים בענן שמרו את היסטוריית השיחות והמסמכים שהועלו על השרתים שלהם, ועבור AI. חברות העובדות עם מידע רגיש, זה הפך למחסום משמעותי בדרך ליישום.

(Open Source) אחת מהפתרונות היציבים והלוגיים ביותר לבעיה זו היא הפעלת מודלים פתוחים של החברה. בניגוד לשירותי ענן, המודלים המקומיים פועלים IT-באופן מקומי, בתוך תשתית ה (LLM) ללא חיבור לאינטרנט, אינם מעבירים נתונים לשרתים חיצוניים ומעניקים לחברות שליטה מלאה על המידע.

כיום מבחינת ביצועים דומה למודלים סגורים] כמו [Open Source LLM] המודל הפתוח הטוב ביותר אך בפיגור של כשלוש שנים, [ChatGPT, Claude].

Epoch AI, 2024. בן קוטייה, חוקר בכיר בארגון מחקר לא-למטרות רווח -

LLaMA חברות הטכנולוגיה הגדולות החלו להנגיש את המודלים הלשוניים הגדולים שלהן לשימוש מקומי. סדרת מסין הפכו לדוגמאות למעבר לארכיטקטורה פתוחה DeepSeek והפרויקט המתפתח במהירות Meta הפתוחה של

גם השיקו מודלים חזקים, חופשיים מהמגבלות של פלטפורמות קנייניות. יוזמות אלו לא Mistral ו-Falcon, לצד זאת, רק האיצו את ההתפתחות של אינטליגנציה מלאכותית עולמית, אלא גם העניקו לחברות, שחשוב להן נושא הפרטיות, חלופות ממשיות בנוגע לעצמאות, גמישות ועמידה בתקני אבטחה.

בסביבה העסקית, במיוחד בתעשיית הבנייה, הגנת נתונים היא לא רק עניין של נוחות, אלא גם של עמידה ברגולציות. עבודה עם מסמכי מכרז, אומדנים, שרטוטים והתכתבויות סודיות דורשת שליטה קפדנית. כאן מודלים לשוניים מקומיים מספקים את הביטחון הנדרש שהנתונים יישארו בתוך גבולות החברה.

	Cloud LLMs (OpenAI, Claude)	Local LLMs (DeepSeek, LLaMA)
Data Control	Data is transmitted to third parties	Data remains within the company's network
License	Proprietary, paid	Open-source (Apache 2.0, MIT)
Infrastructure	Requires internet	Operates in an isolated environment
Customization	Limited	Full adaptation to company needs
Cost	Pay-per-token/request	One-time hardware investment + maintenance costs
Scalability	Easily scalable with cloud resources	Scaling requires additional local hardware
Security & Compliance	Risk of data leaks, may not meet strict regulations (GDPR, HIPAA)	Full compliance with internal security policies
Performance & Latency	Faster inference due to cloud infrastructure	Dependent on local hardware, may have higher latency
Integration	API-based integration, requires internet access	Can be tightly integrated with on-premise systems
Updates & Maintenance	Automatically updated by provider	Requires manual updates and model retraining
Energy Consumption	Energy cost is covered by provider	High power consumption for inference and training
Offline Availability	Not available without an internet connection	Works completely offline
Inference Cost	Pay-per-use model (cost scales with usage)	Fixed cost after initial investment

מודלים מקומיים מספקים שליטה מלאה ואבטחה, בעוד שפתרונות מבוססי ענן מציעים אינטגרציה נוחה ועדכונים 3.33. ריש - אוטומטיים.

יתרונות מרכזיים של מודלים לשוניים פתוחים מקומיים:

■ שליטה מלאה על הנתונים. כל המידע נשאר בתוך החברה, מה שמונע גישה לא מורשית ודליפת נתונים.

■ מבודדות. זה גם IT עבודה עצמאית. נמנעת תלות בחיבור לאינטרנט, דבר שחשוב במיוחד לעבודה בתשתיות

מבטיח עבודה רציפה בתנאים של סנקציות או חסימות של שירותי ענן.

גמישות בשימוש. המודל יכול לשמש ליצירת טקסטים, ניתוח נתונים, כתיבת קוד תוכנה, תמיכה בתכנון וניהול תהליכי עסקים.

התאמה לצרכים עסקיים. ניתן לאמן את המודל על מסמכים פנימיים, מה שמאפשר לקחת בחשבון את המאפיינים הספציפיים של עבודת החברה ותחום עיסוקה. ניתן לחבר את המודל הלשוני המקומי למערכות מה שמאפשר אוטומציה של ניתוח בקשות לקוחות, יצירת דוחות ואפילו חיזוי מגמות BI, או CRM, ERP.

על שרת, לגישה של צוות משתמשים שלם, בעלות של DeepSeek-R1-7B פריסת מודל חינמי ופתוח מבוססי ענן, כמו API דולר בחודש, עשויה להיות זולה יותר מהתשלומים השנתיים עבור 1000 ומאפשרת לחברות לשלוח לחלוטין בנתונים, למנוע את העברתם לאינטרנט, Claude או ChatGPT ו-GDPR ולעזור לעמוד בדרישות רגולציה, כמו

בתעשיות אחרות, מודלים לשוניים מקומיים כבר משנים את הגישה לאוטומציה. בשירותי תמיכה הם עונים על שאלות הם מנתחים קורות חיים ומחפשים מועמדים - HR נפוצות של לקוחות, מפחיתים את העומס על המוקדנים. במחלקות רלוונטיים. במסחר אלקטרוני - הם יוצרים הצעות מותאמות אישית, מבלי לחשוף נתוני משתמשים

בתחום הבנייה צפוי אפקט דומה. בזכות אינטגרציה של מודלים לשוניים עם נתוני פרויקטים ותקנות ניתן להאיץ את הכנת המסמכים, לאוטומט את הכנת האומדנים ולבצע ניתוח חיזוי של עלויות. כיוון מבטיח במיוחד הוא השימוש במודלים לשוניים בשילוב עם טבלאות מובנות ודאטה פריימים

שליטה מלאה על בינה מלאכותית בחברה ואיך להפעיל מודל לשוני גדול משלך

באופן מקומי בתוך מספר שעות בלבד. זה מעניק (LLM) הכלים המודרניים מאפשרים לחברות להפעיל מודל שפה גדול שליטה מלאה על הנתונים והתשתית, מסלק את התלות בשירותי ענן חיצוניים ומפחית את הסיכון לדליפת מידע. פתרון כזה הוא במיוחד רלוונטי לארגונים העובדים עם מסמכי פרויקט רגישים או נתונים מסחריים חסויים

בהתאם למשימות ולמשאבים, זמינים תרחישי פריסה שונים - מפתרונות מוכנים "מהקופסה" ועד לארכיטקטורות גמישות, המאפשר להפעיל מודלים לשוניים בלחיצת כפתור אחת, Ollama וניתנות להרחבה. אחד הכלים הפשוטים ביותר הוא Ollama: ללא צורך בידע טכני מעמיק. התחלה מהירה עם

1. מהאתר הרשמי (Windows / Linux / macOS) הורד את ההפצה עבור מערכת ההפעלה שלך
ollama.com

2. Mistral התקן את המודל דרך שורת הפקודה. לדוגמה, עבור המודל

```
ollama run mistral
```

3. לאחר ההפעלה, המודל מוכן לעבודה - תוכל לשלוח בקשות טקסט דרך הטרימינל או לשלב אותו בכלים אחרים. הפעל את המודל ובצע בקשה:

איך ליצור חישוב עם כל המשאבים לעבודה להתקנת קיר גבס ברוחב 100" ollama run mistral

אפליקציה חינוכית עם ממשק המזכיר את LM Studio עבור מי שמעדיף לעבוד בסביבה ויזואלית מוכרת, קיים ChatGPT:

Imstudio.ai - על ידי הורדת ההפצה מהאתר הרשמי LM Studio התקן את

והורד אותו (GPT-Neo-X או Falcon, למשל) דרך הקטלוג המובנה, בחר במודל

אך לחלוטין מקומי, ChatGPT עבוד עם המודל דרך ממשק אינטואיטיבי, המזכיר את

	Developer	Parameters	GPU Requirements (GB)	Features	Best For
Mistral 7B	Mistral AI	7	8 (FP16)	Fast, supports multimodal tasks (text + images), fully open-source code	Lightweight tasks, mobile devices, laptops
LLaMA 2	Meta	7-70	16-48 (FP16)	High text generation accuracy, adaptable for technical tasks, CC-BY-SA license	Complex analytical and technical tasks
Baichuan 7B/13B	Baichuan Intelligence	7-13	8-16 (FP16)	Fast and efficient, great for large data processing, fully open-source code	Data processing, automating routine tasks
Falcon 7B/40B	Technology Innovation Institute (TII)	7-40	8-32 (FP16)	Open-source, high performance, optimized for fast work	Workloads with limited computational resources
DeepSeek-V3	DeepSeek	671	1543 (FP16) / 386 (4-bit)	Multilingual, 128K token context window, balanced speed and accuracy	Large enterprises, SaaS platforms, multitasking scenarios
DeepSeek-R1-7B	DeepSeek	7	18 (FP16) / 4.5 (4-bit)	Retains 92% of R1 capabilities in MATH-500, local deployment support	Budget solutions, IoT devices, edge computing

קוד פתוח מקומיים LLM השוואת מודלים פופולריים של 3.34. ריש.

Mistral 7B מודלים קטנים, כמו. (3.34. ריש) הבחירה במודל תלויה בדרישות למהירות, דיוק ויכולות חומרה זמינות דורשים DeepSeek-V3 מתאימים למשימות קלות ולמכשירים ניידים, בעוד שמודלים חזקים, כמו Baichuan 7B ו-LLM-משאבים חישוביים משמעותיים, אך מספקים ביצועים גבוהים ותמיכה בשפות רבות. בשנים הקרובות, שוק ה המכסים את כל, LLM יתפתח במהירות - נראה יותר ויותר מודלים קלים ומיוחדים. במקום מודלים אוניברסליים של התוכן האנושי, יופיעו מודלים שהוכשרו על מומחיות צרה. לדוגמה, ניתן לצפות להופעת מודלים המיועדים אך ורק מודלים מיוחדים כאלה יהיו מהירים, מדויקים. CAD לעבודה עם חישובי הנדסה, אומדני בנייה או נתונים בפורמטים ובטוחים יותר לשימוש - במיוחד בסביבה מקצועית, שבה חשובה אמינות גבוהה ועומק נושאי

המקומי, ניתן להתאים אותו למשימות ספציפיות של החברה. לשם כך, נעשה שימוש בטכניקת LLM-לאחר הפעלת ה-שבה המודל עובר אימון נוסף על מסמכים פנימיים, הוראות טכניות, תבניות חוזים או מסמכי (fine-tuning) עדכון נוסף פרויקט.

עוזרי מודלים לשוניים גדולים אינטליגנטיים עם גישה לנתוני החברה: RAG

בעסקים הוא אינטגרציה של המודלים עם נתונים ארגוניים עדכניים LLM השלב הבא בהתפתחות השימוש במודלים של (הפקת תוכן עם תמיכה בהפקת מידע) - הפקה עם תמיכה בהפקת מידע. בארכיטקטורה RAG בזמן אמת. גישה זו נקראת, זו, מודל השפה הופך לא רק לממשק דיאלוגי, אלא לעוזר אינטלקטואלי מלא, המסוגל להתמצא במסמכים, שרטוטים מסדי נתונים ולספק תשובות מדויקות, בהקשר ובספציפיות.

הוא היכולת להשתמש בנתונים פנימיים של החברה מבלי צורך לאמן מחדש את RAG היתרון המרכזי של המודל, תוך שמירה על דיוק וגמישות גבוהה בעבודה עם מידע.

משלבת שני מרכיבים עיקריים RAG טכנולוגיית

■ - שרטוטים, PDF המודל מתחבר לאחסוני נתונים - מסמכים, טבלאות, קבצי: (Retrieval) הפקת מידע ומפיק מידע רלוונטי על פי בקשת המשתמש.

■ על בסיס המידע שהופק, המודל יוצר תשובה מדויקת: (Augmented Generation) הפקת תשובה ומבוססת, תוך התחשבות בהקשר ובספציפיות של הבקשה.

יש לבצע מספר צעדים, RAG-עם תמיכה ב LLM כדי להפעיל

■ הכנת נתונים: אספו את המסמכים, השרטוטים, המפרטים והטבלאות הנדרשים. הם עשויים להיות בפורמטים Excel ועד PDF-ובמבנים שונים, מ

■ הנתונים מומרים, LangChain או LlamaIndex אינדוקציה והווקטורזציה: באמצעות כלים כמו לייצוגים וקטוריים, המאפשרים למצוא קשרים סמנטיים בין קטעי טקסט) על בסיסי נתונים וקטוריים והמרת (נדבר יותר בחלק 8, CAD כמויות גדולות לייצוג וקטורי, כולל פרויקטי

■ בקשה לעוזר: לאחר טעינת הנתונים, ניתן לשאול את המודל שאלות, והוא יחפש תשובות במסגרת בסיס הנתונים הארגוני, ולא בידע הכללי שנאסף מהאינטרנט.

שבה מאוחסנים חוזים, הוראות, אומדנים וטבלאות. באמצעות constructionsite_docs נניח שבחברה יש תיקייה ניתן לסרוק את התיקייה הזו ולבנות אינדוקציה וקטורית: כל מסמך יומר לקבוצת, (3.35 איור) Python סקריפט וקטורים, המשקפים את התוכן הסמנטי של הטקסט. זה הופך את המסמכים ל"מפת משמעויות" שבה המודל יכול -להתמצא ביעילות ולמצוא קשרים בין מונחים וביטויים

לדוגמה, המודל "זוכר" שמילים כמו "החזרה" ו"דרישה" מופיעות לעיתים קרובות בפרק החוזה העוסק במשלוח חומרים - (שורת קוד 11 - 3.35 איור) "לאתר הבנייה. לאחר מכן, אם נשאל שאלה - לדוגמה", מהו המועד להחזרת המוצר ינתח את המסמכים הפנימיים וימצא את המידע המדויק, פועל כעוזר אינטלקטואלי המסוגל לקרוא ולהבין את LLM

תוכן כל הקבצים הארגוניים.

```

1 from llama_index import SimpleDirectoryReader, VectorStoreIndex
2
3 # Load documents from the folder
4 documents = SimpleDirectoryReader("construction_docs").load_data()
5
6 # Creating a vector index for semantic search
7 index = VectorStoreIndex.from_documents(documents)
8
9 # Integration with LLM (e.g. Llama 3)
10 query_engine = index.as_query_engine()
11 response = query_engine.query("What are the return terms in the contracts?")
12 print(response)

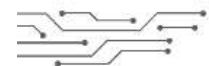
```

קורא את תיקיית הקבצים - בדומה לאופן שבו אדם פותח אותה ומחפש את המסמך הנדרש **LM 3.35** איור.

להרצת הקוד IDE ו-Python נדבר יותר על השימוש ב Python ניתן להריץ את הקוד על כל מחשב שבו מותקן בפרק הבא.

מקומית אינה רק טרנד, אלא החלטה אסטרטגית עבור חברות שמעריכות אבטחה וגמישות. עם זאת, פריסת LLM פריסת בין אם על מחשבים מקומיים של החברה או באמצעות פתרונות מכוונים - היא רק הצעד הראשון. כדי ליישם את LLM, במשימות אמיתיות, חברות צריכות להשתמש בכלים המאפשרים לא רק לקבל תשובות בצ'אטים, אלא גם LLM יכולות זה חשוב להרחבת הפתרונות - תהליכים. LLM-לשמור את הלוגיקה שנוצרה כקוד שניתן להריץ מחוץ להקשר השימוש ב מיד על מספר פרויקטים או אפילו על כל החברה AI מאורגנים כראוי מאפשרים ליישם את הפיתוחים של

משחקת תפקיד חשוב. כלים מודרניים לתכנות מאפשרים לא רק (IDE) בהקשר זה, הבחירה בסביבת פיתוח מתאימה ETL-Pipeline-אלא גם לשלב אותם בתהליכי עסק קיימים, והופכים אותם ל LLM, לפתח פתרונות מבוססי אוטומטיים.



פרק 3.4. עם תמיכה במודלים לשוניים גדולים ושינויים עתידיים בתכנות IDE

מניסויים עם מודלים לשוניים גדולים לפתרונות עסקיים: IDE בחירת

(LLM) כשמתעמקים בעולם האוטומציה, ניתוח נתונים ובינה מלאכותית - במיוחד בעבודה עם מודלים לשוניים גדולים מתאימה. זו תהיה הכלי המרכזי שלכם: המקום שבו יופעל הקוד (IDE) חשוב מאוד לבחור בסביבת פיתוח משולבת - משפיעה לא רק על נוחות העבודה IDE-הן על מחשב מקומי והן בתוך הרשת הארגונית. הבחירה ב, LLM שנוצר על ידי לפתרונות מלאים המוטמעים בתהליכי עסק אמיתיים LLM-אלא גם על כמה מהר תוכלו לעבור מבקשות ניסיוניות ב

(סביבת פיתוח משולבת) היא קומפלקס בנייה אוניברסלי על המחשב שלכם לאוטומציה של תהליכים IDE ועיבוד נתונים. במקום לאחסן בנפרד מסור, פטיש, מקדחה וכלים אחרים, יש לכם מכשיר אחד שיכול עבור מתכנתים היא מרחב IDE. לעשות הכל - לחתוך, לקבע, לקדוח ואפילו לבדוק את איכות החומרים, אחד שבו ניתן לכתוב קוד (באנלוגיה לבנייה - ליצור שרטוטים, (לבדוק את פעולתו) בניית מודל הבניין) למצוא שגיאות (כמו בדיקת עמידות המבנים בבנייה (ולהשיק את הפרויקט המוגמר) מסירת הבית (לשימוש).

פופולריים IDEs סקירה של

■ היא מתאימה מאוד לפרויקטים Python. מקצועית ועוצמתית עבור IDE היא PyCharm® (JetBrains) אינטראקטיביים Jupyter רציניים בזכות כמות גדולה של פונקציות מובנות. עם זאת, התמיכה הבסיסית בקבצי זמינה רק בגרסה בתשלום, ומתחילים עשויים למצוא את הממשק עמוס מדי (IPYNB)

הוא פורמט של מחברות IPYNB (Interactive Python Notebook) קובץ עם סיומת שבו קוד, ויזואליזציות והסברים מאוחדים במסמך אחד, Jupyter® Notebook, אינטראקטיביות של -פורמט זה אידיאלי לבניית דוחות, ניתוחים ותסריטי לימוד

■ ומגוון IPYNB-הוא כלי מהיר, גמיש וניתן להתאמה אישית עם תמיכה הינמית ב VS Code® (Microsoft) ותוספים לעבודה GitHub Copilot תוספים. הוא מתאים גם למתחילים וגם למקצוענים. מאפשר לשלב את עם מודלים לשוניים, מה שהופך אותו לבחירה מצוינת עבור פרויקטים בתחום הבינה המלאכותית ומדע הנתונים. ■ בחירה קלאסית ופופולרית לניסויים ולמידה. מאפשר לכתוב קוד, להוסיף הסברים — Jupyter Notebook וליצור צעדים LLM אידיאלי לבדוק השערות במהירות, לעבוד עם (3.41 איור) ולחזות תוצאות באותו ממשק — Anaconda Navigator-ניתנים לשחזור בניתוח נתונים. לניהול תלותים וספריות מומלץ להשתמש ב Python. ממשק גרפי לניהול סביבות



Jupyter-אלטרנטיבה מבוססת ענן ל – (Kaggle (9.25 איור) וגם פלטפורמת Google Collab™, זו פתרון מצוין להתחלה – ללא התקנה של תוכנה מקומית ועם אפשרות GPU/TPU-המספקת גישה חינם ל (Google של Gemini (LLM ועם Google Drive לעבוד ישירות מהדפדפן. תומך באינטגרציה עם -לאחרונה

Pipeline. הוא אחד הכלים הנוחים והפשוטים ביותר ליצירת תהליכי **IDE: Jupyter Notebook** השוואת **3.42** אזור

Jupyter Notebook במהירות, נסה את **AI** תלויה במשימות שלך. אם אתה רוצה להתחיל לעבוד עם **IDE**-הבחירה ב העיקר – להתחיל. כלים. **VS Code** או **PyCharm**-עבור פרויקטים רציניים עדיף להשתמש ב **Google Collab** או מודרניים מאפשרים להפוך ניסויים לפתרונות עובדים במהירות

המתוארים מאפשרים ליצור פייפליינים לעיבוד נתונים – כלומר, שרשראות של מודולים של קוד) שיכולים **IDE**-כל ה

שכל אחד מהם אחראי על שלב מסוים, לדוגמה, (LLM להיות מיוצרים על ידי

■ תסריטים אנליטיים,

■ שרשראות של חילוף מידע ממסמכים

■ RAG, תגובות אוטומטיות על בסיס

■ יצירת דוחות וחזות נתונים.

בזכות המבנה המודולרי, כל שלב ניתן לייצוג כחלק נפרד: טעינת נתונים → סינון → ניתוח → חזות נתונים → ייצוא תוצאות. ניתן לעשות שימוש חוזר בחלקים אלו, להתאים ולבנות שרשראות חדשות, כמו קונסטרוקטור, רק עבור נתונים

עבור מהנדסים, מנהיגים ואנליסטים, זה פותח אפשרות לתעד את הלוגיקה של קבלת החלטות בצורה של קוד, שיכול גישה זו מסייעת להאיץ משימות שגרתיות, לאוטומט פעולות סטנדרטיות וליצור תהליכים LLM. להיות מיוצר בעזרת ניתנים לשחזור, שבהם כל שלב מתועד וברור לכל חברי הצוות

Apache NiFi (אזור 7.44), Apache Airflow כלים כמו (אזור 7.23), אוטומטיים ETL Pipelines פרטים נוספים על לבניית בלוקים של לוגיקה באוטומציה של תהליכים יידונו בחלקים 7 ו-8 של (אזור 7.46) n8n ו- (אזור 7.45) NiFi. --- הספר.

עם תמיכה במודלים לשוניים גדולים ושינויים עתידיים בתכנות IDE

אינטגרציה של אינטליגנציה מלאכותית בתהליכי פיתוח משנה את הנוף של התכנות. סביבות מודרניות אינן רק עורך טקסט עם הדגשת תחביר – הן הופכות לעוזרים אינטליגנטיים, המסוגלים להבין את הלוגיקה של הפרויקט, להשלים קוד AI: ואפילו להסביר כיצד פועל קטע קוד מסוים. בשוק מופיעים מוצרים המרחיבים את גבולות הפיתוח המוכר בעזרת

■ המייצר קוד על בסיס הערות או, AI עוזר: (VS Code, PyCharm-מתממשק ל) GitHub Copilot תיאורים חלקיים, והופך רמזים טקסטואליים לפתרונות מוכנים

■ מאפשר לא רק להשלים קוד, אלא גם לשאול שאלות על: (AI עם ליבת VS Code פורק של) Cursor הפרויקט, לחפש תלותים וללמוד על בסיס הקוד

■ עם פונקציית הסבר על קוד (PyCharm כולל) JetBrains של IDE-תוסף ל: JetBrains של AI-עוזר ה מורכב, אופטימיזציה ויצירת בדיקות

■ של אמזון AWS עם דגש על אבטחה ותמיכה בשירותי Copilot-מקביל ל: Amazon CodeWhisperer

תכנות בשנים הקרובות יעבור שינויים דרמטיים. המוקד העיקרי יזוז מכתובת קוד שגרתית לעיצוב מודלים וארכיטקטורת ייקח על עצמו משימות שגרתיות: יצירת קוד, בדיקות AI-נתונים - המפתחים יתעסקו יותר בעיצוב מערכות, בעוד שה שבו מכונות לוקחות על עצמן את AI-תיעוד ופונקציות בסיסיות. העתיד של התכנות הוא שיתוף פעולה בין בני אדם ל השגרה הטכנית, ואנשים מתמקדים ביצירתיות

תגיע לרמה חדשה - סביבות הפיתוח ילמדו להתאים IDE תכנות בשפה טבעית יהפוך לעניין יומיומי. התאמה אישית של את עצמן לסגנון העבודה של המשתמש, וחברות, צופות דפוסים, יציעו פתרונות הקשר ולומדות על בסיס פרויקטים קודמים

זה לא מבטל את תפקיד המפתח, אך משנה אותו באופן דרמטי: מכתיבת קוד - לניהול ידע, איכות ותהליכים. אבולוציה כזו תיגע גם בתחום ניתוח העסקים, שבו יצירת דוחות, ויזואליזציות ואפליקציות צ'אט וסוכני LLM ו-AI לתמיכה בקבלת החלטות תתרחש יותר ויותר דרך יצירת קוד ולוגיקה בעזרת ממשק.

ובחרה את סביבת הפיתוח המתאימה, השלב החשוב הבא הוא ארגון הנתונים. תהליך LLM לאחר שהחברה הגדרה צ'אט זה כולל חילוץ מידע ממקורות מפוזרים, ניקוי, המרתו לצורת מבנה ויישומו במערכות ארגוניות.

בגישה המודרנית הממוקדת בנתונים לניהול נתונים, המטרה המרכזית היא להביא את הנתונים לצורת יוניברסלית אחת. שתהיה תואמת למגוון רחב של כלים ואפליקציות. לעבודה עם תהליכי מבנה ונתונים מובנים נדרשות ספריות מיוחדות לפייתון. היא מאפשרת לעבד נתונים Pandas אחת מהספריות החזקות, הגמישות והפופולריות ביותר היא ספריית טבלאיים בנוחות: לסנן, לקבץ, לנקות, להשלים, לבצע אגרגציות ולבנות דוחות.

כלי חיוני לעבודה עם נתונים: Python Pandas

תופסת מקום מיוחד. זו אחת מהספריות הפופולריות והנמצאות בשימוש Pandas, בעולם ניתוח ואוטומציה של נתונים נרחב בשפת התכנות פייתון, המיועדת לעבודה עם נתונים מובנים.

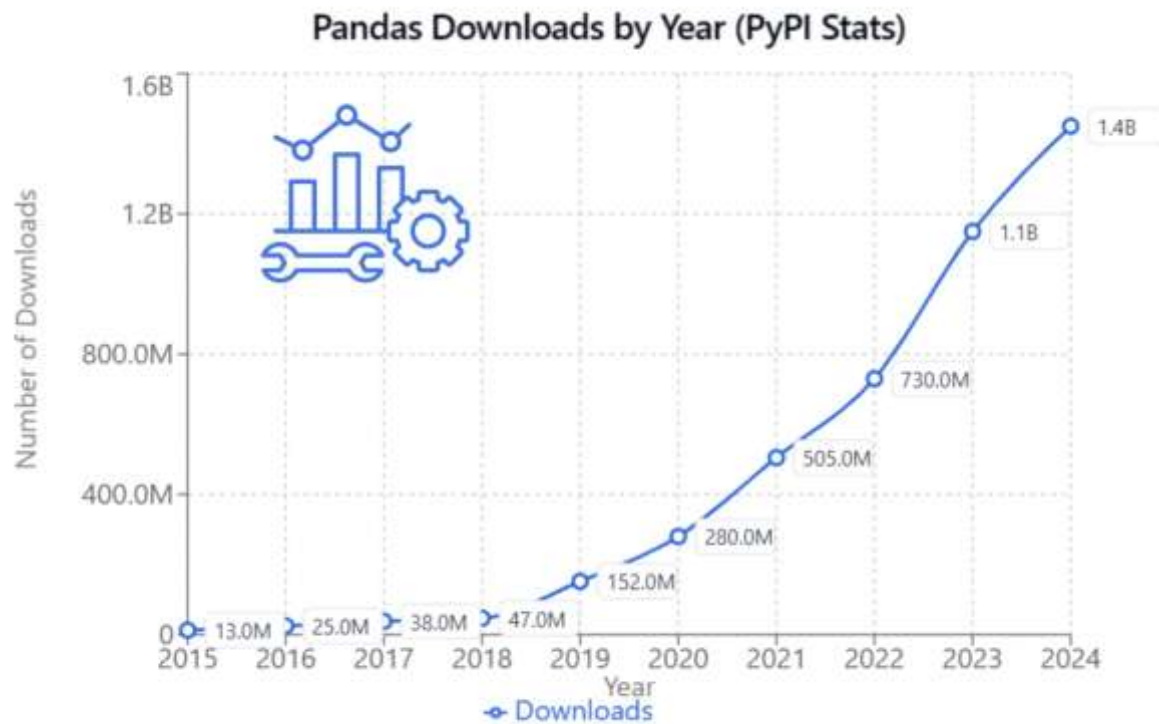
ספרייה היא כמו סט של כלים מוכנים: פונקציות, מודולים, מחלקות. כמו שבאתר בנייה לא צריך כל פעם להמציא פטיש או פלס, כך גם בתכנות ספריות מאפשרות לפתור בעיות במהירות מבלי להמציא מחדש פונקציות ופתרונות בסיסיים.

היא ספריית פייתון עם קוד פתוח, המספקת מבני נתונים בעלי ביצועים גבוהים ואינטואיטיביים Pandas, היא כמו סכין שווייצרית Pandas. פורמט אוניברסלי לעבודה עם טבלאות - DataFrame ובפרט עבור אנליסטים, מהנדסים ומפתחים העובדים עם נתונים.

פייתון – שפת תכנות ברמה גבוהה עם תחביר פשוט, בשימוש נרחב בניתוח נתונים, אוטומציה, למידת מכונה ופיתוח אתרים. הפופולריות שלה מוסברת על ידי קריאות הקוד, תמיכה בפלטפורמות שונות ואקוסיסטמה עשירה של ספריות עד היום נוצרו עבור פייתון יותר מ-137,000 חבילות קוד פתוח, ומספר זה ממשיך לגדול כמעט מדי יום. כל ספרייה, כזו היא מעין מאגר של פונקציות מוכנות: מהפעולות המתמטיות הפשוטות ביותר ועד לכלים מורכבים לעיבוד תמונות. ניתוח נתונים גדולים, עבודה עם רשתות נוירונים ואינטגרציה עם שירותים חיצוניים.

במילים אחרות, דמיינו שיש לכם גישה חופשית ופתוחה למאות אלפי פתרונות תוכנה מוכנים – ספריות, וכלים, אותם תוכלו לשלב ישירות בתהליכי העבודה שלכם. זה כמו קטלוג עצום של אפליקציות. המיועדות לאוטומציה, ניתוח, ויזואליזציה, אינטגרציה ועוד – וכל זה זמין מיד לאחר התקנת פייתון.

פנדס – אחת מהחבילות הפופולריות ביותר באקוסיסטמת פייתון. בשנת 2022, ממוצע ההורדות של ספריית פנדס הגיע ל-4 מיליון ביום, בעוד שבתחילת 2025 מספר זה עלה ל-12 מיליון הורדות ביום, מה שמעיד על הפופולריות הגוברת של LLM. - שלה והשימוש הנרחב בה בניתוח נתונים ובצ'אטים של

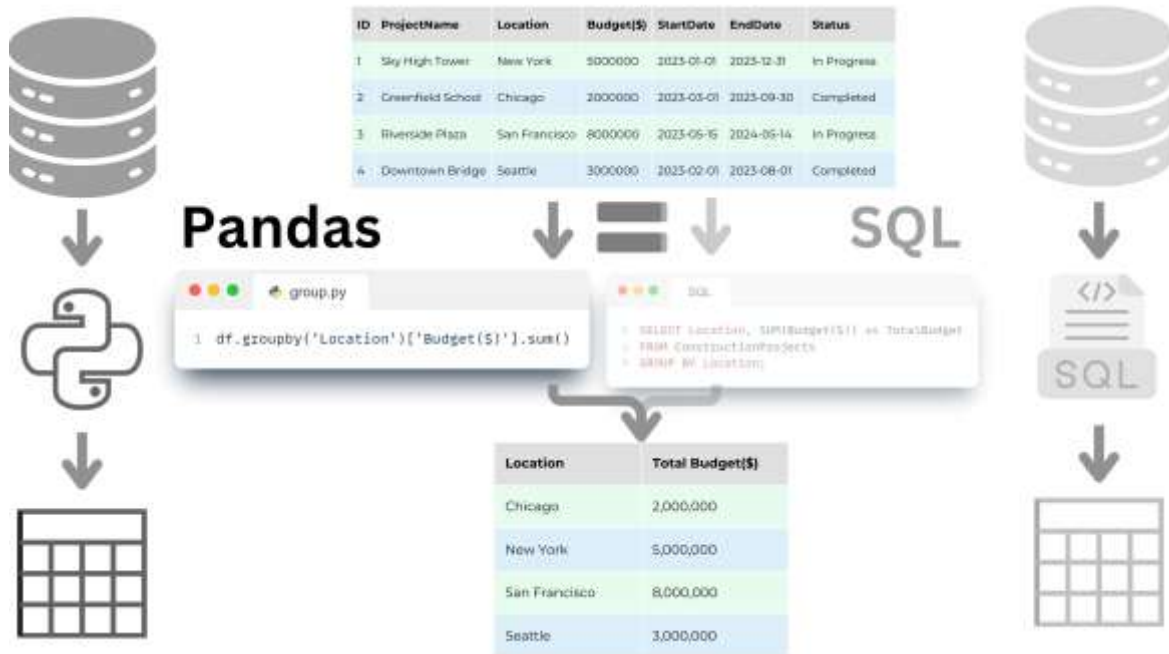


- פנדס – אחת מהספריות המורדות ביותר. בשנת 2024, מספר ההורדות השנתי שלה עבר את ה-1.4 מיליארד

אשר דנו בה בפרק "מסדי נתונים, SQL שפת השאילתות בספריית פנדס דומה בפונקציונליות שלה לשפת השאילתות "SQL רלציוניים ושפת השאילתות

בעולם הניתוח וניהול הנתונים המובנים, פנדס מתבלטת בפשטות, מהירות ועוצמה, ומספקת למשתמשים מגוון רחב של כלים לניתוח ועיבוד מידע בצורה יעילה

המסורתי. הם Excel-ופנדס – מספקים יכולות חזקות לעבודה עם נתונים, במיוחד בהשוואה ל SQL – שני הכלים מותאם לעבודה עם מסדי נתונים רלציוניים, בעוד SQL-תומכים בפעולות כמו דגימה, סינון, כאשר ההבדל הוא ש מה שמאפשר להריץ אותה על כל מחשב, מבלי הצורך ליצור מסדי נתונים, (RAM) שפנדס מעבדת נתונים בזיכרון -ולחקים תשתית נפרדת

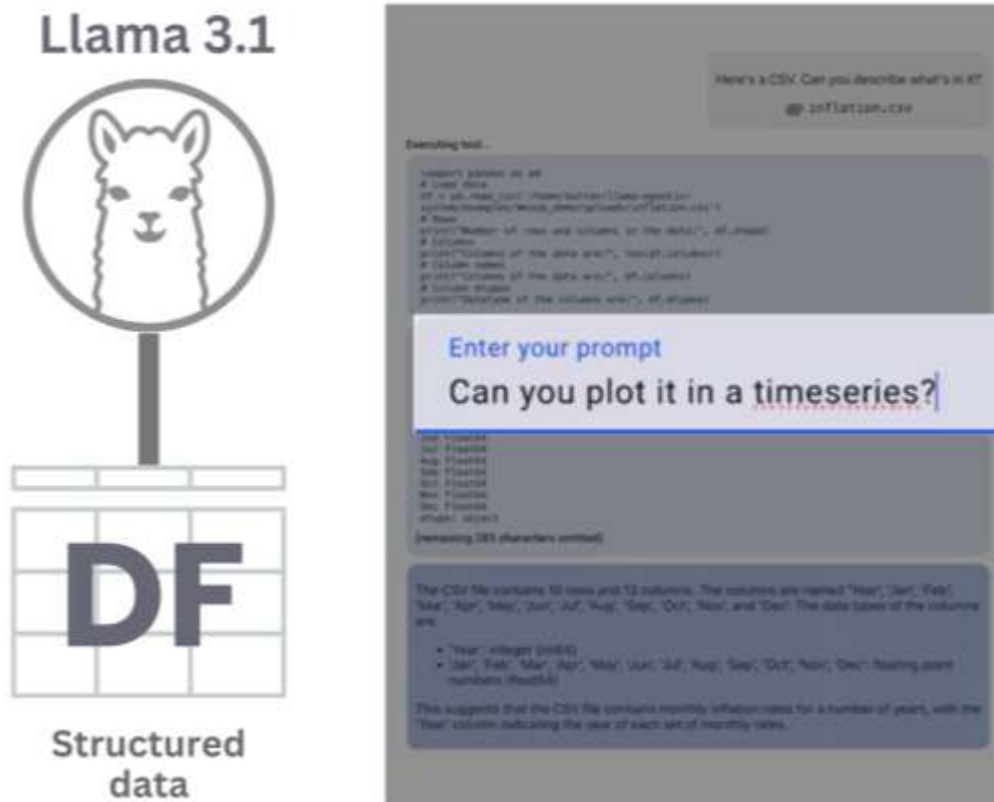


- מציעה גמישות בעבודה עם פורמטים שונים של נתונים, מבלי להיות מוגבלת למסדי נתונים **SQL**, פנדים, בניגוד ל

ומניפולציה של (ETL כולל) פנדים משמשת לעיתים קרובות במחקר מדעי, אוטומציה של תהליכים, יצירת צינורות הוא הסטנדרט לניהול מסדי נתונים ומשמש לעיתים קרובות בסביבות ארגוניות לעבודה **SQL**-נתונים בפייתון, בעוד ש עם כמויות גדולות של נתונים

ספריית פנדים בשפת התכנות פייתון מאפשרת לבצע לא רק פעולות בסיסיות, כמו קריאה וכתיבה של טבלאות, אלא גם משימות מורכבות יותר, כולל שילוב נתונים, קבוצת נתונים וביצוע חישובים אנליטיים מורכבים.

LLM משמשת לא רק במחקר אקדמי ובאנליזת עסקים, אלא גם בשילוב עם מודלים של Pandas היום, ספריית בשנת 2024, שמה LLaMa 3.1 (פייסבוק)™ (כאשר פרסמה את המודל החדש בקוד פתוח Meta® לדוגמה, מחלקת CSV דגש מיוחד על עבודה עם נתונים מובנים, והפכה את עיבוד מסדי הנתונים המובנים) איור 3.45 (בפורמט -ישירות בצ'אט לאחד מהמקרים המרכזיים והראשונים בשחרור שלה Pandas ואינטגרציה עם ספריית



בשנת 2024 היה בניית אפליקציות **Meta-Llama 3.1** ב-**Meta** איור 3.45 אחד מהמקרים הראשונים והמרכזיים שהציגה צוות **Pandas** באמצעות.

הוא כלי חיוני עבור מיליוני מדעני נתונים המעבדים ומכינים נתונים עבור אינטליגנציה מלאכותית **Pandas** הוא לא שינויים בקוד תהיה צעד משמעותי קדימה. מדעני נתונים יוכלו לעבד נתונים בתוך **Pandas** גרטיבית. האצת דקות, ולא שעות, ולקבל כמות נתונים רבה יותר לאימון מודלים גרטיביים של אינטליגנציה מלאכותית [88]. – ג'סן **NVIDIA** חואנג, מייסד ומנכ"ל

בדרך כלל Excel-בעוד ש **Excel** ניתן לנהל ולנתח מערכי נתונים, הרבה מעבר ליכולות של **Pandas** באמצעות יכול בקלות לעבוד עם מערכי נתונים) איור 9.12, איור 9.110 (**Pandas**, מסוגל לעבד עד מיליון שורות נתונים המכילים עשרות מיליוני שורות [89]. יכולת זו מאפשרת למשתמשים לבצע ניתוח נתונים מורכב וויזואליזציה על מערכי יש תמיכה **Pandas**-נתונים גדולים, מה שמספק הבנה מעמיקה ומקל על קבלת החלטות מבוססות נתונים. בנוסף, ל, משתמשים בו מדי יום באינטרנט או לא מקוון (**Kaggle.com**, **Google Collab**, **Microsoft® Azure™ Notebooks**, **Amazon SageMaker**) חזקה מהקהילה [90]: מאות מיליוני מפתחים ואנליסטים ברחבי העולם, ומספקים כמות גדולה של פתרונות מוכנים לכל בעיות עסקיות.

המוצעת על ידי ספריית **DataFrame**, נמצאת צורת נתונים מובנית הנקראת **Python**-בבסיס רוב תהליכי האנליזה ב זהו כלי חזק וגמיש לארגון, ניתוח וויזואליזציה של נתונים טבלאיים **Pandas**.

פורמט אוניברסלי לנתונים טבלאיים DataFrame:

המייצג טבלה דו-ממדית (איור 3.46), (שבה השורות תואמות, Pandas הוא המבנה המרכזי בספריית DataFrame לאובייקטים או רשומות נפרדות, והעמודות תואמות לתכונות, פרמטרים או קטגוריות שלהם. מבנה זה מזכיר ויזואלית אך עולה עליהן בהרבה גמישות, יכולת הרחבה ופונקציונליות, Excel טבלאות.

הוא דרך לייצג ולעבד נתונים טבלאיים, המאוחסנים בזיכרון המחשב DataFrame.

הוא דרך לייצג ולעבד נתונים טבלאיים, המאוחסנים בזיכרון המחשב. בטבלה, השורות יכולות לשקף DataFrame למשל, רכיבי פרויקט בנייה, והעמודות יכולות להיות תכונותיהם: קטגוריות, ממדי, קואורדינטות, עלויות, מועדים ועוד. בטבלה כזו יכולה להיות מידע על פרויקט אחד (איור 4.113), (כמו גם נתונים על מיליוני אובייקטים מאלפי פרויקטים ניתן בקלות לסנן, לקבץ ולאגד כמויות מידע כאלה, Pandas שונים) איור 9.110. (בזכות פעולות הווקטוריות של -במהירות גבוהה

ID	Name	Category	Family Name	Height	BoundingBoxMin_X	BoundingBoxMin_Y	BoundingBoxMin_Z	Level
431144	Single-Flush	OST_Doors	Single-Flush	6.88976378	20.1503	-10.438	9.84252	Level 1
431198	Single-Flush	OST_Doors		6.88976378	13.2281	-1.1207	9.84252	Level 2
457479	Single Window	OST_Windows	Single Window	8.858267717	-11.434	-11.985	9.80971	Level 2
485432	Single Window	OST_Windows	Single Window	8.858267717	-11.434	4.25986	9.80971	Level 2
490150	Single-Flush	OST_Doors	Single-Flush	6.88976378	-1.5748	-2.9565	-1E-16	Level 1
493697	Basic Wall	OST_Walls	Basic Wall		-38.15	20.1656	-4.9213	Level 1
497540	Basic Wall	OST_Walls	Basic Wall		-4.5212	-0.0708	9.84252	Level 1

הוא טבלה דו-ממדית עם רכיבים בשורות ואטריבוטים בעמודות DataFrame איור 3.46 פרויקט בנייה בצורה של

DataFrames, - לפי הערכת נוידיה, כבר היום עד 30% מכלל המשאבים החישוביים משמשים לעיבוד נתונים מובנים וחלק זה ממשיך לגדול

עיבוד נתונים הוא מה שעוסקת בו, כנראה, שליש מכלל החישובים בעולם בכל חברה. עיבוד נתונים והנתונים של פורמט טבלאי, DataFrame-רוב החברות נמצאים ב

ג'נסן חואנג, מנכ"ל נוידיה -

DataFrame ב-Pandas: נפרט כמה תכונות מפתח של

הנתונים מאורגנים בעמודות, שכל אחת מהן יש לה שם ייחודי. עמודות-מאפיינים DataFrame-עמודות: ב יכולות להכיל נתונים מסוגים שונים, בדומה לעמודות בבסיסי נתונים או עמודות בטבלאות

דומה לרשימה או לעמודה בטבלה, שבה לכל, Pandas, היא מבנה נתונים חד-ממדי ב-Pandas Series. ערך יש אינדקס משלו.

יש יותר מ-400 מאפיינים ושיטות, מה שהופך את העבודה עם נתונים לגמישה ב-Pandas Series. להפליא. ניתן ליישם ישירות אחת מתוך ארבע מאות הפונקציות הזמינות על עמודה, לבצע פעולות תומכת Series, מתמטיות, לסנן נתונים, להחליף ערכים, לעבוד עם תאריכים, מחרוזות ועוד. בנוסף בפעולות וקטוריות, מה שמאפשר משמעותית את עיבוד מערכות נתונים גדולות בהשוואה לחישובים מחזוריים. לדוגמה, ניתן בקלות להכפיל את כל הערכים במספר, להחליף נתונים חסרים או ליישם טרנספורמציות מורכבות מבלי לכתוב לולאות מסובכות.

ניתן לאנדקס שורות בערכים ייחודיים. אינדקס זה מאפשר לשנות ולתקן במהירות DataFrame-שורות ב: נתונים בשורות מסוימות.

N-1 מקצה לכל שורה אינדקס מ-0 עד DataFrame, Pandas אינדקס: כברירת מחדל, כאשר נוצר עם זאת, ניתן לשנות את האינדקס כך שיהיה סימונים (DataFrame-הוא מספר השורות ב N כאשר) מיוחדים, כמו תאריכים או תכונות ייחודיות.

פירושו שלכל שורה מוקצה שם ייחודי או תווית, הנקראת אינדקס DataFrame-אינדקס שורות ב DataFrame.

ו-`datetime64`, `bool`, `float`, `int`: תומך במגוון סוגי נתונים, כולל DataFrame: סוגי נתונים יש סוג נתונים משלה, המגדיר אילו פעולות ניתן DataFrame עבור נתוני טקסט. לכל עמודת `object` לבצע על תוכנה.

תומך במגוון רחב של פעולות לעיבוד נתונים, כולל אגרגציה DataFrame: פעולות עם נתונים חלוקה-יישום-שילוב ועוד שיטות רבות, חיבור, (`concat`), מיזוג, (`groupby`), להמרת נתונים.

מאפשר הוספה והסרה של עמודות ושורות, מה שהופך אותו למבנה DataFrame: מניפולציה של גדלים דינמי שניתן לשנותו בהתאם לצרכי ניתוח הנתונים.

ויזואליזציה של נתונים: באמצעות שיטות הויזואליזציה המובנות או על ידי אינטראקציה עם ספריות-ניתן בקלות להמיר את ה, Seaborn או Matplotlib פופולריות לוויזואליזציה של נתונים, כמו לגרפים ודיאגרמות כדי להציג את הנתונים בצורה גרפית DataFrame.

מספק פונקציות לקריאה, ייבוא וייצוא נתונים בפורמטים שונים של קבצים Pandas: קלט ופלט נתונים לנקודת מרכזית לאיסוף DataFrame-מה שהופך את ה, CSV, Excel, JSON, HTML ו-SQL, כמו והפצת נתונים.

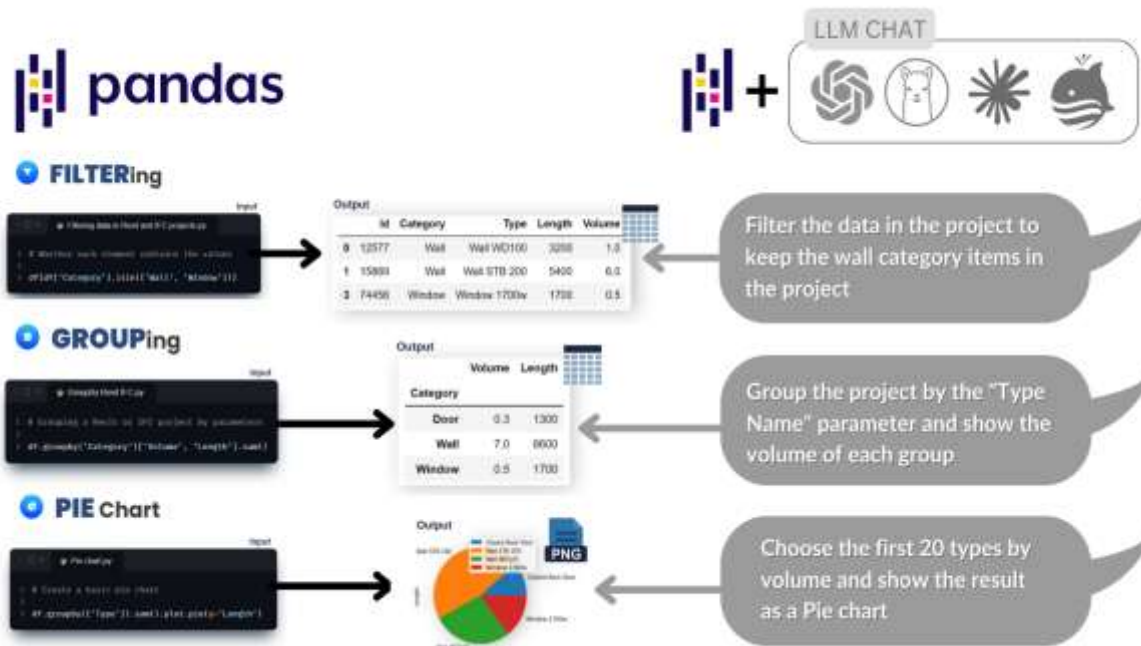
מספק גמישות וביצועים גבוהים יותר בעבודה עם Pandas של CSV ו-XLSX, בניגוד לפורמטים נתונים: הוא מאפשר לעבד כמויות גדולות של מידע בזיכרון, תומך בסוגי נתונים מתקדמים (כולל תאריכים, ערכים לוגיים אינו שומר מידע על CSV-ורצפי זמן, (ומספק אפשרויות רחבות לסינון, אגירה, שילוב וויזואליזציה של נתונים. בעוד ש נשאר DataFrame, לעיתים קרובות עמוס בעיצוב ויש לו יכולת סקלביליות נמוכה XLSX-סוגי נתונים ומבנה, והבחירה האופטימלית לניתוח מהיר, אוטומציה של תהליכים ואינטגרציה עם מודלים של בינה מלאכותית) 3.47). בפרקים הבאים נבחן בפירוט כל אחד מהאספקטים הללו של נתונים, כמו כן בחלק 8 של הספר יידונו בפירוט פורמטים

--ומערכות נתונים) איור 8.12 (Parquet, Apache Orc, JSON, Feather, HDF5 דומים, כגון

	XLSX	CSV	Pandas DataFrame
Storage	Tabular	Tabular	Tabular
Usage	Office tasks, data presentation	Simple data exchange	Data analysis, manipulation
Compression	Built-in	None	None (in-memory)
Performance	Low	Medium	High (memory dependent)
Complexity	High (formatting, styles)	Low	Low
Data Type Support	Limited	Very limited	Extended
Scalability	Low	Low	Medium (memory limited)

3.47 **DataFrame** איור - נתונים DataFrame ותמיכה מתקדמת בסוגי נתונים -

הפכו לסטנדרט דה-פקטו בתחום DataFrame ופורמט Pandas בזכות הגמישות, העוצמה והקלות בשימוש, ספריית הם מתאימים באופן מושלם גם ליצירת דוחות פשוטים וגם לבניית צינורות ניתוח מורכבים. Python-ניתוח הנתונים ב LLM. במיוחד בשילוב עם מודלים של



3.48 במקום קוד, מספיק בקשה טקסטואלית **Pandas** מפשטים את האינטראקציה עם **LLM** איור

ואחרים **ChatGPT, LLaMa, DeepSeek, QWEN** כמו - **LLM** בשימוש פעיל בצ'אטים מבוססי **Pandas** כיום במקרים רבים, כאשר המודל מקבל בקשה הקשורה לעיבוד טבלאות, בדיקת נתונים או ניתוח, הוא מייצר קוד באמצעות -לשפה הטבעית לייצוג נתונים בשיחות עם בינה מלאכותית) איור 3.48 **DataFrame** זה הופך את **Pandas** ספריית ממשטות באופן משמעותי את הניתוח, האוטומציה והאינטגרציה של **Pandas** טכנולוגיות עיבוד נתונים מודרניות, כמו נתונים בתהליכי עסקים. הן מאפשרות לקבל תוצאות במהירות, להפחית את העומס על המומחים ולהבטיח את שחזור הפעולות.

צעדים נוספים: יצירת מסגרת נתונים עמידה

בחלק זה בדקנו את סוגי הנתונים המרכזיים בשימוש בתעשיית הבנייה, הכרנו פורמטים שונים לאחסוןם וניתחנו את בעיבוד מידע. הבנו כי ניהול נתונים יעיל הוא הבסיס לקבלת החלטות, **LLM** ו-**IDE** תפקידם של כלים מודרניים, כולל מושכלות ואוטומציה של תהליכי עסקים. ארגונים המסוגלים למבנה ולסדר את הנתונים שלהם מקבלים יתרון תחרותי משמעותי בשלבי עיבוד והמרת הנתונים.

לסיכום חלק זה, כדאי להדגיש את הצעדים המעשיים העיקריים שיכולים לעזור ליישם את הגישות שנדונו במשימות: הדימויות שלך

ערכו אודיט נתונים בתהליכים שלכם

☐ הרכיבו רישום של כל סוגי הנתונים שאתם משתמשים בהם בפרויקטים

☐ קבעו אילו סוגים ודגמים של נתונים הם הקריטיים ביותר לתהליכי העסקים שלכם

☐ זיהוי אזורים בעייתיים שבהם המידע נשאר לעיתים קרובות לא מובנה, חלש במבנה או לא נגיש

☒ התחילו לפתח אסטרטגיית ניהול נתונים

☐ העלו שאלות לגבי מדיניות וסטנדרטים לעבודה עם סוגי נתונים שונים

☐ נתחו אילו מהתהליכים שלכם ניתן לשפר על ידי המרת נתונים לא מובנים לנתונים מובנים

☐ צרו תקנות לאחסון וגישה לנתונים, תוך התחשבות בביטחון ובפרטיות

☒ הקימו והכירו את הכלים הבסיסיים לעבודה עם נתונים

☐ Jupyter או VS Code מתאים, התואם את הצרכים שלכם) למשל התקינו את IDE בחרו (Notebook)

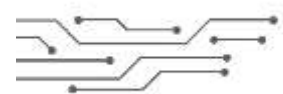
☐ מקומית לעיבוד אישי של הנתונים שלכם LLM נסו להתקין

☐ XLSX לעיבוד נתונים טבלאיים Pandas התחילו להתנסות עם ספריית

☐ את המשימות הטיפוסיות שאתם מעבדים בכלים טבלאיים או בבסיסי נתונים ובקשו מ LLM-תארו ב Pandas לאוטומט את העבודה בעזרת LLM

יישום צעדים דומים יאפשר לכם בהדרגה לשנות את הגישה לעבודה עם נתונים, לעבור ממאגרים מפורזים ולא מובנים הראשון ב DataFrame לאקוסיסטמה אחת, שבה הנתונים הופכים לנכס נגיש וברור. התחילו בקטן – צרו את למשל בעבודה עם Python מקומית, אוטומטו את המשימה השגרתית הראשונה בעזרת LLM הריצו Pandas (Excel-טבלאות ב).

החלק הרביעי של הספר יוקדש לשאלות איכות הנתונים, ארגון, מבניהם ומודולציה. נתמקד במתודולוגיות המאפשרות למאגרים מובנים, המתאימים לניתוח – CAD וטקסטים ועד תמונות ודגמי PDF-להפוך מקורות מידע מפורזים – מ ואוטומציה. כמו כן, נלמד כיצד מתבצע תהליך הפורמליזציה של דרישות הנתונים, כיצד נבנים מודלים קונספטואליים (LLM) ולוגיים בפרויקטים בנייה, וכיצד תהליכים אלו יכולים להיעזר במודלים לשוניים מודרניים





IV חלק

איכות נתונים: ארגון, מבנה, מודל

החלק הרביעי מתמקד במתודולוגיות ובטכנולוגיות המבטיחות את המרת המידע המפוזר, למאגרים מובנים באיכות גבוהה. נבחן בפירוט את תהליכי יצירת ותיעוד דרישות הנתונים כבסיס לארכיטקטורת מידע יעילה בפרויקטים בנייה. יוצגו שיטות מעשיות להפקת מידע עם דוגמאות (CAD תמונות, קבצי טקסט, דגמי PDF מסמכי) מובנה ממקורות שונים וכלים אחרים לאימות ולווידוא (Regex) ליישום. ינותח השימוש בביטויים רגולריים אוטומטי של נתונים. יתואר שלב אחר שלב תהליך המודולציה של נתונים ברמות קונספטואליות, לוגיות ופיזיות תוך התחשבות במאפיינים של תחום הבנייה. יוצגו דוגמאות לאוטומציה של תהליכי מבנה ובדיקת מידע (LLM) קונקרטיות לשימוש במודלים לשוניים. יציעו גישות יעילות להמחשת תוצאות הניתוח, המגבירות את נגישות המידע האנליטי לכל רמות הניהול בפרויקטים בנייה.

פרק 4.1

המרת נתונים לפורמט מובנה

בעידן הכלכלה המונעת נתונים, הנתונים הופכים לא רק למכשול אלא לבסיס לקבלת החלטות. במקום להתאים את המידע לכל מערכת חדשה ולפורמטים שלה, חברות שואפות יותר ויותר ליצור מודל נתונים מובנה אחד, שישמש כמקור האמת האוניברסלי לכל התהליכים. מערכות מידע מודרניות מעוצבות לא סביב פורמטים וממשקים, אלא סביב משמעות הנתונים. שכן המבנה יכול להשתנות, אך משמעות המידע נשארת קבועה הרבה יותר זמן –

המפתח לעבודה יעילה עם נתונים אינו בהמרה ובטרנספורמציה אינסופית שלהם, אלא בארגון הנכון מלכתחילה: יצירת מבנה אוניברסלי, המסוגל להבטיח שקיפות, אוטומציה ואינטגרציה בכל שלב במעגל חיי הפרויקט.

הגישה המסורתית מחייבת בעת יישום כל פלטפורמה חדשה לעסוק בתיקונים ידניים: העברת נתונים, שינוי שמות מאפיינים, והתאמת פורמטים. צעדים אלו אינם משפרים את איכות הנתונים עצמם, אלא רק מסתירים בעיות, ויוצרים מעגל סגור של המרות אינסופיות. כתוצאה מכך, חברות הופכות תלויות בפתרונות תוכנה ספציפיים, והטרנספורמציה הדיגיטלית מואטת.

בפרקים הבאים נבחן כיצד למבנה את הנתונים בצורה נכונה, ולאחר מכן כיצד ליצור מודלים אוניברסליים, למזער את התלות בפלטפורמות ולהתמקד בעיקר - בנתונים כמשאב אסטרטגי, סביבו נבנים תהליכים יציבים

תמונות וטקסטים לפורמטים מובנים, PDF, לומדים להמיר מסמכים

בפרויקטים בנייה, רוב המידע קיים בצורה לא מובנית: אלו מסמכים טכניים, פרוטוקולים של עבודות שבוצעו, שרטוטים מפרטים, ולוחות זמנים. המגוון שלהם - הן בפורמט והן בתוכן - מקשה על אינטגרציה ואוטומציה

תהליך ההמרה לפורמטים מובנים או חצי-מובנים יכול להשתנות בהתאם לסוג הנתונים הנכנסים ולתוצאות הרצויות מעיבוד הנתונים.

המרת נתונים מצורת לא מובנית לצורת מובנית - זו גם אמנות וגם מדע. תהליך זה משתנה בהתאם לסוג הנתונים הנכנסים בעיבוד נתונים ואנליטיקה, במטרה (3.25 איור) ולמטרות הניתוח ולעיתים תופס חלק משמעותי מעבודתו של המהנדס לקבל סט נתונים נקי ומסודר



המרת מסמך סרוק לא מובנה לפורמט טבלאי מובנה 4.11 איור.

זהו תהליך שלב אחר שלב, הכולל את השלבים - (4.11 איור) תמונות וטקסטים לפורמט מובנה, PDF, הפיכת מסמכים הבאים:-

■ בשלב זה נטען המסמך או התמונה המקורית, המכילה נתונים לא מובנים. זה יכול (Extract) חילוץ נתונים צילום, שרטוט או סקיצה, PDF להיות, למשל, מסמך.

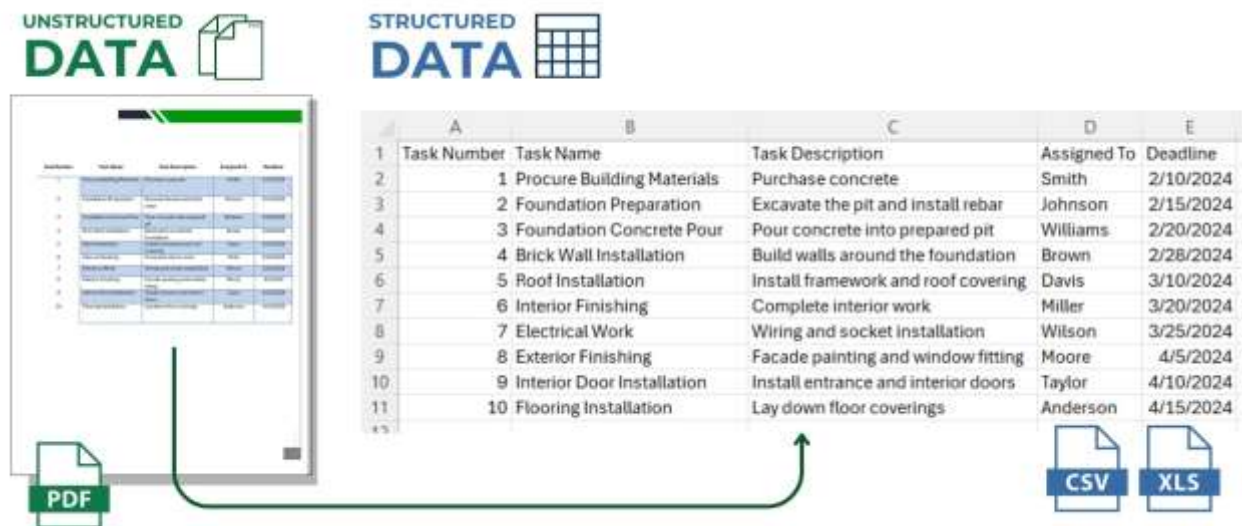
■ לאחר מכן מגיע שלב המרת הנתונים הלא מובנים לפורמט מובנה. לדוגמה, זה: (Transform) המרת נתונים או שיטות עיבוד אחרות (OCR) יכול לכלול זיהוי ופירוש טקסט מתמונות באמצעות זיהוי תווים אופטי.

■ השלב האחרון כולל שמירה של הנתונים המעובדים בפורמטים שונים: (Load) טעינה ושמירה של נתונים לעבודה נוספת, כאשר הבחירה בפורמט תלויה בדרישות ובהעדפות, CSV, XLSX, XML, JSON, כגון הספציפיות.

משחק תפקיד מרכזי בעיבוד אוטומטי של נתונים, עליו (ETL (Extract, Transform, Load) תהליך זה, הידוע בשם וצינור: חילוץ, המרה, טעינה. "בהמשך נבחן דוגמאות כיצד מסמכים בפורמטים שונים ETL" נדון בהרחבה בפרק מומרים לנתונים מובנים.

לטבלה PDF דוגמה להמרת מסמך

כדי להדגים את PDF אחת מהשימושים השכיחות ביותר בפרויקטים בנייה - עיבוד מסמכי דרישות טכניות בפורמט CSV והמרתה לפורמט PDF המעבר מנתונים לא מובנים לפורמט מובנה, נבחן דוגמה מעשית: חילוץ טבלה מתוך מסמך - (4.12 איור) Excel או.



נפוצים בקלות ומשתלבים בקלות במערכות ניהול נתונים שונות XLSX ו-CSV פורמטים, PDF-בניגוד ל 4.12 איור

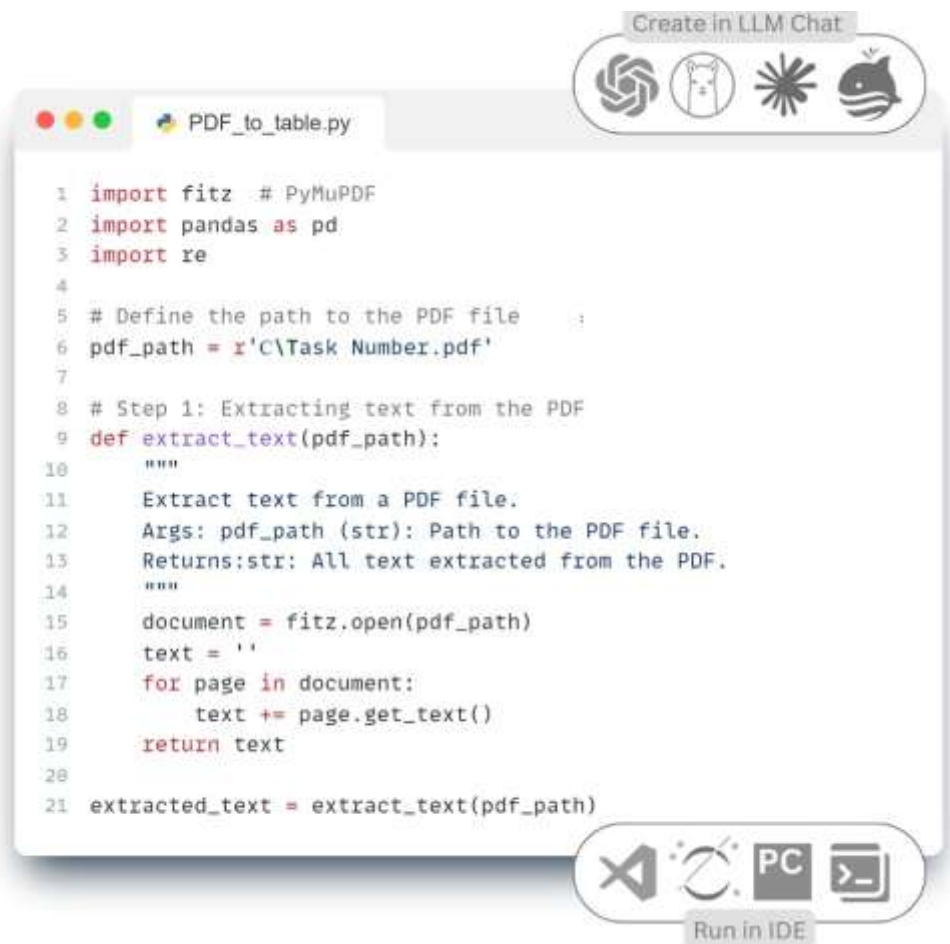
מפשטים באופן ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN, כמו LLM מודלים לשפה משמעותי את עבודת המומחים עם נתונים, מפחיתים את הצורך בלימוד מעמיק של שפות תכנות ומאפשרים פתרון של משימות רבות באמצעות בקשות טקסטואליות

או בפורומים וערוצי שיחה StackOverFlow לכן, במקום לבזבז זמן בחיפוש פתרונות באינטרנט) בדרך כלל באתר המודרניים, הן LLM ייעודיים (או לפנות למומחים בתחום עיבוד הנתונים, אנו יכולים לנצל את האפשרויות של לפורמט טבלי PDF באינטרנט והן במקומי. מספיק להגיש בקשה, והמודל יספק קוד מוכן להמרת מסמך

LLM (ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN) שלחו את הבקשה הטקסטואלית הבאה לכל מודל (או כל מודל אחר QWEN):

שמכיל טבלה. הקוד צריך לקבל כארגומנט את PDF כתוב בבקשה קוד להפקת טקסט מקובץ DataFrame התיב לקובץ ולהחזיר את הטבלה המופקת בפורמט

שכן שפה זו בשימוש נרחב לעיבוד, Python-ברוב המקרים תינתן בצורת קוד ב LLM-התשובה של מודל הנתונים, אוטומציה ועבודה עם פורמטים שונים של קבצים:



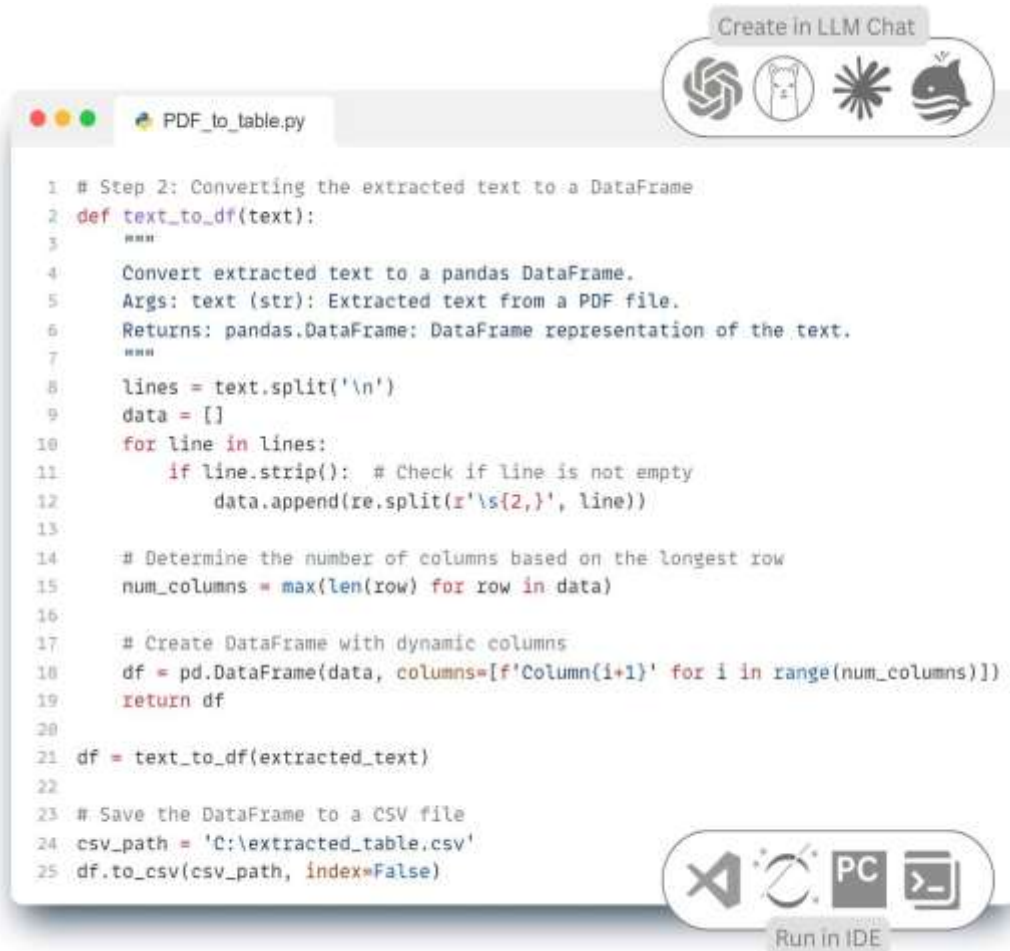
PDF. מפיקה טקסט מקובץ (Pandas, Fitz) והספריות והחבילות שלו Python בצורת קוד LLM תשובת 4.13. ריס

PyCharm, הפופולריות שדיברנו עליהן קודם, במצב לא מקוון IDE-ניתן להרצה באחת מה (ריס) קוד זה עם Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse או בכלים מקוונים Python, JupyterLab עם תוסף Python, IntelliJ IDEA, Thonny, Wing IDE, PyDev, TOSF Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker. פופולריים

Python (הפופולרית) שעליה דיברנו בהרחבה בפרק Pandas בשלב "המרה" נשתמש בספריית DataFrame-כלי בלתי ניתן להחלפה לעבודה עם נתונים, ("כדי לקרוא את הטקסט המופק ל: Pandas CSV או XLSX בקובץ טבלאי DataFrame-ולשמור את ה

כמו כן, הוסף קוד PDF ל-DataFrame אני צריך קוד שימיר את הטבלה המתקבלת מקובץ CSV בקובץ DataFrame-לשמירת ה

LLM: תשובת



4.14. ריש CSV ושמירת הטבלה בקובץ DataFrame-ל-PDF-המרת הטבלה המופקת מ

- מתרחשת שגיאה - למשל, בשל חוסר בספריות או נתיב שגוי לקובץ (ריס. 4.14, ריס. 4.13) אם בעת הרצת הקוד המודל ינתח את הודעת LLM. ניתן פשוט להעתיק את טקסט השגיאה יחד עם הקוד המקורי ולשלוח אותו שוב למודל השגיאה, יסביר מה הבעיה ויציע תיקונים או צעדים נוספים

הופכת למחזור שלם: בקשה → תשובה → בדיקה → משוב → תיקון LLM AI-כך, האינטראקציה עם ה
ללא צורך בידע טכני מעמיק -

IDE, שאנו יכולים להריץ מקומית בכל Python ועשרה שורות LLM באמצעות בקשה טקסטואלית רגילה בצ'אט קל לקריאה על ידי מכונה ומשתלב במהירות PDF, אשר, בניגוד למסמך CSV, לפורמט טבלאי PDF המרתנו מסמך בכל מערכות ניהול הנתונים.

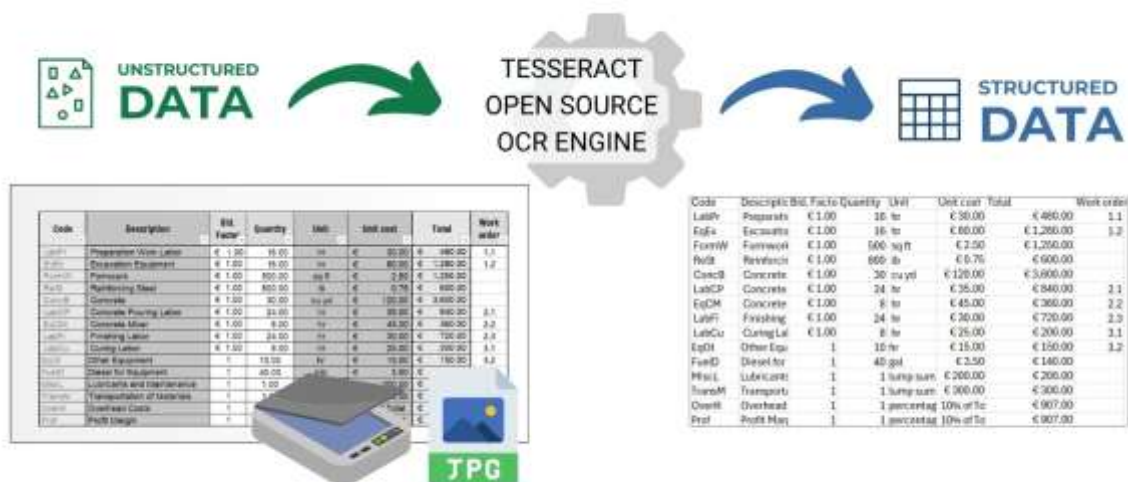
על עשרות ואפילו אלפי LLM על ידי העתקתו מכל צ'אט, (ריס. 4.14, ריס. 4.13) אנו יכולים ליישם את הקוד הזה CSV. חדשים בשרת, ובכך לאוטומט את תהליך המרת זרם המסמכים הלא-מובנים לפורמט טבלאי מובנה PDF מסמכי

לא תמיד מכילים טקסט, לרוב מדובר במסמכים סרוקים שיש לעבד אותם כדימויים. למרות שדימויים PDF אך מסמכי הם מטבעם לא מובנים, פיתוח ויישום ספריות זיהוי מאפשרים לחלץ, לעבד ולנתח את תוכנם, מה שמאפשר לנו לנצל את הנתונים הללו במלואם בתהליכי העסקים.

לפורמט מובנה JPEG, PNG המרת תמונת

דימויים הם אחת הצורות הנפוצות ביותר של נתונים לא מובנים. בבנייה ובתחומים רבים אחרים, כמות עצומה של מידע נשמרת בצורה של מסמכים סרוקים, שרטוטים, תמונות ושרטוטים. נתונים כאלה מכילים מידע יקר ערך, אך לא ניתן, או מסד נתונים. דימויים מכילים הרבה מידע מורכב, שכן תוכנם, צבעיהם Excel לעבד אותם ישירות כמו טבלת ומרמיהם מגוונים, ועבור חילוף מידע שימושי נדרשת עיבוד מיוחד.

הקושי בשימוש בדימויים כמקור נתונים טמון בהיעדר מבנה. דימויים אינם מעבירים משמעות בדרך ישירה, שקל לכמת או טבלת מסד נתונים. כדי להמיר נתונים Excel אותה, באופן שהמחשב יכול להבין או לעבד מיד, כפי שעושה טבלת לא מובנים מדימויים לצורת מבנה, יש להשתמש בספריות מיוחדות המסוגלות לפרש את המידע הוויזואלי הכלול בהם. (איור 4.15).



OCR. איור 4.15 המרת מסמכים סרוקים ודימויים לפורמטים מובנים אפשרית בעזרת כלים מיוחדים של

(זיהוי תווים אופטי. (טכנולוגיה זו מאפשרת לזהות אותיות OCR לחילוף טקסט מדימויים נעשה שימוש בטכנולוגיית והופכת אותם לטקסט שניתן לעריכה ולקריאה על ידי מכונה, PDF ומספרים בסריקות של מסמכים, בתמונות ובקבצי בשימוש כבר זמן רב באוטומציה של ניהול מסמכים וכיום הן משתלבות בקלות בכל תהליכי העסקים OCR טכנולוגיות

HP™ קוד פתוח, שפותח במקור על ידי Tesseract, הוא OCR אחד הכלים הפופולריים ביותר של Python. וביישומי הוא תומך ביותר מ-100 שפות ומצטיין בדיוק גבוה בזיהוי. Google™ וכיום נתמך על ידי

צ'אט לכתוב דוגמת קוד לחילוץ נתונים מטבלה סרוקה או מצולמת בצורה מובנית LLM-בואו נבקש מ

(ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN) צ'אט LLM-שלחו בקשה טקסטואלית ל:

↩ DataFrame-המכילה טבלה ל JPEG כתוב קוד המאפשר להמיר תמונת

🗒 לזיהוי טקסט בדימויים Pytesseract ברוב המקרים תציע להשתמש בספריית LLM-התשובה מ



```

1 import pytesseract
2 from PIL import Image
3 import pandas as pd
4
5 # Load the image from file
6 image_path = '/mnt/data/ConstructionDDC,'
7 image = Image.open(image_path)
8
9 # Use Tesseract to convert the image to text
10 text = pytesseract.image_to_string(image, lang='eng')
11
12 # Split the text into lines and then columns
13 rows = text.split('\n')
14 data = [row.split('\t') for row in rows if row.strip()]
15
16 # Create a DataFrame using pandas
17 # Assuming the first row is the header
18 df = pd.DataFrame(data[1:], columns=data[0])
19
20 # Display the DataFrame
21 df.head() # Show only the first few rows for brevity
22

```

איור 4.16 המרת טקסט, שהופק מטבלת דימויים או תמונות, לייצוג טבלאי מובנה.

(Python עבור Tesseract) pytesseract משתמש בספריית LLM-בדוגמה זו – הקוד) איור 4.16, (שהתקבל מ

להמרת הטקסט הזה לצורת מבנה Pandas (זיהוי תווים אופטי (ובספריית OCR להמרת דימוי לטקסט באמצעות DataFrame, כלומר.

תהליך ההמרה כולל בדרך כלל עיבוד מקדים לשיפור איכות הדימוי, ולאחר מכן מוחלים אלגוריתמים שונים לזיהוי דפוסים, חילוך תכונות או זיהוי אובייקטים. כתוצאה מכך, מידע ויזואלי לא מובנה מומר לנתונים מובנים.

ודימויים הם מקורות מרכזיים של מידע לא מובנה, האלוף האמיתי בנפח הוא הטקסט שנוצר במיילים PDF-למרות ש צ'אטים, פגישות ומסנג'רים. נתונים אלה לא רק רבים - הם מפוזרים, לא פורמליים ומאוד לא מובנים.

המרת נתוני טקסט לפורמט מובנה

עם טבלאות (איור 4.12) (וגרסאות סרוקות של טפסים טבלאיים) (איור 4.15), (חלק משמעותי PDF בנוסף למסמכי מהמידע במסמכי פרויקט מוצג בטקסט. זה יכול להיות משפטים מחוברים במסמכים טקסטואליים, כמו גם רישומים פגומים המפוזרים על גבי שרטוטים ודיאגרמות. בתנאים המודרניים של עיבוד נתונים, אחת מהמשימות הנפוצות היא -להמיר טקסט כזה לפורמט מובנה, המתאים לניתוח, ויזואליזציה וקבלת החלטות.

האלמנט המרכזי בתהליך זה הוא טקסונומיה - מערכת סיווג המאפשרת לארגן מידע לפי קטגוריות ותתי קטגוריות על בסיס תכונות משותפות.

טקסונומיה היא מבנה סיווג היררכי, המשמש לקיבוץ וארגון אובייקטים. בהקשר של עיבוד טקסט, היא משמשת כבסיס לחלוקה שיטתית של רכיבים לקטגוריות משמעותיות, מה שמקל על הניתוח ומעלה את איכות עיבוד הנתונים.

יצירת טקסונומיה מלווה בשלבים של חילוך ישויות, קטגוריזציה וחיבור להקשר. כדי לדמות את תהליך חילוך המידע PDF: מנתוני טקסט, יש לבצע את הצעדים הבאים, הדומים לאלה שכבר השתמשנו בהם במבנה נתונים מתוך מסמכי

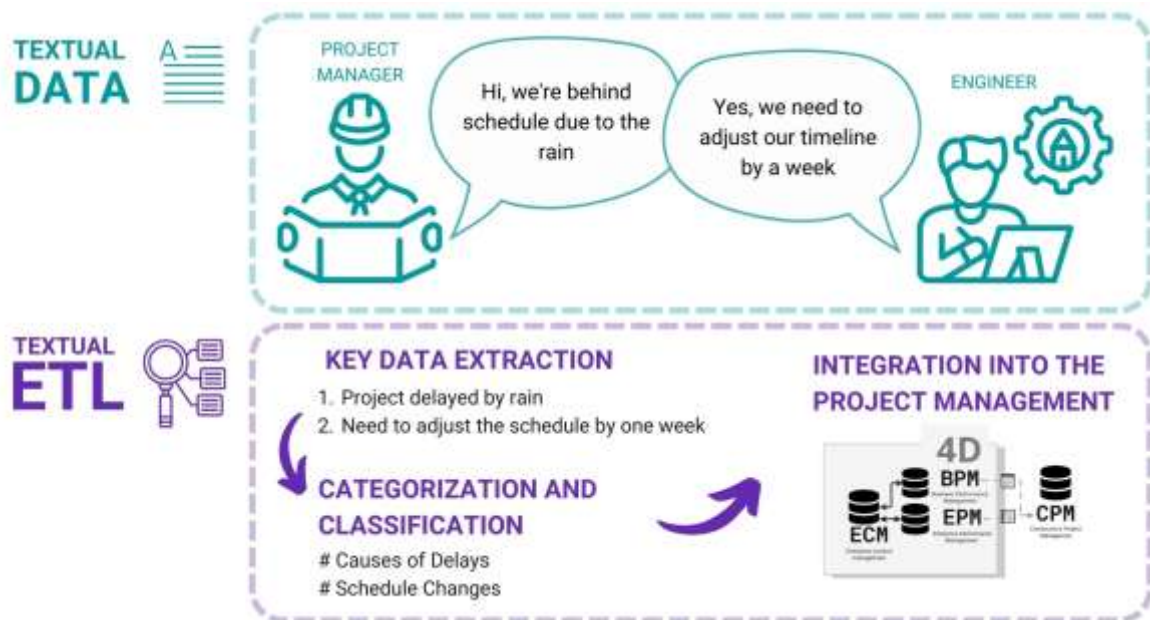
יש לנתח את הנתונים הטקסטואליים כדי לחלץ מידע על עיכובים ושינויים בלוח: (Extract) חילוך נתונים הזמנים של הפרויקט

מחלקים את המידע שהתקבל לקטגוריות, לדוגמה, סיבות לעיכובים: (Transform) קטגוריזציה וסיווג ושינויים בלוח הזמנים

בסוף מכינים את הנתונים המובנים לאינטגרציה עם מערכות ניהול נתונים חיצוניות: (Load) אינטגרציה

נבחן מצב: יש לנו דיאלוג בין מנהל פרויקט למהנדס, שבו נדונות בעיות בלוח הזמנים. המטרה שלנו היא לחלץ את האלמנטים המרכזיים (סיבות לעיכוב, התאמות בלוחות הזמנים (ולהציגם בצורה מובנית) (איור 4.17).

כדי לדמות את חילוך הנתונים ולאחר מכן DataFrame נבצע את החילוך על בסיס מילות מפתח צפויות, ניצור חדשה, שתכיל עמודות עבור תאריך, אירוע (למשל, סיבת העיכוב (ופעולה) למשל, שינוי בלוח DataFrame ניצור (הזמנים).



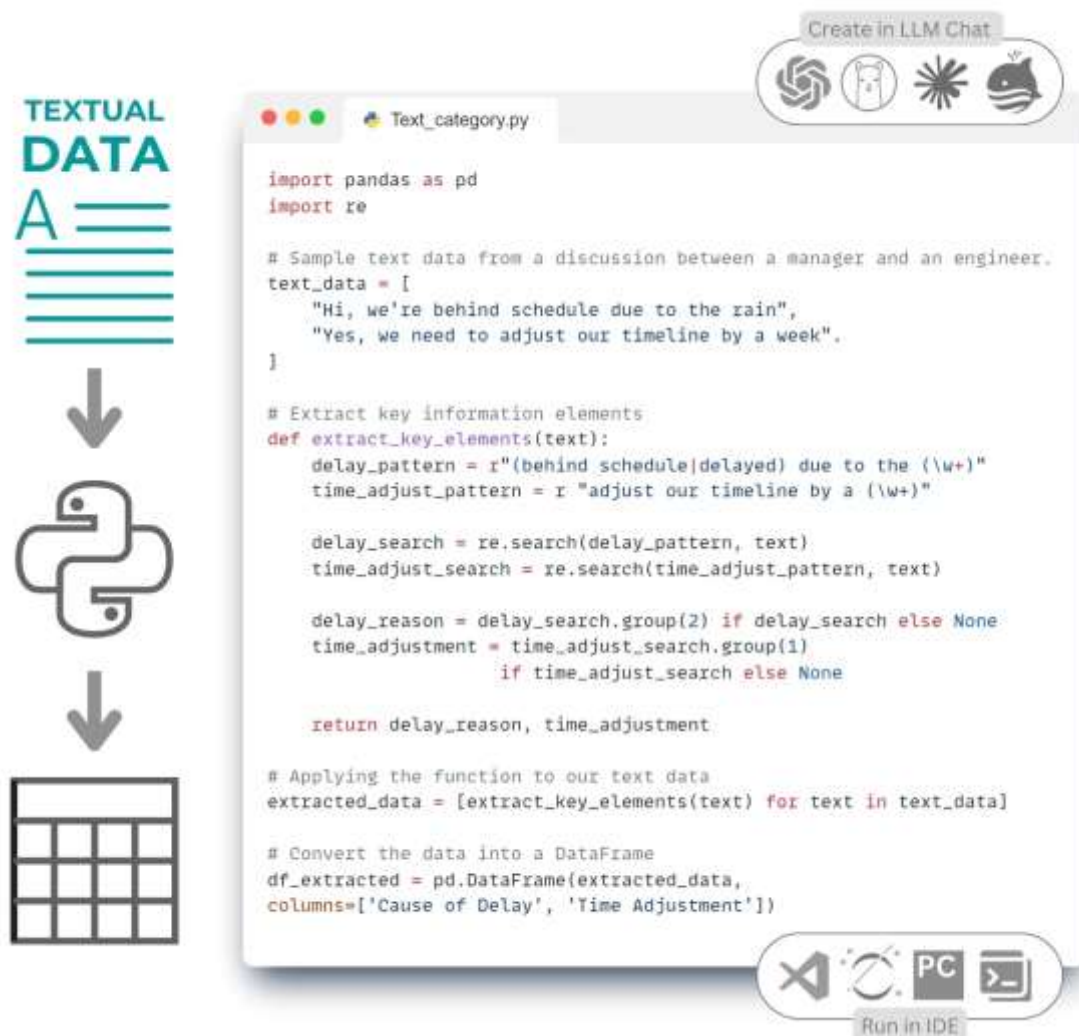
איור 4.17 חילוץ מידע מרכזי מהטקסט על הצורך בהתאמת לוחות זמנים ואינטגרציה של שינויים במערכת ניהול הפרויקט.

נציג קוד לפתרון הבעיה באמצעות בקשה טקסטואלית באחת מהמודלים הלשוניים, כמו בדוגמאות הקודמות

LLM: שלחו בקשה טקסטואלית בכל צ'אט

יש לי שיחה בין מנהל: שלום, אנחנו מאחרים בלוח הזמנים בגלל גשם "ומהנדס": כן, אנחנו צריכים להתאים את לוחות הזמנים בשבוע. "אני צריך תסריט שינתח שיחות טקסט דומות בעתיד, יחלץ מהן מנתונים אלה. לאחר DataFrame סיבות לעיכובים והתאמות בלוחות הזמנים, ולאחר מכן ייצור CSV צריך להישמר בקובץ DataFrame-מכן, ה

ובספריית (re - Regex) תוך שימוש בביטויים רגולריים Python בדרך כלל תכלול קוד LLM-התגובה מ Pandas (pd):

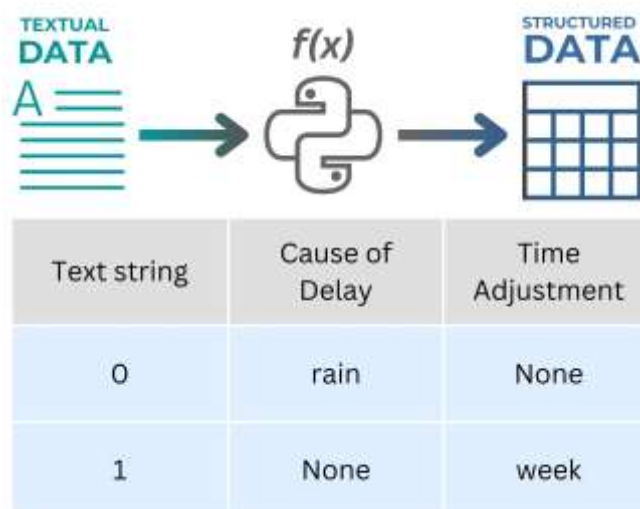


איור 4.18 הוצאת מידע מפתח מהטקסט על הצורך בתיקון מועדים בצורה של טבלה.

בדוגמה זו (איור 4.17) הנתונים הטקסטואליים, הכוללים התכתבות בין מנהל הפרויקט למהנדס, מנותחים כדי לזהות ולהוציא מידע ספציפי שיכול להשפיע על ניהול פרויקטים עתידיים עם דיאלוגים דומים. באמצעות ביטויים רגולריים נקבעים באמצעות תבניות הסיבות ("Regex על ביטויים רגולריים נדבר בפרק "דרישות מובנות וביטויים רגולריים") לעיכובי הפרויקט והתיקונים הנדרשים בלוח הזמנים. הפונקציה שנכתבה בדוגמה זו מוציאה משורות הטקסט או את סיבת העיכוב או את התיקון בזמן, בהתבסס על תבניות: היא מוציאה את המילה שאחר "בגלל" כסיבת העיכוב או את המילה שאחר "על" כתיקון הזמן.

אם בשורה מוזכר עיכוב בגלל מזג האוויר, אז הסיבה שנקבעת היא "גשם"; אם בשורה מוזכר תיקון בלוח הזמנים לתקופה "מסוימת, אז התקופה הזו מוצאת כתיקון הזמן (איור 4.19). (היעדר כל אחת מהמילים הללו בשורה מביא לערך "לא

עבור העמודה המתאימה.



לאחר ביצוע הקוד מכילה מידע על קיום עיכובים ותיקונים **DataFrame**, איור 4.19 טבלה מסכמת, שהתקבלה בצורה של נדרשים בזמן.

מבנה ופרמטריזציה של תנאים מהטקסט (דיאלוג, מכתב, מסמך (מאפשרים לתקן במהירות עיכובים בבנייה: לדוגמה חוסר בעובדים יכול להשפיע על קצב העבודה במזג אוויר רע, ולכן חברות, בדיעת פרמטרי העיכוב מהדיאלוגים) איור בין המנהל בשטח למנהל הפרויקט - יכולות מראש לחזק את הצוות במקרה של תחזית לא טובה (4.19).

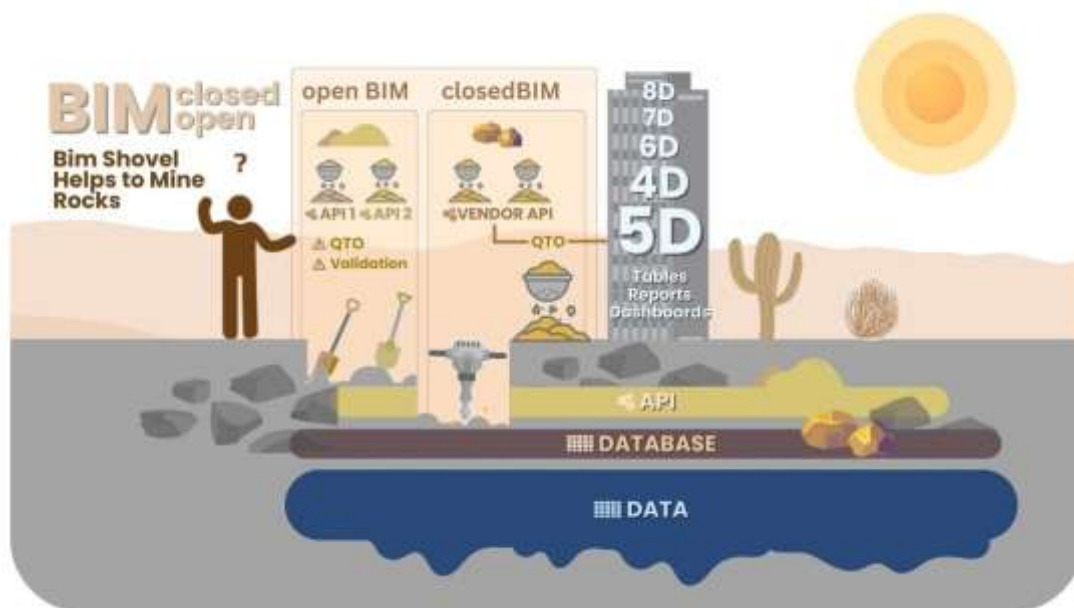
המרת מסמכים ודימויים לפורמט מובנה יכולה להתבצע בעזרת כלים יחסית פשוטים, פתוחים וחינמיים, המבוססים על קטגוריזציה.

CAD קטגוריזציה של רכיבים היא גם חלק מרכזי בעבודה עם נתוני פרויקטים, במיוחד בהקשר של שימוש בתוכנות (BIM).

לפורמט מובנה CAD (BIM) המרת נתוני

CAD היא משימה מורכבת יותר, מכיוון שהנתונים שנשמרים ממסדי נתונים CAD (BIM) מבנה וקטגוריזציה של נתוני כמעט תמיד מוצגים בפורמטים סגורים או בפרמטריים מורכבים, לעיתים משולבים בו זמנית רכיבי נתונים (BIM) (גיאומטריים) חצי-מובנים (ורכיבי מידע מטה) חצי-מובנים או מובנים.

בדרך כלל מוגנים ואינם זמינים לשימוש ישיר, אלא אם כן CAD (BIM) הפורמטים הנייטיביים של נתונים במערכות של המפתח עצמו) איור 4.110. (בידוד הנתונים הזה יוצר מחסנים API נעשה שימוש בתוכנה מיוחדת או בממשקי סגורים - מה שנקרא "סילואים", המגבילים את חילופי המידע החופשי ומעכבים את יצירת תהליכים דיגיטליים רציפים. - בחברה.



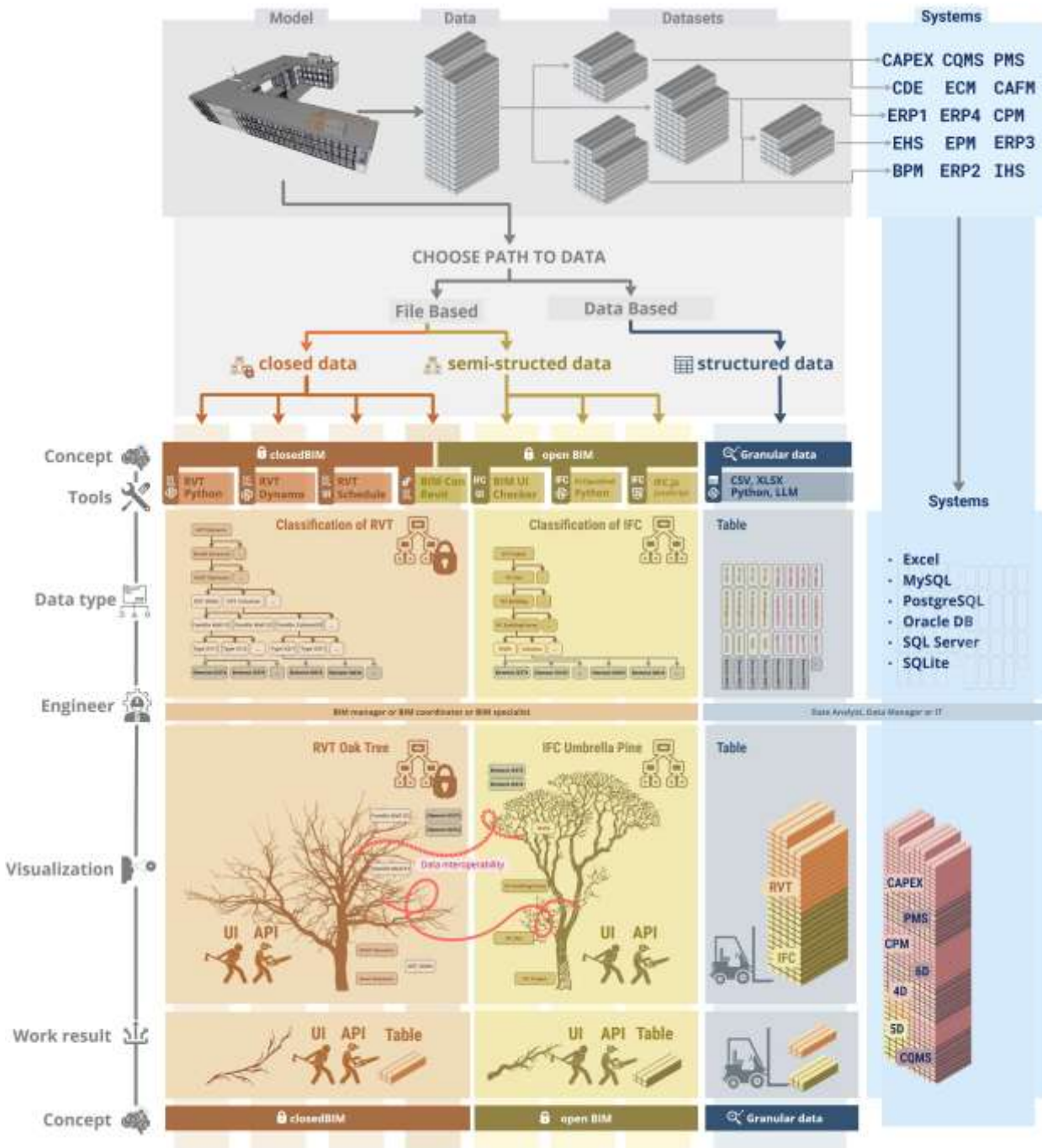
או כלים של הספק API יכולים לגשת לנתונים הנייטיביים דרך חיבורי CAD (BIM)-איור 4.110 אנשי מקצוע ב.

המידע על תכונות ואטריבוטים של רכיבי הפרויקט נאסף במערכת סיווג היררכית (BIM) CAD בפורמטים מיוחדים של שבה ישויות עם תכונות מתאימות ממוקמות, בדומה לפירות של עץ פרי, בקצות הענפים האחרונים של סיווג הנתונים. - (איור 4.111)

חילוף נתונים מהיררכיות דומות אפשרי בשני דרכים: או ידנית, על ידי לחיצה על כל צומת, כאילו מעבדים עץ, קוטעים מציעה גישה - (API) את הענפים הנבחרים של הקטגוריות והסוגים. האפשרות האלטרנטיבית - שימוש בממשקי תכנה יעילה יותר, אוטומטית, להשגת ולקיבוץ נתונים, והופכת אותם בסופו של דבר לטבלה מסודרת לשימוש במערכות אחרות.

Dynamo, pyRvt, ניתן להשתמש בכלים שונים, כגון CAD (BIM) לחילוף טבלאות נתונים מסודרות מפרויקטי IFC. לפורמט IFCjs או IfcOpSh או פתרונות קוד פתוח, כגון, ACC, Pandamo (Pandas + Dynamo).

CAD כלים מודרניים לייצוא ולהמרת נתונים מאפשרים, כדי לפשט את עיבוד והכנת הנתונים, להפריד את תוכן מודלי לשני מרכיבים מרכזיים: מידע גיאומטרי ונתונים אטריבוטיביים (איור 4.113 -) מטא-מידע שמתאר את תכונות רכיבי הבניין (איור 3.116). (שני שכבות הנתונים הללו נשאות מקושרות זו לזו באמצעות מזהים ייחודיים, המאפשרים, התאמה מדויקת של כל רכיב עם תיאור הגיאומטריה) דרך פרמטרים או פוליגונים (עם האטריבוטים שלו: שם, חומר, שלב ביצוע, עלות ועוד. גישה זו מבטיחה את שלמות המודל ומאפשרת שימוש גמיש בנתונים הן עבור ויזואליזציה) נתוני הגיאומטריה של המודל (והן עבור משימות אנליטיות או ניהוליות) מסודרות או חלשות, (תוך עבודה עם שני סוגי הנתונים. -בנפרד או במקביל

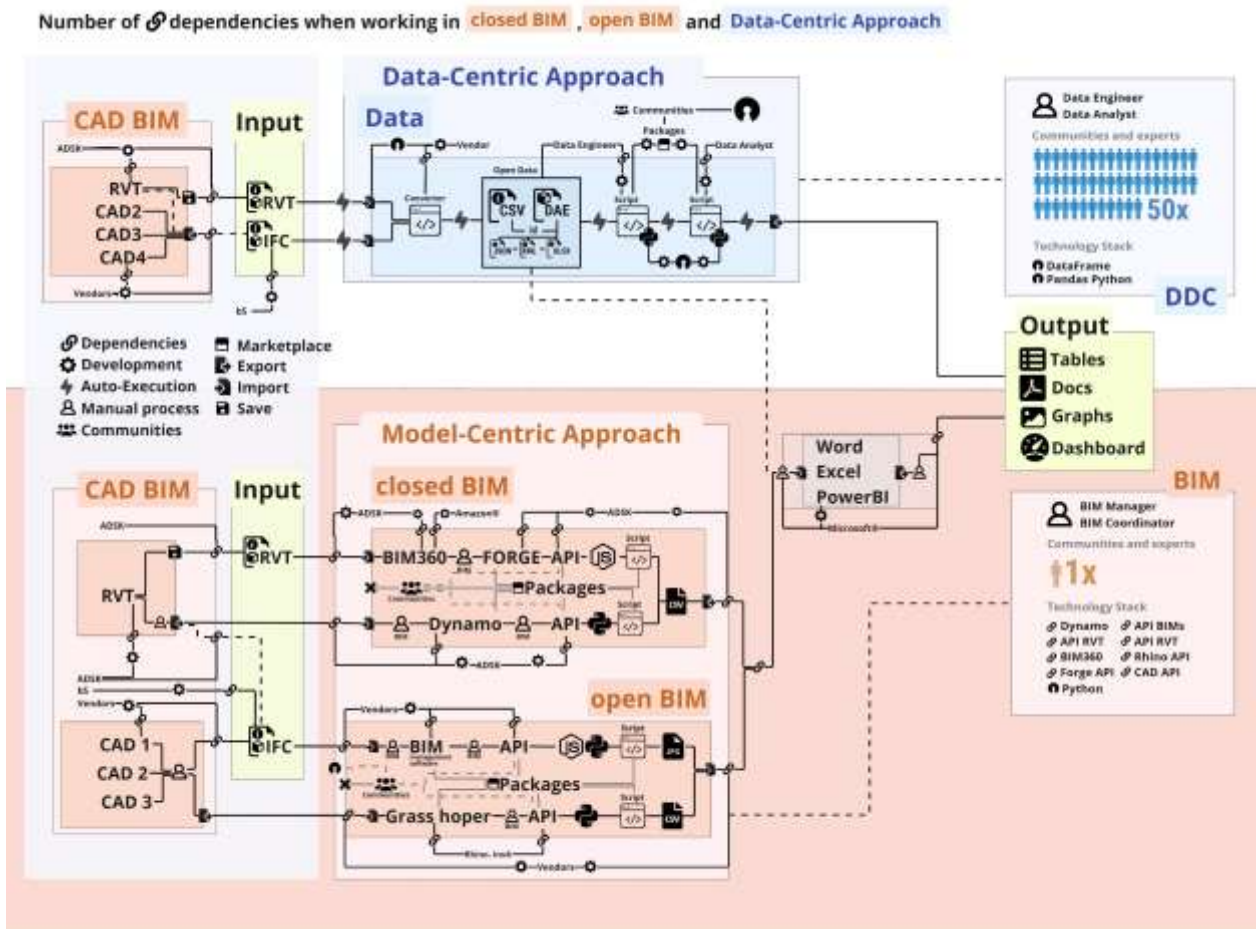


4.111 מוצג למשתמש בצורה של עצי סיווג **CAD (BIM)** איור סוג המידע ממסדי נתונים של

SDK (Software Development Kit) עם התפתחות טכנולוגיות ההנדסה ההפוכה והופעת ערכות פיתוח תוכנה הפכו לפשוטות הרבה יותר CAD (BIM) זמינות והמרת נתונים מפורמטים סגורים של תוכנות - CAD להמרת נתוני כיום קיימת אפשרות להמיר באופן חוקי ובטוח נתונים מפורמטים סגורים לפורמטים אוניברסליים, המתאימים לניתוח והמאבק על ("Open DWG") ולשימוש במערכות אחרות. על ההיסטוריה של הופעת הכלים הראשונים להנדסה הפוכה "של ספקים דיברנו בפרק "נתונים מסודרים: יסוד הדיגיטליזציה CAD שליטה בפורמטים של

כלים להנדסה הפוכה מאפשרים להשיג באופן חוקי נתונים מפורמטים סגורים, מפרקים את המידע מפורמט מעורב לסוגי נתונים ופורמטים הנדרשים למשתמש, ומקלים על עיבוד וניתוחם (BIM) CAD.

עושה את המידע לזמין, ומאפשר שימוש בנתונים CAD השימוש בהנדסה הפוכה וגישה ישירה למידע ממסדי נתונים של פתוחים ובכלים פתוחים, כמו גם ניתוח נתונים באמצעות כלים סטנדרטיים, בניית דוחות, ויזואליזציות ואינטגרציה עם מערכות דיגיטליות אחרות) איור 4.112.



מאפשרת למזער את התלות בפורמט תוכנה ולעבור לגישה מונחת נתונים CAD איור 4.112 גישה ישירה לנתונים.

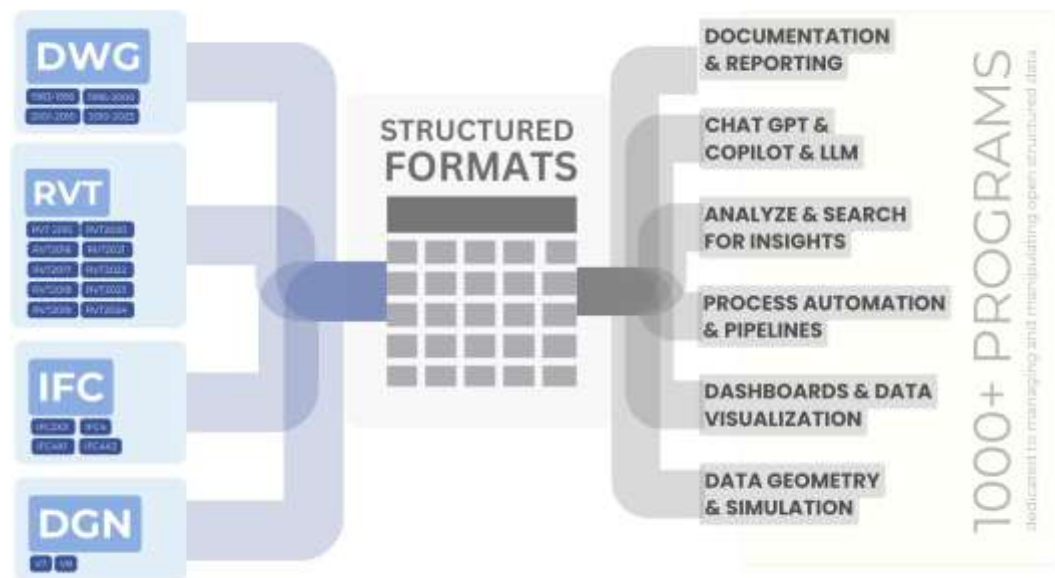
הפך אפשרי, פורמטים RVT ומאז 2018 עבור DGN מאז 2008 עבור פורמט DWG, מאז 1996 עבור פורמט להמיר בנוחות וביעילות לכל פורמט אחר, כולל פורמטים מסודרים, באמצעות כלים להנדסה, CAD סגורים של נתוני הגדולות וחברות ההנדסה הגדולות בעולם משתמשות בכלי (BIM) CAD הפוכה) איור 4.113. (כיום כמעט כל חברות [92].- CAD (BIM) להנדסה הפוכה כדי לחלץ נתונים מפורמטים סגורים של ספקי SDK



לכל מודל נתונים נוח **CAD (BIM)** השימוש בכלי הנדסה הפוכה מאפשר להמיר את בסיסי הנתונים של תוכנות **4.113** איור.

לנתוני **CAD (BIM)** המרת נתונים מפורמטים סגורים ובעלים לפורמטים פתוחים והפרדת פורמטים מעורבים של אטריבוט גיאומטריים ומטא-מידע מקלה על תהליך העבודה עם הנתונים, ומאפשרת ניתוח, מניפולציה ואינטגרציה עם **4.114** איור) מערכות אחרות.

CAD הגענו לרמה שבה אין צורך לבקש אישור מספקי **CAD (BIM)** בעבודה המודרנית עם נתוני **CAD**. כדי לגשת למידע מפורמטים של **(BIM)**



CAD (BIM) מאפשרים להמיר באופן חוקי נתונים מפורמטים בעלים של בסיסי נתונים **SDK** כלים מודרניים של **4.114** איור.

CAD, פרויקטים ממשיכות להתפתח תחת השפעת שחקנים מרכזיים בשוק – ספקי **CAD** מגמות מודרניות בעיבוד נתוני העובדים על חיזוק מעמדם בעולם הנתונים ויוצרים פורמטים וקונספטים חדשים.

עוברים לנתונים מובנים CAD ספקי פתרונות

החל משנת 2024 מתרחש שינוי טכנולוגי משמעותי בתחום התכנון והבנייה בשימוש ובעיבוד נתונים. במקום גישה שפותחה בשנת BIM מתמקדים בקידום קונספטים חדשים. גישות כמו CAD חופשית לנתוני פרויקטים, יצרני מערכות (שפותחה בשנת 2012) (מפנות בהדרגה את מקומן לפתרונות טכנולוגיים מודרניים, המתחילים open BIM ו-2002) [93]: CAD להתקדם על ידי ספקי

■ המעבר לשימוש בנתונים "גרנולריים", המאפשרים ניהול מידע יעיל ומעבר לניתוח נתונים

■ לארגון גמיש של נתונים Entity-component-system (ECS) והטמעת גישת USD הופעת פורמט

■ שימוש פעיל בבינה מלאכותית בעיבוד נתונים, אוטומציה של תהליכים וניתוח נתונים

■ פיתוח אינטרופראביליות - שיפור האינטראקציה בין תוכניות, מערכות ובסיסי נתונים שונים

שיווק, מציאות ועתיד נתוני פרויקטים: CAD ו-BIM" כל אחד מהאספקטים הללו יידון בפירוט בחלק השישי של הספר הגדולים שואפים כיום לחשוב מחדש על CAD בבנייה. "בפרק זה נציין בקצרה את הכיוון הכללי של השינויים: ספקי דרכי מבנה המידע הפרויקטלי. אחד השינויים המרכזיים הוא הנטישה של מודל אחסון קבצים קלאסי לטובת ארכיטקטורת [93] נתונים גרנולרית, הממוקדת בניתוח, המבטיחה גישה מתמדת לרכיבים נפרדים של המודל

מהות המתרחש היא שהתעשייה מתנערת בהדרגה מפורמטים מגושמים, מיוחדים ופרמטריים, המצריכים שימוש בליבות גיאומטריות, לטובת פתרונות יותר אוניברסליים, קריאים למכונה וגמישים.

(תיאור סצנה אוניברסלי, שפותח במקור בתעשיית הגרפיקה הממוחשבת USD אחד מהדחפים לשינויים היה הפורמט Isaac Sim וגם) NVIDIA Omniverse אך כבר זכה להכרה גם ביישומים הנדסיים, בזכות הפיתוח של פלטפורמת מציע מבנה פשוט יותר, מאפשר לתאר גיאומטריה USD, הפרמטרי IFC-עבור סימולציות וויזואליזציות. בניגוד למה שמקל על עיבוד המידע ומאיץ את האינטגרציה שלו בתהליכים דיגיטליים, JSON ותכונות של אובייקטים בפורמט בצורה של (פרטים נוספים בחלק 6 של הספר – BREP-NURBS מלבד) הפורמט החדש מאפשר לאחסן גיאומטריה מה שהופך אותו לנוח יותר עבור תהליכים אוטומטיים ועבודה, JSON-ותכונות אובייקטים – ב, MESH פוליגוני -במערכות אקולוגיות בענן

אך רובם נשארים, (NWD, SVF, CP2, CPXML, למשל) כבר משתמשים בפורמטים דומים CAD ו-ERP כמה ספקי סגורים ואינם זמינים לשימוש חיצוני, מה שמגביל את אפשרויות האינטגרציה והשימוש החוזר בנתונים. בהקשר זה DWG בזמנו – אלטרנטיבה פתוחה לפורמטים קנייניים, כמו DXF יכול למלא את אותה תפקיד כמו USD

General Information 				Comparison / Notes
Year of format creation	1991	2016	IFC focuses on construction data, USD on 3D graphics	
Creator-developer	TU Munich	Pixar	IFC was founded in Germany, USD in America	
Prototypes and predecessors	IGES, STEP	PTEX, DAE, GLTF	IFC evolved from IGES/STEP, USD from PTEX/DAE/GLTF	
Initiator in Construction	ADSK	ADSK	ADSK initiated the adoption of both formats in construction	
Organizer of the Alliance	ADSK	ADSK	ADSK organized both alliances	
Name of the Alliance	bS (IAI)	AOUSD	Different alliances for each format	
Year of Alliance Formation	1994	2023	The IFC alliance was formed in 1994, AOUSD for USD in 2023	
Promoting in the construction	ADSK and Co	ADSK and Co	ADSK and Co actively promotes both formats in bS (IAI) since the introduction	

Purpose and Usage 				Comparison / Notes
Purpose	Semantic description and interoperability	Data simplification, visualization unification	IFC for semantics and exchange; USD for simplification and visualization	
Goals and Objectives	Interoperability and semantics	Unification for visualization and data processing	IFC focuses on semantics; USD on visualization	
Use in Other Industries	Predominantly in construction	In film, games, VR/AR, and now in construction	USD is versatile and used in various fields	
Supported Data Types	Geometry, object attributes, metadata	Geometry, shaders, animation, light, and camera	USD supports a wider range of data types suitable for complex visualizations; IFC focuses on construction-specific data	

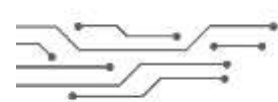
- לענות על הביקוש לאינטרופראביליות ועצמאות נתוני פרויקטים מהליבות הגיאומטריות CAD פורמט, כניסיון של ספקי USD

בניגוד ל) USD, GLTF, OBJ, XML המעבר של המפתחים הגדולים לפורמטים פתוחים ומפשטים כמו ופורמטים דומים משקף מגמה גלובלית וביקוש בתעשייה (NWD, CP2, SVF, SVF2, CPIXML) לפשט את הנתונים ולהגביר את זמינותם. בשנים הקרובות ניתן לצפות לעזיבה הדרגתית של תקנים ופורמטים פרמטריים מורכבים התלויים בליבות גיאומטריות לטובת פתרונות קלים ומסודרים יותר. מעבר - זה יאיץ את הדיגיטציה של תעשיית הבניין, יקל על אוטומציה של תהליכים ויפשט את חילופי הנתונים

לקידום פורמטים פתוחים חדשים, אנשי מקצוע בתעשיית הבניין יכולים CAD למרות התוכניות האסטרטגיות של ספקי בעזרת כלים, CAD (BIM) סגורות, מבלי להזדקק לשימוש בכלים של CAD גם לקבל גישה מלאה לנתונים ממערכות של הנדסה הפוכה.

כל המגמות הללו מובילות בהכרח למעבר ממודלים תלת-ממדיים מסורבלים ומונוליטיים לנתונים אוניברסליים ומסודרים ולשימוש בפורמטים שכבר הוכיחו את עצמם בתעשיות אחרות. ברגע שהצוותים לא רק אובייקטים ויזואליים או קבוצות קבצים, אלא כבסיסי CAD הפרויקטים מתחילים לראות במודלי נתונים המכילים ידע ומידע – הגישה לתכנון ולניהול משתנה באופן דרמטי

וקיבלו גישה לבסיסי נתונים, CAD לאחר שהצוותים למדו לחלץ נתונים מסודרים ממסמכים, טקסטים, שרטוטים ומודלי השלב הקריטי הבא הוא מודל הנתונים והבטחת איכותם. שלב זה משפיע במידה רבה על מהירות העיבוד והטרנספורמציה של המידע, אשר בסופו של דבר ישמש לקבלת החלטות במשימות יישומיות ספציפיות



פרק 4.2. סיווג ואינטגרציה: שפה אחידה לנתוני בנייה

מהירות קבלת החלטות תלויה באיכות הנתונים

האדריכלות המודרנית של נתוני פרויקטים עוברת שינויים יסודיים. הענף עובר ממודלים מסורבלים ומבודדים ופורמטים סגורים לעבר מבנים גמישים יותר, קריאים למכונה, הממוקדים בניתוח, אינטגרציה ואוטומציה של תהליכים. עם זאת, המעבר לפורמטים חדשים בפני עצמו אינו מבטיח יעילות - במרכז תשומת הלב נמצא בהכרח איכות הנתונים עצמם.

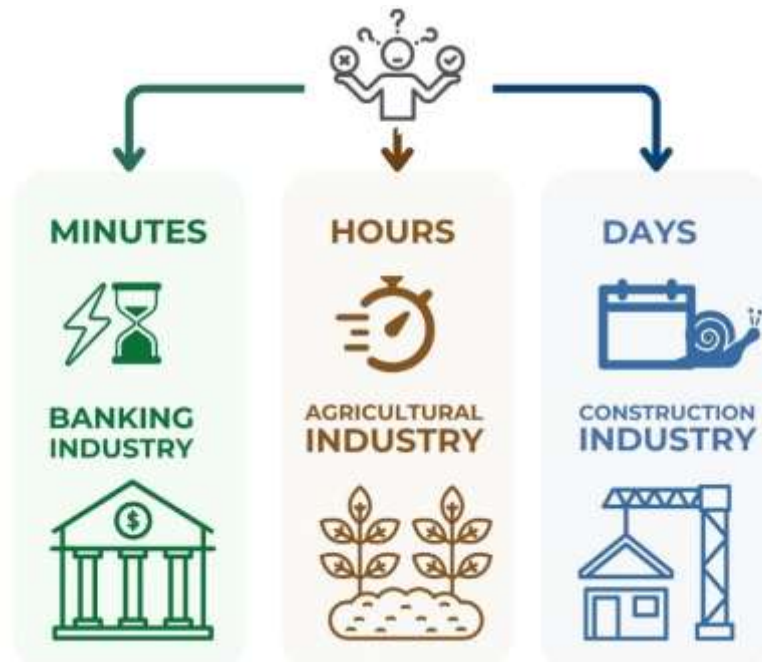
בעמודים של ספר זה אנו מדברים רבות על פורמטים, מערכות ותהליכים. אך כל המאמצים הללו מאבדים את משמעותם ללא רכיב מפתח אחד - נתונים שניתן לסמוך עליהם. איכות הנתונים היא אבן הפינה של הדיגיטציה, אליה נחזור בכל החלקים הבאים.

חברות הבנייה המודרניות - במיוחד הגדולות - משתמשות בעשרות, ולעיתים אלפי מערכות ומסדי נתונים שונים. מערכות אלו צריכות לא רק להתמלא באופן קבוע במידע חדש, אלא גם לפעול בצורה יעילה זו עם זו. כל הנתונים החדשים הנוצרים כתוצאה מעיבוד המידע המתקבל, משתלבים בסביבות אלו ומשמשים לפתרון בעיות עסקיות ספציפיות.

על בסיס - HiPPO ואם בעבר ההחלטות לגבי בעיות עסקיות ספציפיות התקבלו על ידי מנהיגים בכירים - מה שנקרא ניסיון ואינטואיציה, כיום, בתנאים של גידול חד בנפח המידע, גישה זו הופכת למעוררת מחלוקות. במקומה מגיעה אנליטיקה אוטומטית, הפועלת עם נתונים בזמן אמת.

הדיון "המסורתי-ידני" על תהליכים עסקיים ברמת המנהיגים יוסט לעבר אנליטיקה תפעולית, אשר דורשת תגובות מהירות לבקשות עסקיות.

העידן שבו רואי חשבון, מנהלי פרויקטים ומעריכי עלויות יצרו ידנית דוחות וטבלאות סיכום, כמו גם תצוגות נתונים על פרויקטים במשך מספר ימים ושבועות, הולך ונעלם. כיום, מהירות וזמן קבלת ההחלטות הופכים לגורם מפתח במיתרון תחרותי.



בתעשיית הבנייה, חישובים והחלטות לוקחים מספר ימים, בניגוד לתעשיות אחרות, שבהן זה קורה בשעות או דקות.

ההבדל העיקרי בין תעשיית הבנייה לתעשיות אחרות המתקדמות יותר מבחינה דיגיטלית הוא ברמת האיכות והסטנדרטיזציה הנמוכה של הנתונים. גישות מיושנות ליצירה, העברה ועיבוד מידע מאטות את התהליכים ויוצרות כאוס. חוסר בסטנדרטים אחידים לאיכות הנתונים מונע את יישום האוטומציה המקיפה.

אחת הבעיות המרכזיות היא איכות הנתונים המקוריים הנמוכה, כמו גם חוסר בתהליכים פורמליים להכנה ובדיקה שלהם. ללא נתונים מהימנים ומוסכמים, אינטגרציה יעילה בין המערכות אינה אפשרית. זה מוביל לעיכובים, שגיאות והגדלת עלויות בכל שלב מחיי הפרויקט.

בחלקים הבאים של הספר נבחן בפירוט כיצד ניתן לשפר את איכות הנתונים, לסטנדרטיזציה של תהליכים ולצמצם את הדרך מקבלת מידע לנתונים איכותיים, מאומתים ומוסכמים.

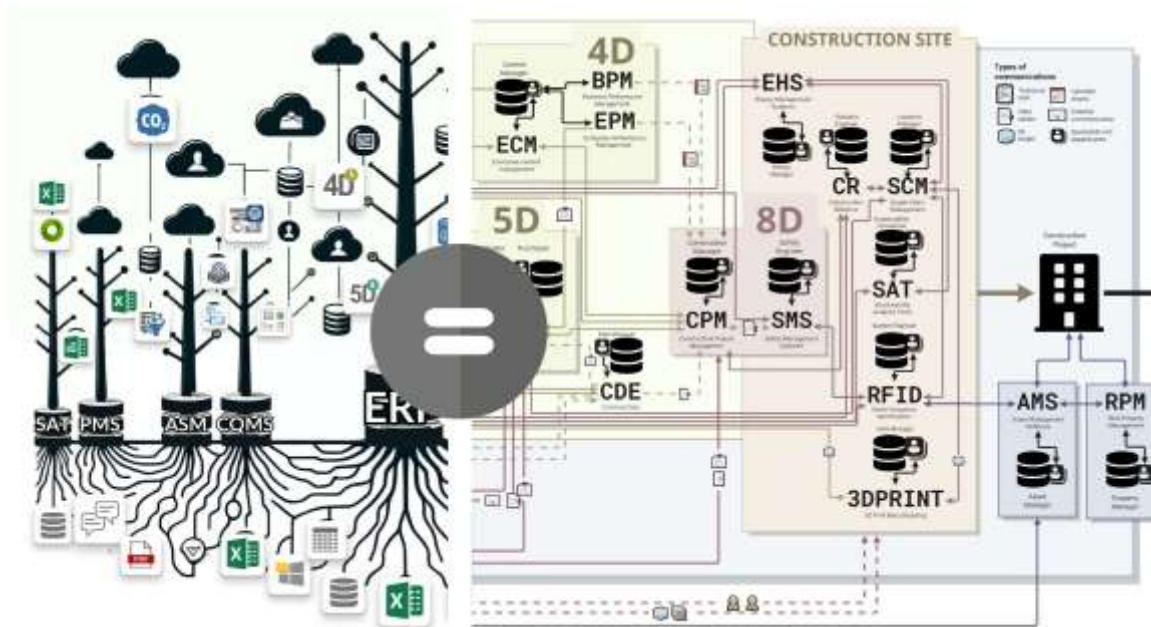
סטנדרטיזציה ואינטגרציה של נתונים

עבודה יעילה עם נתונים דורשת אסטרטגיה ברורה לסטנדרטיזציה. רק כאשר יש דרישות ברורות לגבי המבנה ואיכות הנתונים ניתן לאוטומט את הבדיקה שלהם, לצמצם את מספר הפעולות הידניות ולהאיץ את קבלת ההחלטות המושכלות בכל שלב של הפרויקט.

נתונים, CAD, קבצי PDF, בפרקטיקה היומיומית של חברת הבנייה, יש צורך לעבד מאות קבצים מדי יום: מיילים, מסמכי אשר יש לשלבם בתהליכי העבודה של החברה, IoT, מחיישני

צריך ללמוד כיצד להפיק חומרים מזינים, (4.22 איור) היער האקולוגי של החברה, המורכב ממסדי נתונים וכלים מנתונים המגיעים בפורמטים שונים, כדי להשיג את התוצאות הנדרשות לחברה.

כדי להתמודד ביעילות עם זרם הנתונים, אין צורך להעסיק צבא של מנהלים, בראש ובראשונה יש לפתח דרישות וסטנדרטים מחמירים לנתונים ולהשתמש בכלים מתאימים לבדיקה אוטומטית, אחידות ועיבוד שלהם.



הבטחת חיים בריאים של האקולוגיה של החברה דורשת אספקה איכותית ובזמן של משאבים למערכות שלה 4.22 איור.

כדי לאוטומט את תהליך בדיקת הנתונים ואחידותם (לצורך אינטגרציה אוטומטית לאחר מכן, יש להתחיל בתיאור ההדרישות המינימליות הנדרשות לנתונים עבור כל מערכת ספציפית. דרישות אלו מגדירות:

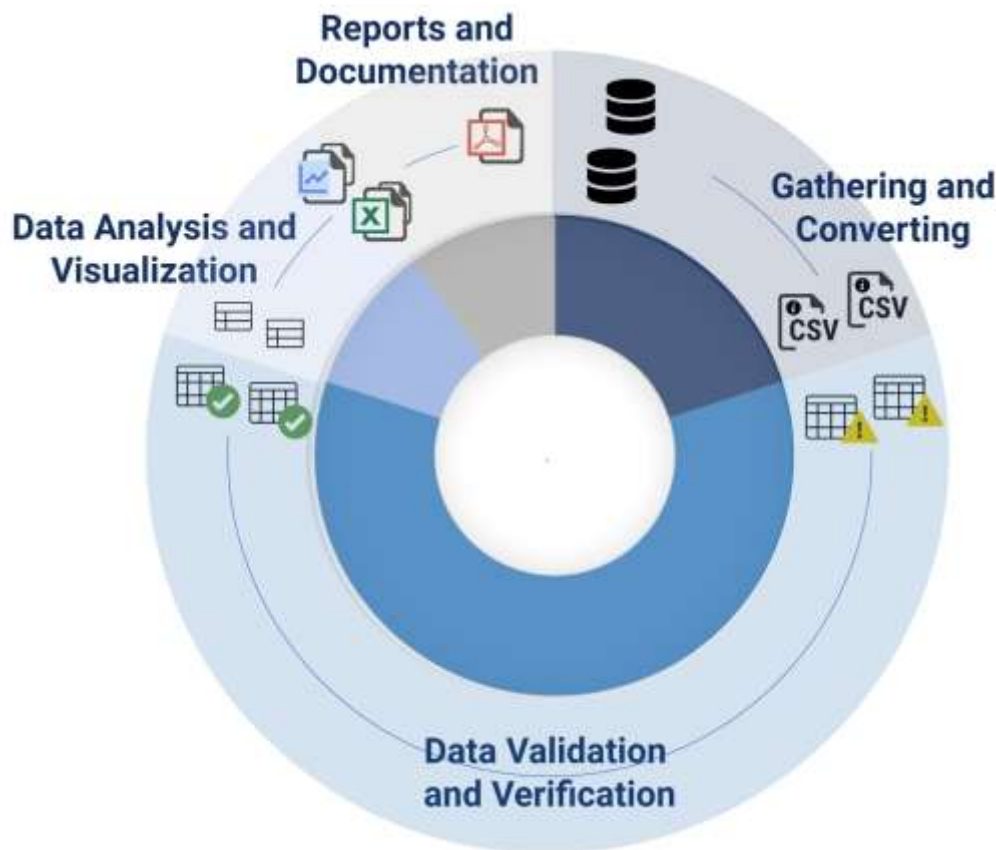
- מה בדיוק יש להשיג
- (באיזה פורמט) מבנה, פורמט)
- אילו מאפיינים הם חובה
- אילו סבירות של דיוק ושלמות הן מקובלות

דרישות הנתונים מתארות את קריטריוני האיכות, המבנה והשלמות של המידע המתקבל ומעובד. לדוגמה, עבור טקסטים אובייקטים (7.216 איור - 7.214 איור) חשוב שהעיצוב יהיה מדויק בהתאם לסטנדרטים של התעשייה, PDF במסמכי (7.310 איור, 7.39 איור) צריכים להיות בעלי מאפיינים נכונים (מידות, קודים, חיבורים למדריכים) CAD במודלים ועבור סריקות של חוזים, חשוב שיהיו תאריכים ברורים ואפשרות לחילוץ אוטומטי של הסכום והתנאים המרכזיים (איור 4.110 - 4.17).

ניסוח דרישות לנתונים ובדיקת התאמתם באופן אוטומטי הוא אחד השלבים המורכבים ביותר, אך קריטיים ביותר. זהו השלב שלרוב תופס את רוב הזמן בתהליכי העבודה.

מתבזבז (BI) כפי שהוזכר בחלק השלישי של הספר, בין 50% ל-90% מהזמן של אנשי מקצוע בתחום ניתוח הנתונים תהליך זה כולל איסוף, אימות, ולידציה, אחידות ומבנה של (3.25 איור) לא על ניתוח, אלא על הכנת הנתונים הנתונים.

לפי סקר שנערך בשנת 2016 [95], אנשי מקצוע בתחום עיבוד הנתונים במגוון רחב של תחומים דיווחו כי רוב הזמן לאסוף מערכות נתונים קיימות ולארגן: (4.23 איור) שלהם כ-80% (מתבזבז על מה שהם הכי פחות אוהבים לעשות) לאחד, למבנה (אותן). כך, פחות מ-20% מהזמן שלהם נשאר למשימות יצירתיות, כמו חיפוש דפוסים ומגמות, שיביאו לתובנות וגילויים חדשים.



בדיקת ואבטחת איכות הנתונים היא השלב היקר ביותר, הארוך והמורכב ביותר בהכנת הנתונים לאינטגרציה 4.23 איור במערכות אחרות.

ניהול מוצלח של נתונים בחברת בנייה דורש גישה מקיפה, הכוללת פרמטריזציה של משימות, ניסוח דרישות לאיכות הנתונים ושימוש בכלים מתאימים לבדיקה אוטומטית שלהם.

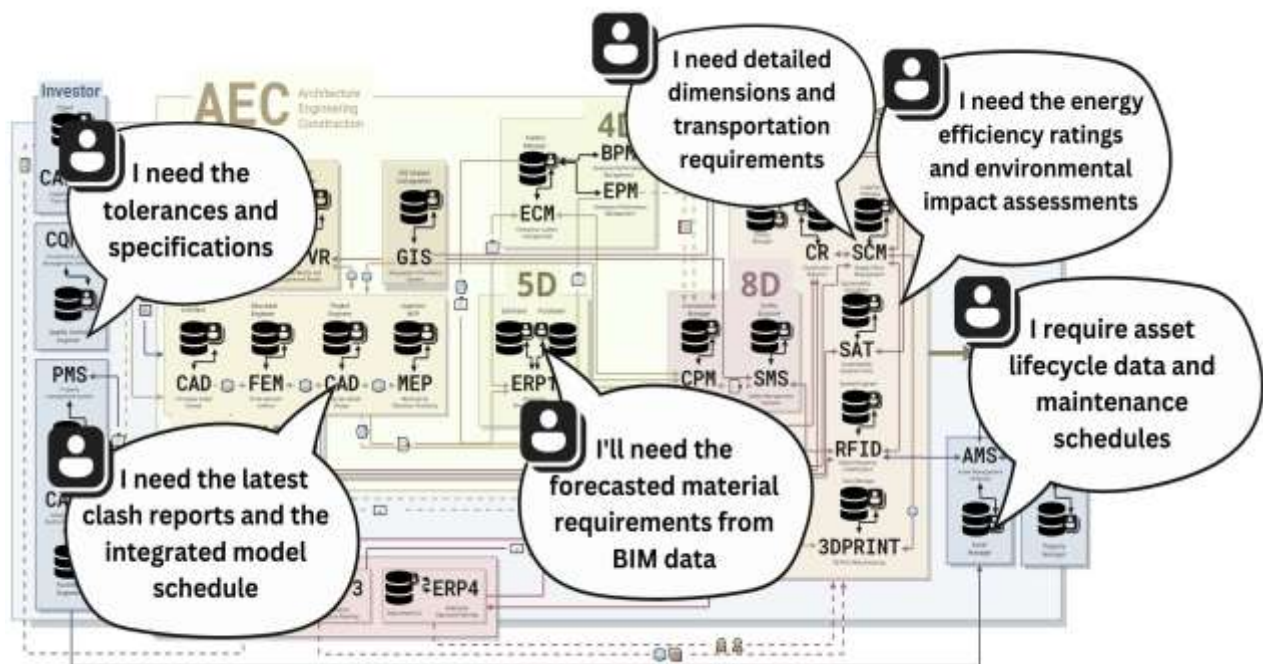
תאימות דיגיטלית מתחילה בדרישות

עם העלייה במספר המערכות הדיגיטליות בתוך החברות, עולה גם הצורך בהתאמה של הנתונים ביניהן. מנהלים האחראים, שונות לעיתים קרובות נתקלים בקושי לעמוד בקצב הגובר של המידע ומגוון הפורמטים. בתנאים אלו IT על מערכות הם נאלצים לבקש מהמומחים ליצור נתונים בצורה המתאימה לשימוש באפליקציות ופלטפורמות אחרות.

זה, בתורו, דורש מהמהנדסים ועובדים העוסקים ביצירת נתונים, להתאים את עצמם למגוון דרישות, לעיתים קרובות ללא שקיפות והבנה ברורה של היכן וכיצד הנתונים הללו ייושמו בהמשך. חוסר בגישות סטנדרטיות לעבודה עם מידע מביא לאובדן יעילות ולהגברת העלויות בשלב הבדיקה, אשר לעיתים קרובות, בשל המורכבות והחוסר בסטנדרטיזציה של הנתונים, מתבצעת ידנית.

שאלת הסטנדרטיזציה של הנתונים אינה רק שאלה של נוחות או אוטומציה. מדובר בהפסדים כספיים משנת 2016, ההפסדים השנתיים עקב איכות נתונים נמוכה בארצות הברית IBM ישרים. לפי דוח של וחברות ייעוץ אנליטיות אחרות מראים כי עלות MIT מסתכמים ב-3.1 טריליון דולר. בנוסף, מחקרים של איכות נתונים נמוכה יכולה להגיע ל-15%–25 מההכנסות של החברה.

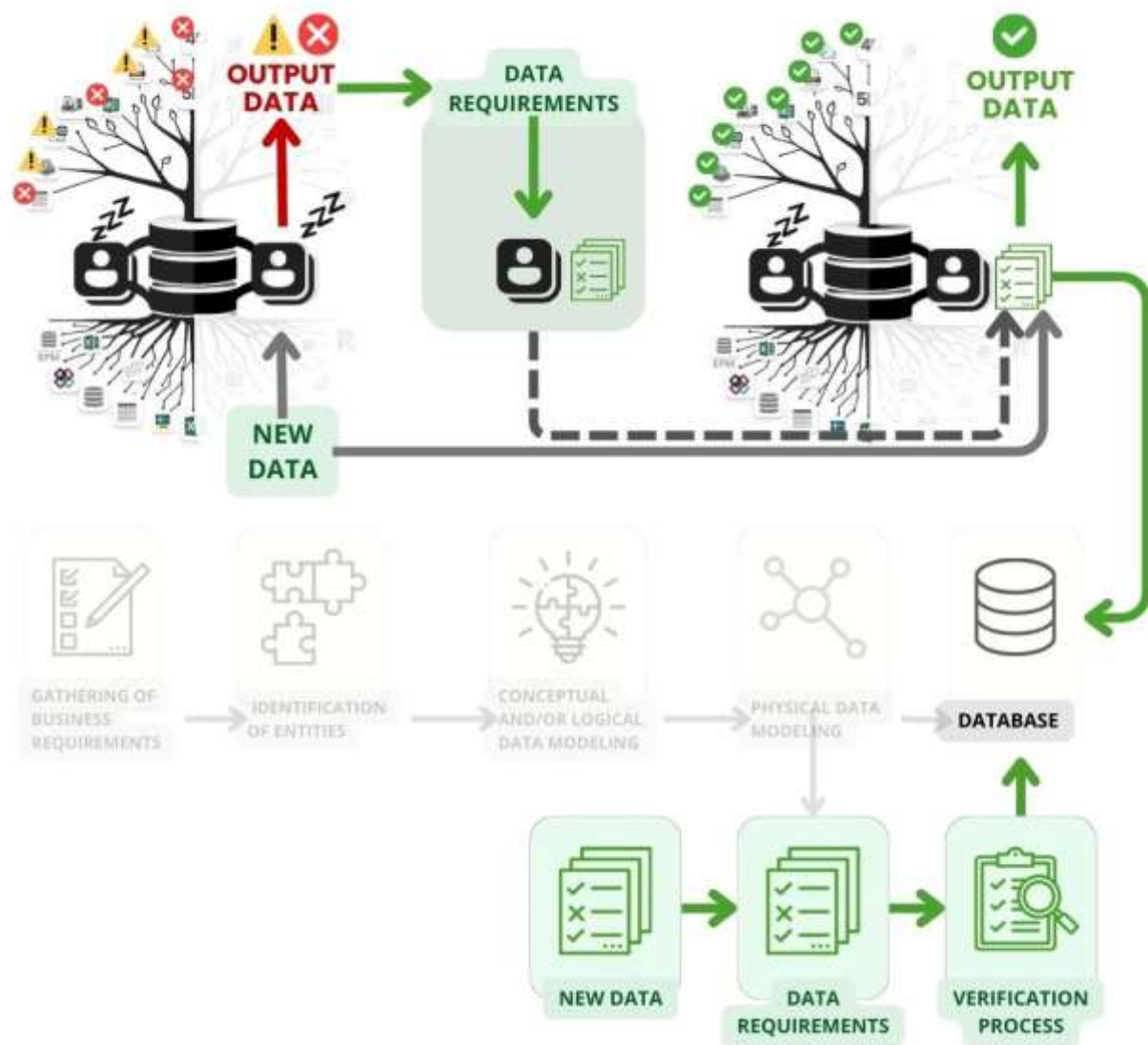
בתנאים אלו, הופך להיות קריטי להחזיק בדרישות ברורות לנתונים ובתיאורים של אילו פרמטרים, באיזה פורמט ובאיזה רמת פירוט יש לכלול באובייקטים הנוצרים. ללא פורמליזציה של דרישות אלו, אין אפשרות להבטיח את איכות הנתונים ואת התאמתם בין המערכות ובין שלבי הפרויקט.



עסק מבוסס על אינטראקציה בין תפקידים שונים, כאשר כל אחד מהם דורש פרמטרים וערכים מסוימים, קריטיים לביצוע משימות עסקיות.

כדי לנסח דרישות נכונות לנתונים, יש להבין את תהליכי העסקים ברמת הנתונים. פרויקטים בנייתיים משתנים לפי סוג חישוב, (ERP 4D) תכנון לוח זמנים, (CAD (BIM)) היקף ומספר המשתתפים, וכל מערכת – בין אם מדובר במידול דורשת עבור הישויות הנכנסות שלה פרמטרים ייחודיים משלה – (SCM) או לוגיסטיקה (ERP 5D) עלויות

בהתאם לצרכים אלו, מנהלי העסקים צריכים או לתכנן מבני נתונים חדשים בהתחשב בדרישות שנקבעו, או להתאים טבלאות ומסדי נתונים קיימים. איכות הנתונים הנוצרים תהיה תלויה ישירות בכמה מדויקת ונכונה נוסחו הדרישות.



איכות הנתונים תלויה באיכות הדרישות שנוצרות עבור מקרים ספציפיים של שימוש בנתונים.

מכיוון שכל מערכת מציבה את הדרישות הספציפיות שלה לנתונים, הצעד הראשון בניסוח דרישות כלליות צריך להיות קטגוריזציה של כל האלמנטים המשתתפים בתהליכי העסקים. זה אומר שיש צורך בחלוקה של אובייקטים לקטגוריות וקבוצות קטגוריות, המתאימות למערכות או למשימות יישומיות ספציפיות. עבור כל קבוצה כזו, יש לפתח דרישות נפרדות לגבי המבנה, האטריבוטים ואיכות הנתונים.

עם זאת, במציאות, יישום גישה זו נתקל בקושי משמעותי: חוסר בשפה אחידה לקיבוץ נתונים. סיווגים מפוצלים, שכפול מזהים וחוסר תאימות בפורמטים מביאים לכך שכל חברה, כל תוכנה ואפילו כל פרויקט יוצרים מודלים ודירוגים מבודדים משלהם. כתוצאה מכך נוצר מגדל בבל דיגיטלי, שבו להעברת מידע בין מערכות נדרשים המרות רבות למודלים ודירוגים הנדרשים, לעיתים קרובות מבוצעות ידנית. ניתן להתגבר על מחסום זה רק על ידי מעבר למדריכים אוניברסליים וסטנדרטים של דרישות.

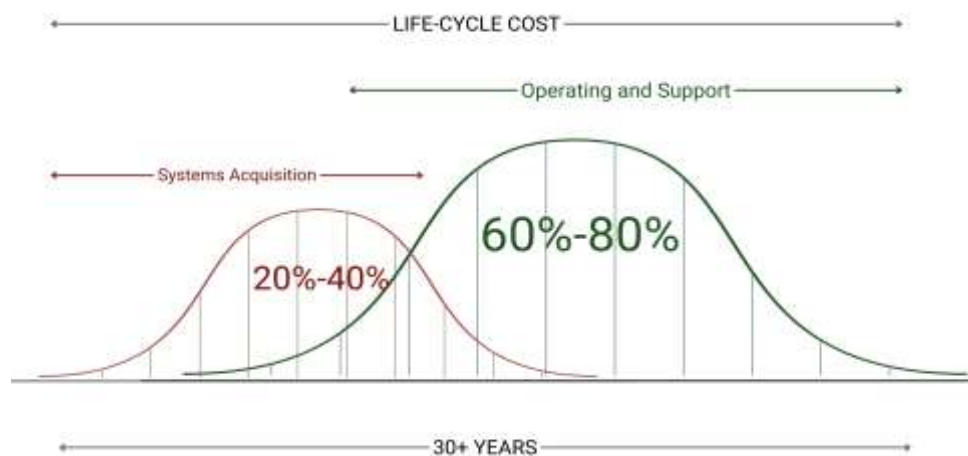
שפה אחידה לבנייה: תפקיד המדרגים בטרנספורמציה דיגיטלית

בהקשר של דיגיטציה ואוטומציה של תהליכי בדיקה ועיבוד, תפקיד מיוחד ממלאות מערכות סיווג של רכיבים - מעין מילונים דיגיטליים, "המבטיחים אחדות בתיאור ובפרמטריזציה של אובייקטים. בדיוק המדריכים יוצרים את "השפה" האחידה "המאפשרת לקבץ נתונים לפי משמעות השימוש ולשלב נתונים בין מערכות שונות, רמות ניהול ושילובי מחזור חיי הפרויקט.

ההשפעה המורגשת ביותר של המדריכים היא בכלכלה של מחזור חיי הבניין, שבה ההיבט החשוב ביותר הוא אופטימיזציה של עלויות תפעול ארוכות טווח. מחקרים מראים כי הוצאות תפעול מהוות עד 80% מהעלות הכוללת של בעלות על בניין, מה שמכפיל את העלויות הראשוניות של הבנייה (איור 4.26) [98]. [משמעות הדבר היא שההחלטה על הוצאות - עתידיות מתעצב במידה רבה כבר בשלב התכנון

צריכות להיות נקודת המוצא בעת (CAFM, AMS, PMS, RPM) לכן, הדרישות מהנדסי התפעול גיבוש הדרישות לנתונים בשלב התכנון (איור 1.24). (מערכות אלו צריכות להיחשב לא כשלב סופי של - הפרויקט, אלא כחלק בלתי נפרד מהאקוסיסטמה הדיגיטלית של הפרויקט - מהקונספט ועד לפרוק

מדריך מודרני אינו רק מערכת קודים לקיבוץ. זהו מנגנון הבנה הדדית בין אדריכלים, מהנדסים, מערכי עלויות בדומה לכך שהטייס האוטומטי של רכב צריך לזהות באופן חד משמעי IT. לוגיסטיקאים, שירותי תפעול ומערכות אובייקטים בכביש בדיוק גבוה, מערכות הבנייה הדיגיטליות ומשתמשיהן צריכות לפרש באופן חד משמעי את אותו רכיב פרויקט עבור מערכות שונות דרך סוג הרכיב



איור 4.26 הוצאות תפעול וטכניות עלולות על עלות הבנייה פי שלושה, מהוות 60%–80% מכלל הוצאות מחזור חיי הבניין (על פי [99] חומרים -).

רמת הפיתוח של המדריכים מתוארת ישירות עם עומק הדיגיטציה של החברה ובשלב הדיגיטלי שלה. ארגונים עם רמת דיגיטציה נמוכה מתמודדים עם פיצול נתונים, חוסר תאימות של מערכות מידע וככתוצאה מכך, חוסר תאימות וחוסר יעילות של המדריכים. בחברות כאלה, אותו רכיב יכול לעיתים קרובות להיות בעל מזהים שונים לקיבוץ במערכות שונות, מה שמקשה באופן קריטי על האינטגרציה הסופית ומקשה על אוטומציה של תהליכים

במערכת התקצוב ובמערכת התפעול (איור CAD, לדוגמה, אותו חלון בפרויקט יכול להיות מסומן בדרכים שונות במודל

- ניתוח חיזוי של סיכונים פוטנציאליים והתנגשויות
- אופטימיזציה של תהליכי רכישה ושרשראות לוגיסטיות
- יצירת דיגיטליים כפולים של בניינים ומבנים
- אינטגרציה עם מערכות "עיר חכמה" ואינטרנט של הדברים

הזמן להמרה מוגבל - עם התפתחות הטכנולוגיות של למידת מכונה וראייה ממוחשבת, הבעיה של סיווג אוטומטי, שלא הצליחה להיפתר במשך עשורים, תיפתר בשנים הקרובות, וחברות בנייה ועיצוב שלא הספיקו להסתגל לכך, מסתכנות לחוות את גורל מוניות השירות שהודרו על ידי פלטפורמות דיגיטליות.

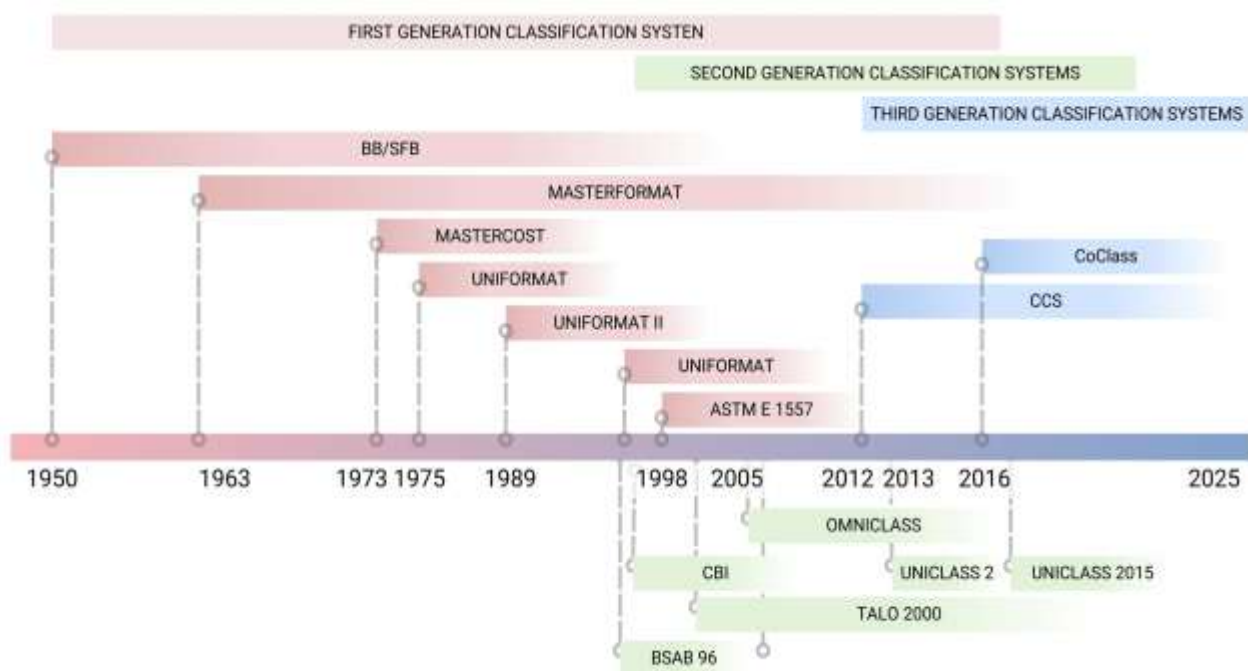
פרטים נוספים על אוטומציה של חישובי עלויות וזמנים, כמו גם על נתונים גדולים ולמידת מכונה, יידונו בחלק החמישי והתשיעי של הספר. שאלות הסיכון לחוות את גורל מוניות השירות והאובריציה של תעשיית הבנייה נבחנות בפירוט בחלק העשירי של הספר.

בהבנת התפקיד המרכזי של מסווגים בהמרה הדיגיטלית של תעשיית הבנייה, יש לפנות להיסטוריה של האבולוציה שלהם. ההקשר ההיסטורי מאפשר להבין כיצד התפתחו הגישות לסיווג ואילו מגמות קובעות את מצבם המודרני.

אבולוציה של מערכות סיווג: CoClass ו-Uniclass, OmniClass, Masterformat

היסטורית, מסווגים של רכיבי בנייה ועבודות התפתחו בשלושה דורות, שכל אחד מהם שיקף את רמת הטכנולוגיות והזמינות ואת הצרכים הרלוונטיים של התעשייה בתקופה מסוימת (איור 4.28):

- הדור הראשון) תחילת שנות ה-50 - סוף שנות ה-80 (-) מדריכים מודפסים, מסווגים היררכיים, שנעשה בהם (Masterformat, Sfb), למשל) שימוש מקומי
- הדור השני) סוף שנות ה-90 - אמצע שנות ה-2010 (-) טבלאות ומסדי נתונים מובנים, המיועדים ליישום ב-Excel ו-Access (ASTM E 1557, OmniClass, Uniclass 1997).
- CAD (BIM), אינטגרציה עם API הדור השלישי) משנות ה-2010 ועד היום (-) שירותים דיגיטליים וממשקי (Uniclass 2015, CoClass). אוטומציה



איור 4.28 שלושה דורות של מסווגים בתעשיית הבנייה.

בעשורים האחרונים נצפתה ירידה במורכבות ההיררכית (איור 4.29) של המסווגים: אם מערכות מוקדמות, כמו CoClass, השתמשו עד 7 רמות של הטמעה לתיאור 6887 קטגוריות, הפתרונות המודרניים, כמו OmniClass, מוגבלים ל-3 רמות עם 750 קטגוריות. זה מפשט את העבודה עם נתונים, תוך שמירה על הפרטים הנדרשים הנמצא בשימוש נרחב בבריטניה כסטנדרט, מאחד 7210 קטגוריות בארבע רמות בלבד, מה שהופך אותו לנוח, ורכישות ממשלתיות CAD עבור פרויקטי

Classifier	Table / Objects	Number of classes	Nesting depth
OmniClass	Table 23 Products	6887	7 levels
Uniclass 2015	Pr — Products	7210	4 levels
CoClass, CCS	Components	750	3 levels

איור 4.29 עם כל דור חדש של מסווגים, מורכבות הקטגוריזציה פוחתת באופן משמעותי.

במערכות הערכת עלויות בנייה במדינות שונות, בשל השוני במסווגים, אפילו רכיב טיפוסי כמו קיר יסוד מבטון יכול להיות מתואר בדרכים שונות לחלוטין (איור 4.210). (ההבדלים הללו משקפים את המאפיינים הלאומיים של הפרקטיקה הבנייתית, את מערכות המדידה המיושמות, את הגישות לסיווג חומרים, כמו גם את הדרישות התקניות והטכניות החלות בכל מדינה.



איור 4.210 אותו רכיב במדינות שונות משמש בפרויקטים דרך תיאורים ומסווגים שונים.

המגוון של סיווגים של אותם אלמנטים מקשה על שיתוף פעולה בינלאומי, הופך את השוואת העלויות והיקף העבודה בפרויקטים בינלאומיים למורכב, ולעיתים אף בלתי אפשרי. נכון לעכשיו, ברמה הגלובלית אין סיווג אוניברסלי אחד. כל מדינה או אזור מפתחים מערכות משלהם, בהתבסס על תקנים מקומיים, שפה ותרבות עסקית.

- (דנמרק): (מערכת סיווג עלויות לאורך כל מחזור חיי הפרויקט) תכנון, בנייה, תפעול. (הדגש הוא על לוגיקת CCS התפעול והתחזוקה, אך היא כוללת גם ניהול תקציבים ומשאבים).
- (נורבגיה): (מסווגת פרויקטים לפי פונקציות, אלמנטים מבניים ושלבי מחזור חיי הפרויקט. משמשת NS 3451 לניהול פרויקטים, הערכת עלויות ותכנון אורך טווח).
- (ארצות הברית): (מערכת לארגון מפרטי הבנייה לפי קטגוריות) למשל: בטון, חשמל, גימור (MasterFormat). (UniFormat-בניגוד ל) מתמקדת בדיסציפלינות ובסוגי עבודות, ולא באלמנטים פונקציונליים.
- (מאחד BIM): (בריטניה): (אחד הסיווגים המפורטים ביותר, מיועד לרכישות ממשלתיות ולפרויקט UniClass 2 נתונים על פרויקטים, עבודות, חומרים ומרחבים במערכת אחת).
- (לניהול מידע על פרויקטים: מספריית רכיבים (בארצות הברית CSI פותח על ידי) תקן בינלאומי: OmniClass ולכלים דיגיטליים אחרים (BIM) CAD-ועד מפרטים אלקטרוניים. מתאים לאחסון נתונים לטווח ארוך, תואם ל).
- (החלפת מידע בנוגע לבניית תפעול - תקן בינלאומי להחלפת נתונים בין שלבי התכנון, הבנייה והתפעול: COBie, מוכן לתפעול. "מתמקד בהעברת מידע) למשל BIM כחלק מהקונספט של "מודל, BS 1192-4:2014-כלול ב (מפרטי ציוד, אחריות, פרטי קשר של קבלנים).

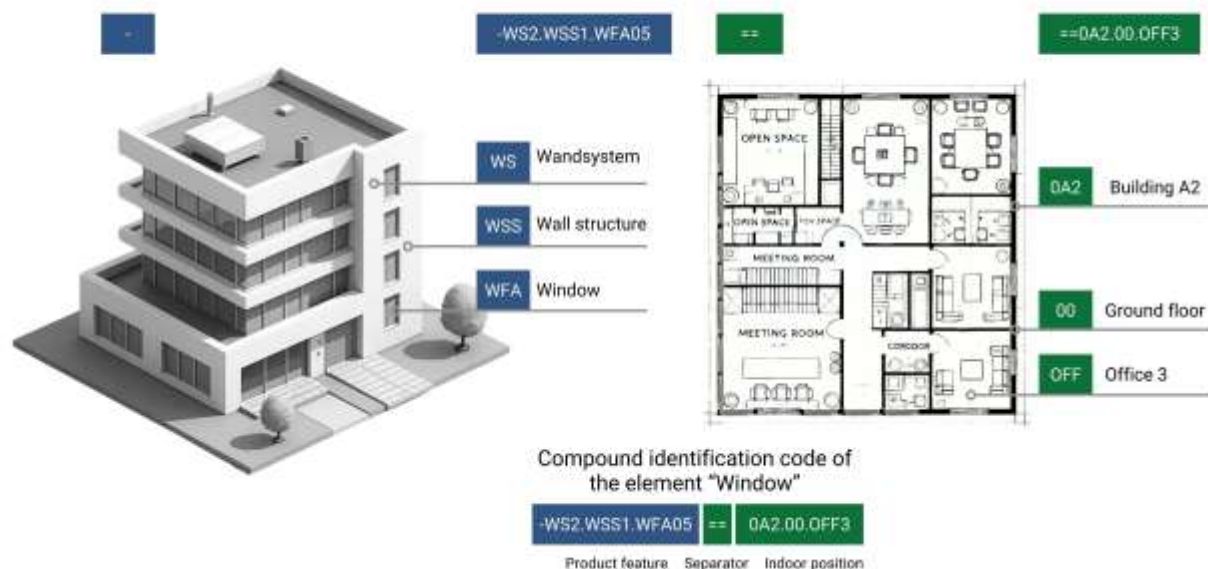
הגלובליזציה של תעשיית הבנייה צפויה להוביל לאיחוד הדרגתי של מערכות סיווג אלמנטים בנייה, מה שיקטין באופן משמעותי את התלות בתקנים לאומיים מקומיים. תהליך זה עשוי להתפתח בדומה להתפתחות התקשורת באינטרנט, שבה פרוטוקולים אוניברסליים להעברת נתונים בסופו של דבר החליפו פורמטים מקומיים מפוזרים, והבטיחו תאימות גלובלית של מערכות.

דרך חלופית להתפתחות עשויה להיות המעבר הישיר למערכות סיווג אוטומטיות המבוססות על טכנולוגיות למידת מכונה טכנולוגיות אלו, המפותחות כיום בעיקר בתחום התחבורה האוטונומית, מציעות פוטנציאל משמעותי ליישום על כמויות CAD גדולות של נתוני תכנון.

כיום, המצב אינו מוגבל רק לקלאסטרזציה הלאומית של סיווגים. בשל המון מאפיינים שאינם נלקחים בחשבון ברמה הממשלתית, כל חברה נאלצת לעסוק בעצמה באיחוד ובסטנדרטיזציה של קטגוריות אלמנטים ומשאבים עימם היא עובדת.

בדרך כלל, התהליך הזה מתחיל בקטן - מטבלאות מקומיות של אובייקטים או מערכות סימון פנימיות. עם זאת, המטרה האסטרטגית היא לעבור לשפה אחידה שמתארת את כל האלמנטים, שתהיה מובנת לא רק בתוך החברה, אלא גם מחוצה לה - באידיאל, מסונכרנת עם מסווגים בינלאומיים או תעשייתיים. גישה כזו מקלה על אינטגרציה עם שותפים חיצוניים - מערכות דיגיטליות ותורמת להקניית תהליכים רציפים אחידים במסגרת מחזור חיי האובייקטים.

ניתנות להרחבה, יש להשתמש או במסווגים ברמה הלאומית, או לבנות מבנה זיהוי IT לפני המעבר לאוטומציה ולמערכות לוגי וחד משמעי של אלמנטים. כל אובייקט - בין אם זה חלון, דלת או מערכת הנדסית - צריך להיות מתואר כך שניתן יהיה לזהות אותו ללא שגיאות בכל מערכת דיגיטלית של החברה. זה קריטי במעבר משרטוטים שטוחים למודלים - דיגיטליים, המכסים גם את שלב התכנון וגם את הפעלת הבניינים.



- דוגמה לקוד זיהוי מורכב של אלמנט בנייה חלון על בסיס מסווג ומיקום בבניין

אחד מהדוגמאות למסווגים פנימיים יכול להיות פיתוח קוד זיהוי מורכב. קוד כזה מאחד מספר רמות מידע: ייעוד קומה 0, חדר 3. מבנה A2 (פונקציונלי של האלמנט) "למשל", חלון בקיר, ("סוגו, וכן את הקישור המרחבי המדויק - בניין רב-רמות כזה מאפשר ליצור מערכת ניווט אחידה במודלים דיגיטליים ובמסמכים, במיוחד בשלבי בדיקה והמרת נתונים שבהם נדרשת קבוצת אלמנטים חד משמעית. חד משמעיות הזיהוי של האלמנט מבטיחה עקביות בין המחלקות ומפחיתה - את הסיכונים להכפלה, שגיאות ואובדן מידע

מסווג בנוי היטב הוא לא רק מסמך טכני, אלא היסוד של המערכת הדיגיטלית של החברה

מבטיח תאימות נתונים בין מערכות

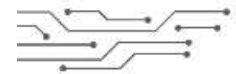
■; מקטין את העלויות לחיפוש ועיבוד מידע

■; מעלה את השקיפות והניהוליות

■. יוצר בסיס להרחבה ואוטומציה

תיאור סטנדרטי של אובייקטים, באמצעות שימוש במסווגים לאומיים או קודי זיהוי מורכבים, הופך לבסיס לנתונים עקביים, החלפת מידע אמינה והטמעה של שירותים אינטליגנטיים - החל מרכישות אוטומטיות ועד דיגיטליים כפולים.

לאחר סיום שלב מבנה הנתונים המגוונים ובחירת המסווג שיישמש לזיהוי וקיבוץ אלמנטים, השלב הבא הוא מודל נתונים נכון. תהליך זה כולל הגדרת פרמטרים מרכזיים, בניית מבנה נתונים לוגי ותיאור הקשרים בין האלמנטים



פרק 4.3

מודל נתונים ומרכז מצוינות

מודל נתונים: מודל קונספטואלי, לוגי ופיזי

ניהול נתונים יעיל (מובנים ומסווגים קודם לכן) אינו אפשרי ללא מבנה אחסון ועיבוד מתוכנן היטב. כדי להבטיח גישה, ועקביות של מידע בשלבי האחסון והעיבוד, חברות משתמשות במודל נתונים - מתודולוגיה המאפשרת לתכנן טבלאות מסדי נתונים וקשרים ביניהם בהתאם לדרישות העסקיות.

מודל הנתונים הוא היסוד שעליו נבנית כל מערכת אקולוגית דיגיטלית. ללא תיאור של מערכות, דרישות ומודל נתונים, מהנדסים ומומחים המייצרים נתונים אינם יודעים ואינם מבינים היכן הנתונים שיצרו ייעשה בהם שימוש.

כמו בבניית בניין, שבה אין אפשרות להתחיל להניח לבנים ללא תכנית, יצירת מערכת אחסון נתונים דורשת הבנה ברורה, של אילו נתונים ייעשה בהם שימוש, כיצד הם קשורים זה לזה, ומי יעבוד איתם. ללא תיאור של תהליכים ודרישות מהנדסים ומומחים המייצרים נתונים מאבדים את המבט על היכן וכיצד הנתונים הללו ייושמו בהמשך.

הוא מאפשר לנסח דרישות, למבנה מידע ולהקל על התקשורת בין בעלי IT-מודל הנתונים משמש כקשר בין העסק ל העניין. במובן זה, מודל הנתונים דומה לעבודה של אדריכל, אשר לפי רעיון הלקוח מפתח תכנית לבניין, ולאחר מכן מעביר אותה לבונים - מנהלי ומפתחי בסיסי נתונים - לצורך ביצוע (יצירת בסיסי נתונים).

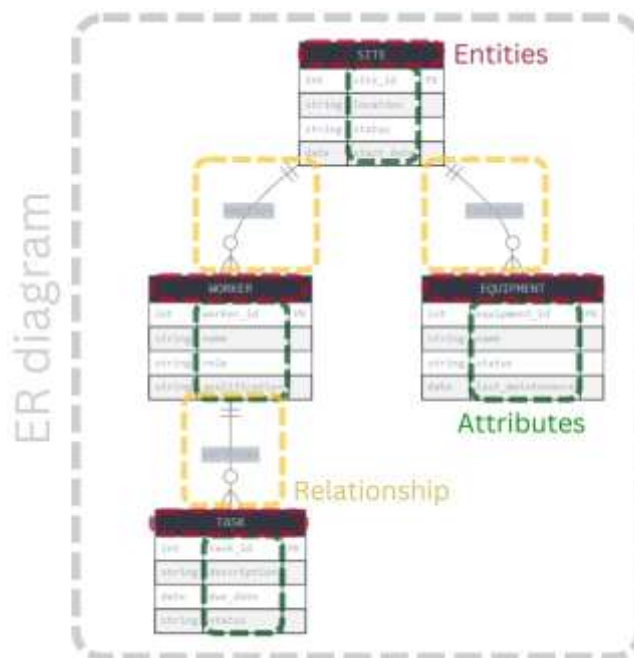
כך, כל חברת בנייה, בנוסף למבנה ולקטגוריזציה של אלמנטים ומשאבים, צריכה לרכוש את האמנות של "בניית" בסיסי נתונים (טבלאות) (וללמוד ליצור קשרים ביניהם, כאילו מחברים לבנים לקיר ידע חזק ואמין של נתוני החברה. המונחים --:המרכזיים במודל הנתונים כוללים

■ ישויות - אלו הם אובייקטים, נתונים עליהם יש לאסוף. בשלב מוקדם של התכנון, ישות יכולה להיות אלמנט בודד (למשל, "דלת", "ובמודל התקציב - קבוצת אלמנטים המאוחדים לפי קטגוריות) (למשל, "דלתות פנימיות").

■ תכונות - אלו הן מאפיינים של ישויות, המתארים פרטים חשובים:מידות, תכונות, עלות הרכבה, לוגיסטיקה ופרמטרים נוספים.

■ קשרים - מראים כיצד ישויות אינטראקציה זו עם זו. הם יכולים להיות מסוגים שונים: "אחד לאחד", "רבים לאחד", "רבים לרבים".

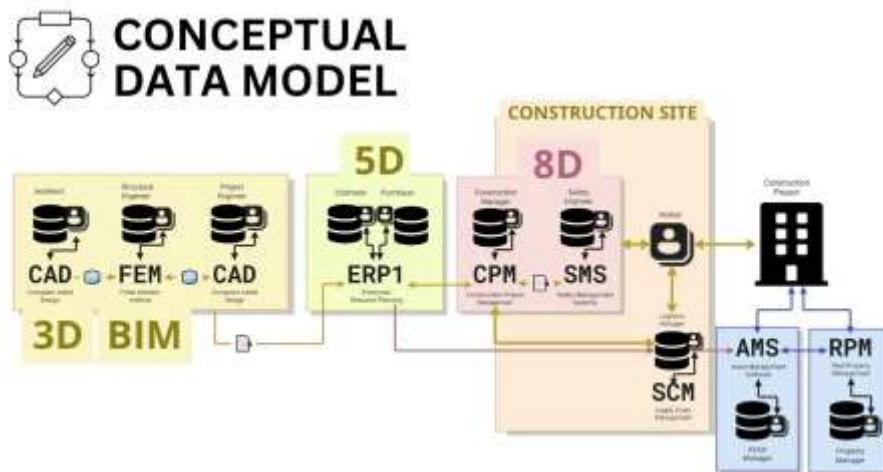
■ סכמות ויזואליות, שבהן מוצגות ישויות, תכונות - ER (Entity-Relationship diagrams) דיאגרמות יכולות להיות קונספטואליות, לוגיות ופיזיות - כל אחת משקפת רמת פירוט ER וקשרים ביניהם. דיאגרמות שונה.



של מבנה בסיס הנתונים הקונספטואלי עם ישויות, תכונות וקשרים **ER** דיאגרמת.

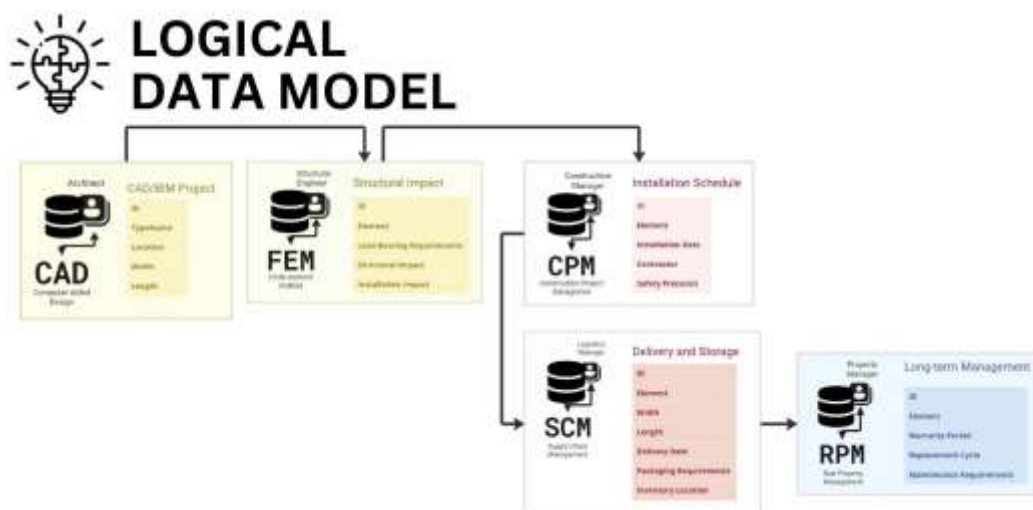
תהליך תכנון הנתונים והגדרת הקשרים ביניהם מחולק באופן מסורתי לשלוש מודלים עיקריים. כל אחד מהם מבצע ה: בהצגת מבנה הנתונים 抽象 פונקציות מסוימות, תוך הבחנה ברמת הפירוט ובדרגת הה

המודל הקונספטואלי של הנתונים: מודל זה מתאר את הישויות העיקריות ואת הקשרים ביניהן, מבלי להיכנס לפרטים של תכונות. בדרך כלל הוא משמש בשלב ההתחלתי של התכנון. בשלב זה ניתן לבצע סקיצות של בסיסי נתונים ומערכות, כדי להראות את הקשר בין מחלקות ומומחים שונים.



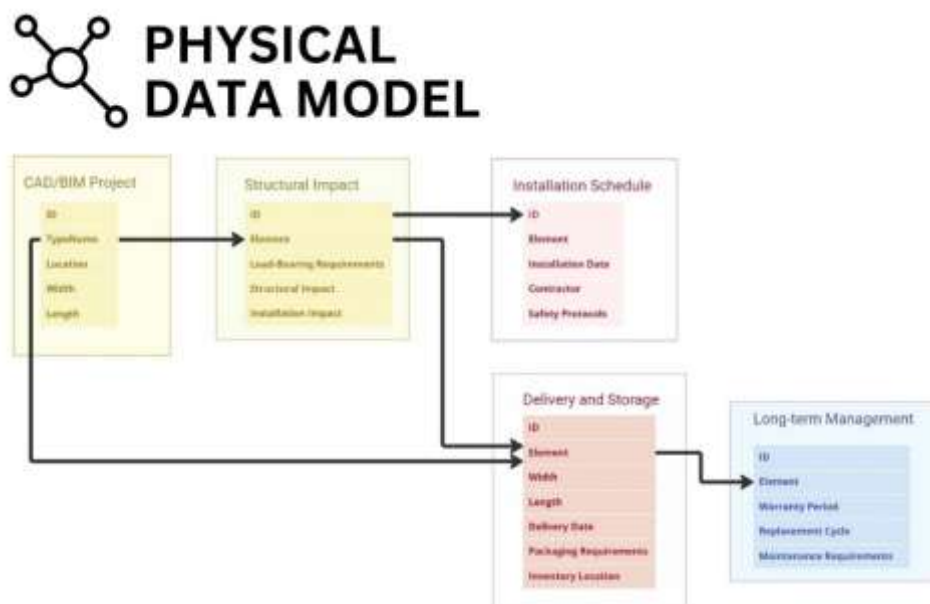
דיאגרמה קונספטואלית מתארת את תוכן המערכת: ייצוג ברמה גבוהה של הקשרים, ללא פרטים טכניים.

מודל נתונים לוגי: בהתבסס על המודל הקונספטואלי, המודל הלוגי כולל תיאור מפורט של ישויות, תכונות, מפתחות וקשרים, הממחישים את המידע העסקי ואת הכללים.



איור 4.33 המודל הלוגי של הנתונים מתאר בפירוט את סוגי הנתונים, הקשרים והמפתחות, אך ללא יישום מערכת.

מודל נתונים פיזי: מודל זה מתאר את המכנים הנדרשים ליישום מסד הנתונים, כולל טבלאות, עמודות וקשרים. הוא מתמקד בביצועי מסד הנתונים, אסטרטגיות אינדוקסציה ואחסון פיזי כדי ליעל את הפריסה הפיזית של מסדי נתונים.



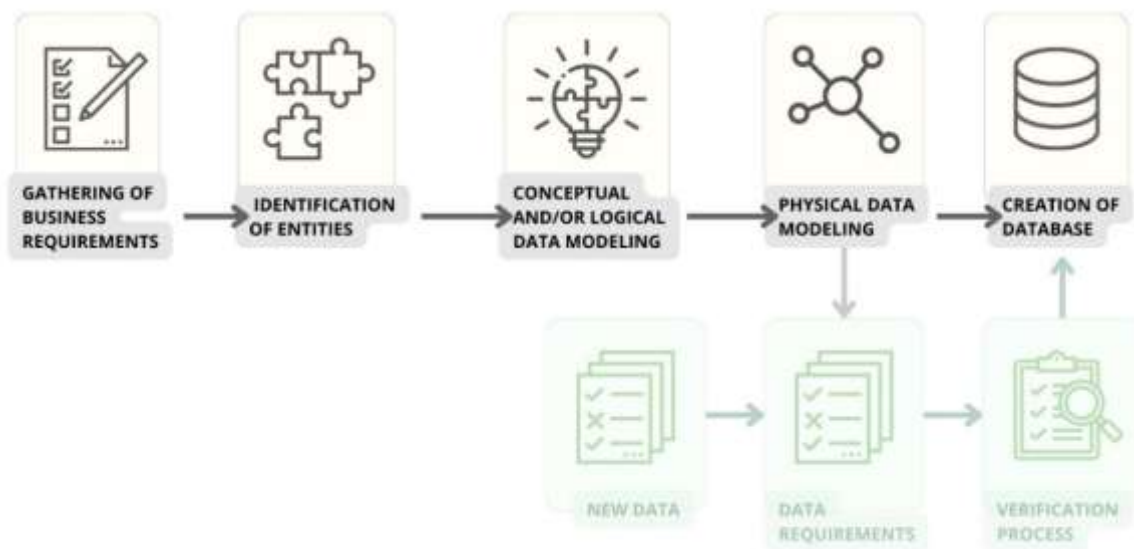
איור 4.34 המודל הפיזי של הנתונים מגדיר כיצד תיושם המערכת, כולל טבלאות ופרטים ספציפיים של מסד הנתונים.

ה משחקת תפקיד מרכזי בבניית תמונת פיתוח מסדי נתונים ועיצוב קשרים טבלאיים, הבנת רמות הה ארכיטקטורת מערכת יעילה.

מתודולוגיה יעילה למידול נתונים מאפשרת לשלב בין משימות עסקיות ליישום טכני, מה שהופך את כל שרשרת התהליכים ליותר שקופה וניהולית. מידול נתונים הוא לא משימה חד פעמית, אלא תהליך הכולל צעדים רציפים (איור

4.35):-

- איסוף דרישות עסקיות: מוגדרות משימות מפתח, מטורות וזרמי מידע. זהו שלב של אינטראקציה פעילה עם מומחים ומשתמשים.
- זיהוי ישויות: מבודדות אובייקטים עיקריים, קטגוריות וסוגי נתונים שחשוב לקחת בחשבון במערכת העתידית.
- פיתוח מודל קונספטואלי ולוגי: תחילה מתועדות ישויות מפתח וקשריהן, ולאחר מכן תכונות, כללים ומבנה מפורט.
- מידול פיזי: מתוכנן היישום הטכני של המודל: טבלאות, שדות, קשרים, מגבלות, אינדקסים.
- יצירת מסד נתונים: השלב הסופי הוא יישום המודל הפיזי במערכת ניהול מסדי נתונים שנבחרה, ביצוע בדיקות והכנה לפעולה.



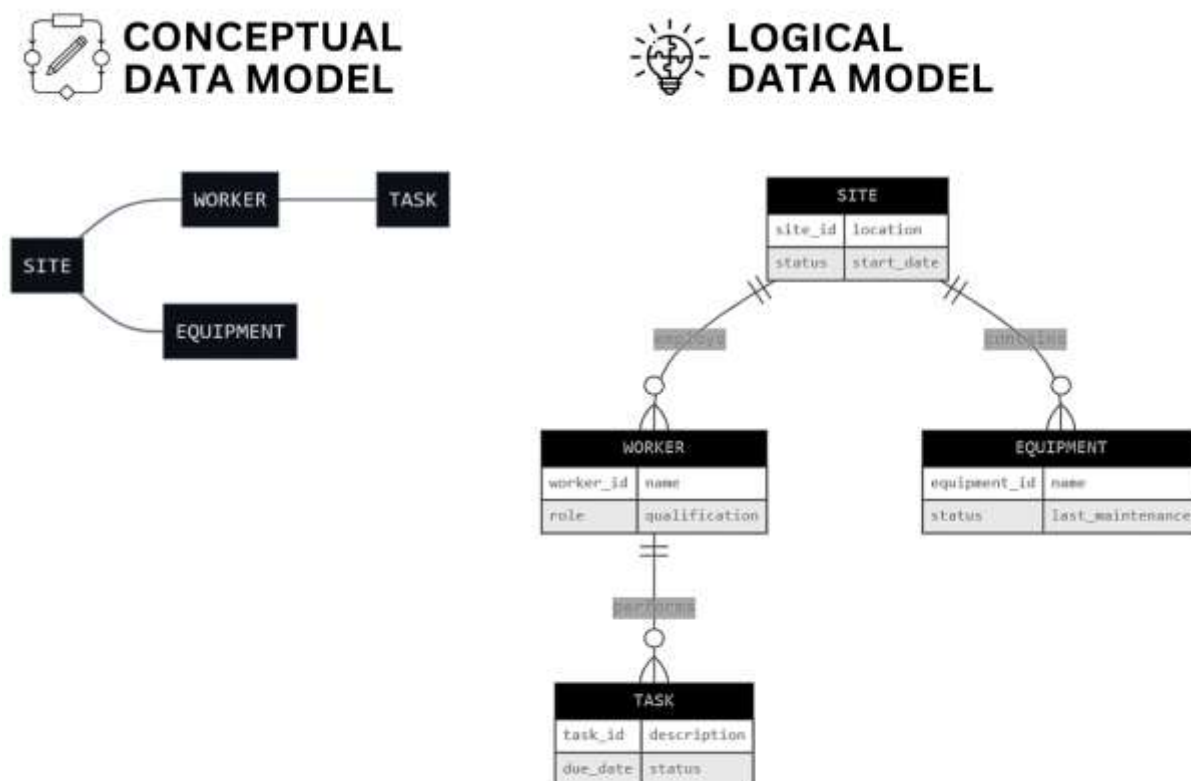
איור 4.35 יצירת מסדי נתונים ומערכות ניהול נתונים עבור תהליכים עסקיים מתחילה בהגדרת דרישות ומידול נתונים.

תהליכי מידול נתונים בנויים היטב מאפשרים להבטיח שקיפות בזרמי המידע, מה שחשוב במיוחד בפרויקטים מורכבים כמו ניהול פרויקט בנייה או אתר בנייה. נבחן כיצד המעבר מהמודל הקונספטואלי ללוגי, ולאחר מכן לפיזי, מסייע לסדר את התהליכים.

מודל נתונים מעשי בהקשר של בנייה

ניקח לדוגמה את משימת מידול הנתונים של ניהול אתר בנייה ונמיר את הדרישות של המנהל לעיצוב מודל לוגי מובנה עובדים (SITE) בהתבסס על הצרכים הבסיסיים של ניהול הבנייה, נגדיר את הישויות המרכזיות עבור אתר בנייה כל ישות (EQUIPMENT_USAGE) ושימוש בציוד (TASK) משימות (EQUIPMENT) ציוד (WORKER), זה יכול להיות תיאור המשימה, מועד סיום TASK מכילה קבוצת תכונות המשקפות מאפיינים חשובים. לדוגמה, עבור 'שם, תפקידו באתר, תעסוקה נוכחית וכו' - WORKER סטטוס, עדיפות; עבור

במודל הלוגי נקבעות הקשרים בין ישויות אלו, המראות כיצד הן אינטראקציה זו עם זו בתהליכי עבודה אמיתיים (איור 4.36). לדוגמה, הקשר בין האתר לעובדים מצביע על כך שבאתר אחד יכולים לעבוד מספר עובדים, והקשר בין העובדים למשימות משקף שעובד אחד יכול לבצע מספר משימות.

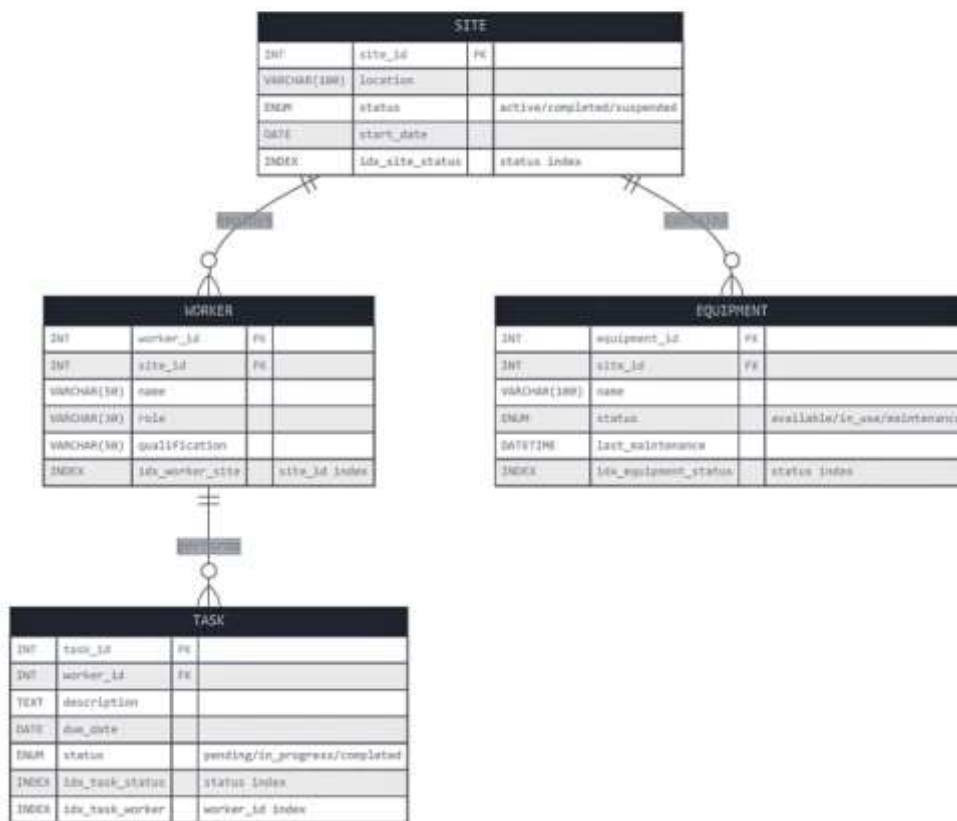


איור 4.36 המודל הקונספטואלי והלוגי של הנתונים, שנוצר על פי דרישות המנהל לתיאור התהליכים באתר הבנייה.

(VARCHAR, INT, DATE), במעבר למודל הפיזי מתווספות פרטים טכניים של היישום: סוגי נתונים ספציפיים מפתחות ראשיים וזרים לקשרים בין הטבלאות, כמו גם אינדקסים לשיפור ביצועי מסד הנתונים (איור 4.37).

לדוגמה, עבור הסטטוסים יש לקבוע סוגים ספציפיים עם ערכים אפשריים, וכדי לשפר את ביצועי החיפוש יש להוסיף זה הופך את התיאור הלוגי של המערכת לתוכנית יישום. status ו-worker_id אינדקסים על שדות מפתח, כמו קונקרטי של מסד הנתונים, מוכנה ליצירה והטמעה.

PHYSICAL DATA MODEL



איור 4.37 המודל הפיזי של הנתונים מתאר את הישויות באתר הבנייה דרך הפרמטרים המינימליים הנדרשים.

המודל הפיזי לעיתים שונה מהמודל הלוגי. בממוצע, חלוקת הזמן על המודלים נראית כך: כ-50% מוקדש למודל הקונספטואלי (איסוף דרישות, דיון בתהליכים, זיהוי ישויות, (10% – למודל הלוגי) הבהרת תכונות וקשרים (ו-40% – למודל הפיזי) יישום, בדיקה, התאמה למערכת ניהול מסדי נתונים)

איוון זה מוסבר בכך שהשלב הקונספטואלי מניח את היסוד למבנה הנתונים, בעוד שהמודל הלוגי רק מבהיר את הקשרים והתכונות. המשאבים הגדולים ביותר נדרשים למודל הפיזי, שכן בשלב זה הנתונים מוטמעים בפלטפורמות ובכלים קונקרטיים.

יצירת בסיס נתונים באמצעות מודלים לשוניים גדולים

עם מודל הנתונים ותיאור הישויות דרך הפרמטרים, אנו מוכנים ליצור מסדי נתונים – מאגרי מידע שבהם נוכל לאחסן

את המידע המתקבל לאחר שלב הסטרוקטוריזציה בתהליכים קונקרטיים

בדוגמת שפת התכנות SQLite ננסה ליצור דוגמה למסד נתונים פשוט אך פונקציונלי עם מינימום קוד באמצעות "SQL מסדי נתונים רלציוניים נבחנו בפירוט בפרק "מסדי נתונים רלציוניים מובנים ושפת השאילתות Python.

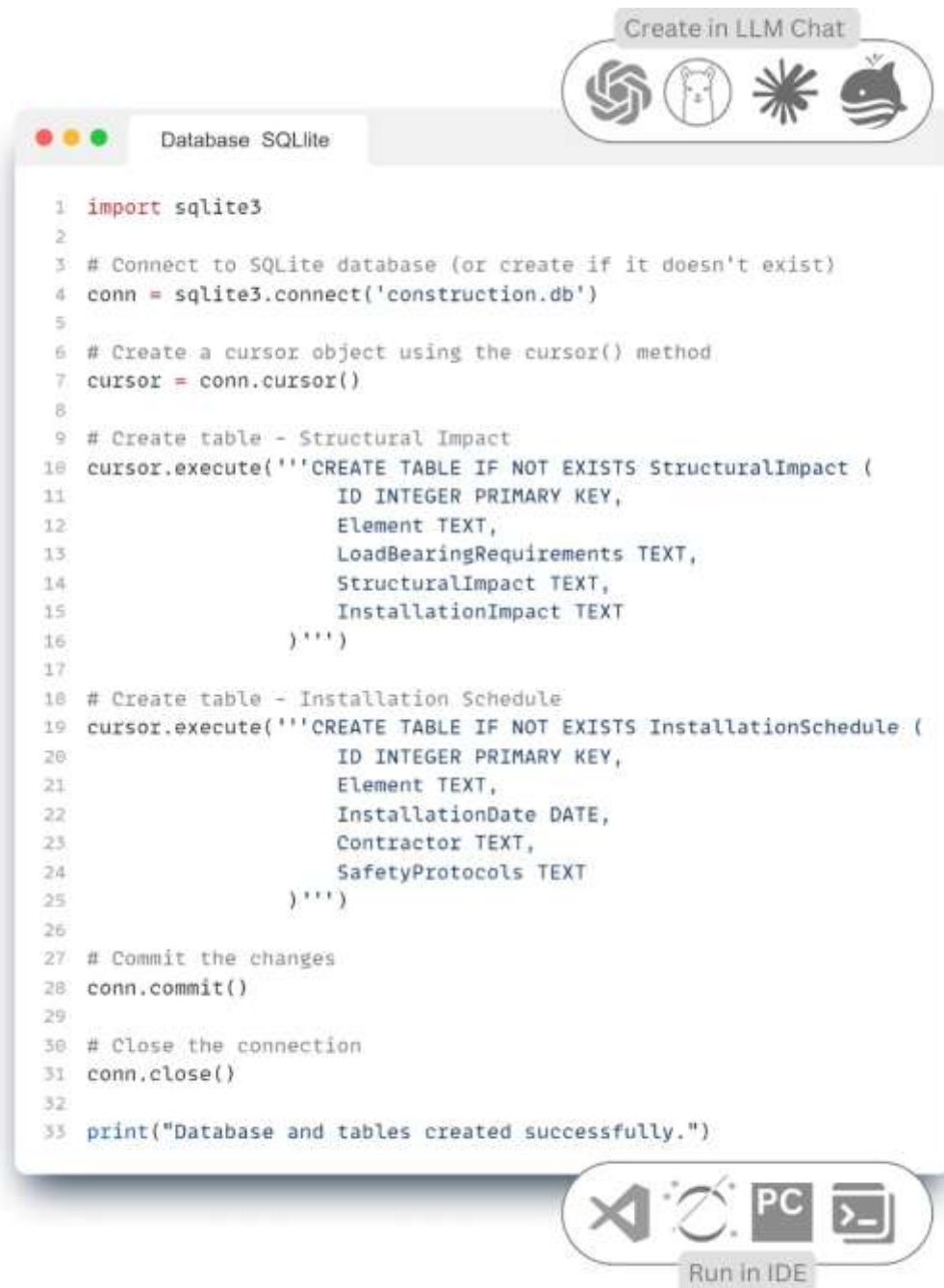
SQL היא מסד נתונים קל, קוד פתוח, מוטמע, חוצה פלטפורמות, התומך בפונקציות בסיסיות של SQLite. ונמצאת כמעט בכל סמארטפון EPM, ERP, CAD (BIM)-בשימוש נרחב ב SQLite.

על בסיס המודל הפיזי שנוצר למעלה) איור 4.34 (עבור SQLite ליצור עבורנו מסד נתונים קטן LLM נבקש ממודל - (לוח זמנים להרכבה) CPM-השפעה מבנית (ו FEM מערכות

🗣️ ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, צ'אט LLM-שלחו בקשה טקסטואלית ל: (או כל אחד אחר QWEN

עבור מערכות "השפעה מבנית "ו"לוח זמנים להרכבה" מהאיור המצורף SQLite צור מסדי נתונים -> (הכנס לשיחה את הסכימה כהדמיה) איור 4.34)

LLM: תשובת



איור 4.38 הקוד שנוצר יוצר את מסדי הנתונים והטבלאות מהתמונה שהועלתה, אליהם אנו יכולים כעת לטעון נתונים.

שהיא בדרך כלל כלולה בסביבת SQLite-הקוד הזה ליצירת בסיס נתונים יכול להתבצע בסביבת פייתון עם תמיכה בפייתון הסטנדרטית. ניתן להריץ אותו על כל מחשב מקומי עם פייתון, המסוגל להריץ סקריפטים וליצור קבצים. לאחר דומה) construction.db בשם SQLite ההפעלה והביצוע, הקוד הזה ייצור על הדיסק הקשיח קובץ - בסיס נתונים "עם גיליונות עבודה, (שיכיל את הטבלאות "השפעה מבנית "ו"לוח זמנים להתקנה Excel לקובץ

ניתן יהיה להוסיף נתונים באמצעות שאילתות, construction.db לאחר יצירת הטבלאות החדשות בבסיס הנתונים או ייבוא, מה שיאפשר מאוחר יותר להתחיל ביצירת עיבוד נתונים אוטומטי. ניתן לייבא נתונים לבסיס הנתונים SQL מבסיסי נתונים אחרים ואחסונים API או לייצא דרך Excel גיליונות, CSV מקבצי SQLite

כדי להקים תהליכים יציבים של מודל נתונים וניהול בסיסי נתונים יעיל, חברות זקוקות לאסטרטגיה ברורה, כמו גם לתיאום מסודר בין צוותים טכניים וצוותי עסקים. בתנאים של פרויקטים מפוזרים ומקורות נתונים רבים, לעיתים קשה להבטיח עקביות, סטנדרטיזציה ובקרת איכות בכל הרמות. אחת מהפתרונות המרכזיים עשויה להיות הקמת מרכז מצוינות (Data Modeling Center of Excellence, COE) פנימי למודל נתונים

במודל נתונים (CoE) מרכז מצוינות

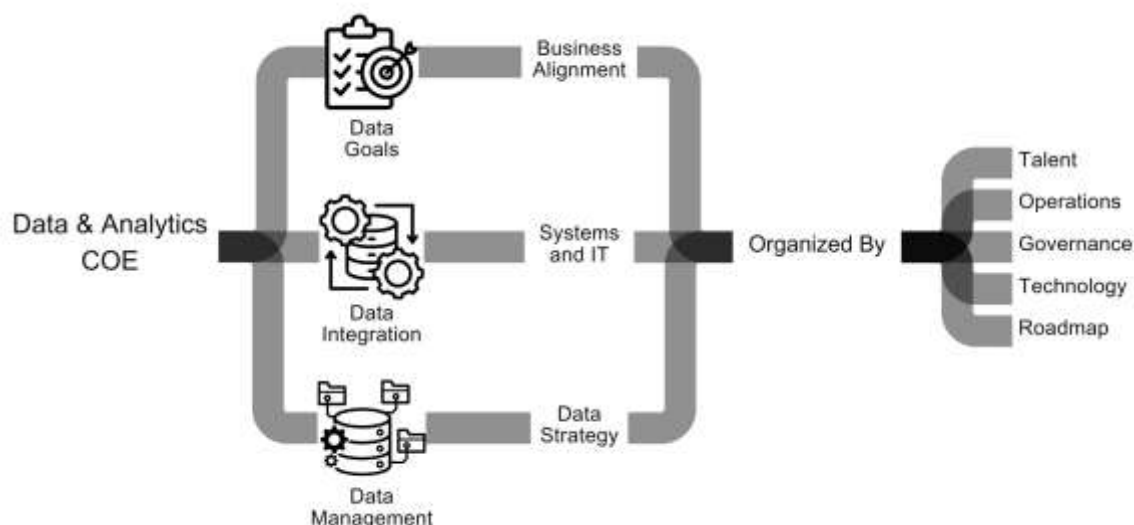
בתנאים שבהם נתונים הופכים לאחד מהנכסים האסטרטגיים המרכזיים, חברות צריכות לא רק לאסוף ולאחסן מידע (Center of Excellence, CoE) בצורה נכונה - חשוב ללמוד לנהל נתונים בצורה מערכתית. מרכז מצוינות למיון ומודל נתונים הוא יחידה ארגונית המבטיחה עקביות, איכות ויעילות בכל העבודה עם נתונים בארגון (Excellence, CoE)

הוא הלב של תמיכה מקצועית והבסיס המתודולוגי לשינויים דיגיטליים בחברה (CoE) מרכז מצוינות הוא יוצר תרבות עבודה עם נתונים ומאפשר לארגונים לבנות תהליכים, מקבלים החלטות לא על סמך אינטואיציה או מידע מקומי, אלא על סמך נתונים מובנים, מאומתים ומייצגים

מרכז מצוינות לנתונים בדרך כלל מוקם מצוותים חוצי פונקציות, העובדים לפי עקרון "שתי פיצות". עיקרון זה, שהוצע על ידי ג'ף בזוס, מציין כי גודל הצוות צריך להיות כזה שניתן להאכיל אותו בשתי פיצות, כלומר לא לעלות על 6-10 צריך לכלול עובדים CoE-אנשים. גישה זו מסייעת להימנע מבירוקרטיה מיותרת ומגבירה את הגמישות בעבודה. צוות ה, עם מגוון מיומנויות טכניות: מניתוח נתונים ולמידת מכונה ועד מומחיות בתחומים עסקיים ספציפיים. עם ידע טכני מעמיק מהנדסי נתונים צריכים לא רק ליעל תהליכים ולמודל נתונים, אלא גם לתמוך בעמיתים, מקצרים את הזמן למשימות -שגרתיות

כמו שבטבע, יציבות המערכת האקולוגית מתבצעת על ידי מגוון ביולוגי, כך בעולם הדיגיטלי גמישות והתאמה מושגות דרך מגוון גישות לעבודה עם נתונים. עם זאת, מגוון זה צריך להתבסס על כללים ומושגים אחידים.

ניתן להשוות ל"מצבי אקלים" של מערכת אקולוגית יערית, המגדירים אילו סוגי נתונים ישגשו (CoE) מרכז מצוינות תורם לבחירה הטבעית של שיטות עבודה CoE-ואילו ייפסלו אוטומטית. על ידי יצירת "אקלים" נוח לנתונים איכותיים, ה, מיטביות ומודלים, אשר בהמשך הופכים לסטנדרטים בארגון



בנתונים ובאנליטיקה מאחד את המומחיות בהיבטים המרכזיים של ניהול נתונים, אינטגרציה (CoE) מרכז המצוינות 4.39 איור שלהם ופיתוח אסטרטגיה.

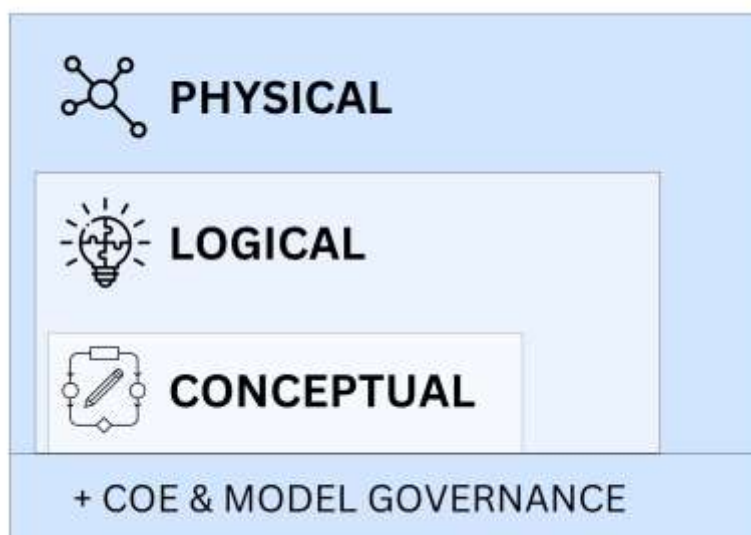
לספק לחבריו רמה מספקת של אוטונומיה CoE-כדי לזרז את מחזורי האינטגרציה ולהשיג תוצאות טובות יותר, על ה בקבלת החלטות. זה במיוחד חשוב בסביבה דינמית, שבה שיטת ניסוי וטעייה, משוב מתמשך ושחרורים תכופים יכולים להביא תועלת משמעותית. עם זאת, אוטונומיה זו יעילה רק כאשר יש תקשורת ברורה ותמיכה מצד ההנהלה הבכירה. ללא חזון אסטרטגי ותיאום מלמעלה, אפילו הצוות המיומן ביותר עלול להיתקל במכשולים בעת יישום היוזמות שלו.

או ההנהלה הבכירה של החברה שאחראית לכך שהגישה למידול נתונים לא תוגבל לפרויקט CoE-זהו ה אחד או שניים, אלא תהיה משולבת במערכת הכוללת של ניהול מידע ותהליכי עסקים.

אחראי על פיתוח (Data Governance), בנוסף למשימות הקשורות למידול נתונים וניהולם (CoE) מרכז המצוינות סטנדרטים וגישות אחידים לפריסה ולתפעול של תשתיות נתונים. בנוסף, הוא יוצר תרבות של שיפור מתמיד (4.310 איור) אופטימיזציה של תהליכים ושימוש יעיל בנתונים בארגון.

יכולה להיות מחולקת באופן גס למספר בלוקים מרכזיים CoE-הגישה המערכתית לניהול נתונים ומודלים בתוך ה

- מפתח ומיישם מתודולוגיות המאפשרות לאחד CoE-סטנדרטיזציה של תהליכים וניהול מחזור חיי המודלים: ה את יצירת המודלים וניהולם. זה כולל: יצירת תבניות מבניות, שיטות לבקרת איכות ומערכות לניהול גרסאות המבטיחות רציפות נתונים בכל שלב של העבודה.
- מוגדרות תפקידים מרכזיים בתהליך מידול הנתונים. כל CoE-ניהול תפקידים וחלוקת אחריות: במסגרת ה משתתף בפרויקט מקבל פונקציות ואזורי אחריות מוגדרים בבירור, מה שמסייע לעבודה מסונכרנת של הצוותים ומפחית את הסיכונים לאי התאמה בנתונים.
- בקרת איכות וביקורת: ניהול יעיל של נתוני בנייה דורש ניטור מתמשך של איכותם. מיושמים מנגנונים אוטומטיים לבדוק נתונים, לזהות שגיאות ופריטים חסרים.
- אחראי על יצירת מערכת אחידה של סיווג וזיהוי, סטנדרטים CoE-ניהול מטא-נתונים ואדריכלות מידע: ה. לשמות ותיאורי ישויות, מה שקריטי לאינטגרציה בין מערכות.



CoE-מידול נתונים וניהול איכות נתונים הם אחת המשימות המרכזיות של ה **4.310** איור

בנתונים הוא לא רק קבוצת מומחים, אלא מנגנון מערכתית, היוצר תרבות מונעת נתונים חדשה (CoE) מרכז המצוינות ומבטיח גישה אחידה לעבודה עם נתונים בכל החברה. בזכות אינטגרציה נכונה של תהליכי מידול במערכת הכוללת של מסייע לעסק לשפר באופן מתמיד את מוצריו ותהליכי CoE-ניהול מידע, סטנדרטיזציה, סיווג ובקרת איכות נתונים, ה.העסק, להגיב במהירות לשינויים בשוק ולקבל החלטות מושכלות על בסיס אנליטיקה מהימנה

גישה המבטיחה אספקה רציפה, אוטומציה - DataOps מרכזים כאלה יעילים במיוחד בשילוב עם עקרונות מודרניים של בפרק השמיני, בפרק "טכנולוגיות מודרניות לעבודה עם נתונים DataOps ובקרת איכות נתונים. נדון בפרטים על "בתעשיית הבנייה

בפרקים הבאים נעבור מאסטרטגיה לפרקטיקה - נתחפש באופן סמלי למרכז עיבוד נתונים: נבחן בכמה דוגמאות כיצד מתבצע פרמטריזציה של המשימה, איסוף דרישות ותהליך אוטומטי של אימות



פרק 4.4

מערכת דרישות ואימות מידע

איסוף וניתוח דרישות: המרת תקשורת לנתונים מובנים

איסוף וניהול דרישות הוא הצעד הראשון להבטחת איכות הנתונים. למרות ההתפתחות של כלים דיגיטליים, רוב הדרישות עדיין מנוסחות בצורה לא מובנית: דרך מכתבים, פרוטוקולי פגישות, שיחות טלפון ודיונים בעל פה. צורת תקשורת זו, מקשה על אוטומציה, בדיקה ושימוש חוזר במידע. בפרק זה נבחן כיצד להמיר דרישות טקסטואליות למבנים פורמליים. תוך הבטחת שקיפות ומערכתיות של משימות עסקיות.

מחקרי חברת גרטנר "איכות נתונים: שיטות עבודה מומלצות להשגת מידע מדויק" מדגישים את החשיבות הקריטית של איכות הנתונים להצלחת יוזמות בתחום הנתונים והאנליטיקה. הם מציינים כי איכות נתונים נמוכה עולה לארגונים בממוצע לפחות 12.9 מיליון דולר בשנה, וכי נתונים אמינים ואיכותיים נדרשים ליצירת חברה המנוהלת על ידי נתונים.

חוסר בדרישות מובנות גורם לכך שאותו אלמנט (ישות) ופרמטרים שלו יכולים להישמר במערכות שונות בגרסאות שונות. זה לא רק מפחית את היעילות של התהליכים, אלא גם גורם לאובדן זמן, שכפול מידע וצורך בבדיקת נתונים מחדש לפני השימוש בהם. כתוצאה מכך, אפילו חוסר אחד - פרמטר אבוד או אלמנט מתואר לא נכון - יכול להאט את קבלת ההחלטות ולגרום לבזבזת לא יעיל של משאבים.

מחסור בברזל גרם לאובדן הסוס. מחסור בסוס גרם לאובדן הרוכב. מחסור ברוכב גרם לאובדן ההודעה. מחסור בהודעה גרם לאובדן הקרב. מחסור בקרב גרם לאובדן הממלכה. והכל בגלל מחסור בברזל בקלב.

פתגם -

ניתוח ואיסוף דרישות לתהליך מילוי ואחסון נתונים מתחיל בזיהוי כל בעלי העניין. בדומה לפתגם שבו אובדן של מסמר אחד מביא לשרשרת של תוצאות קריטיות, בעסק - אובדן של משתתף אחד, דרישה חסרה או אובדן אפילו של פרמטר אחד יכולים להשפיע באופן משמעותי לא רק על תהליך עסקי בודד, אלא גם על כל האקוסיסטמה של הפרויקט והארגון כולו. לכן, חשוב מאוד לזהות גם את האלמנטים, הפרמטרים והתפקידים שנראים לא משמעותיים, אך לאחר מכן עשויים להתברר כקריטיים ליציבות העסק.

נניח שלחברה יש פרויקט שבו הלקוח מציע בקשה חדשה - להוסיף חלון נוסף בצד הצפוני של הבניין. בתהליך הקטן, מנהל הבנייה, CAD (BIM) בקשת הלקוח להוספת חלון חדש בפרויקט הנוכחי "מעורבים האדריכל, המזמין, מומחה" מהנדס בקרת איכות, מהנדס בטיחות, מנהל הבקרה ומנהל הנדל"ן, ERP, מנהל הלוגיסטיקה, אנליסט

אפילו בתהליך קטן יכולים להיות מעורבים עשרות מומחים שונים. כל משתתף בתהליך צריך להבין את הדרישות של המומחים עימם הוא קשור ברמת הנתונים.

- התקשורת בין הלקוח למומחים בשרשרת התהליך מתבצעת כך (4.41 איור) ברמת הטקסט

- 🔊 "המזמין": החלטנו להוסיף חלון נוסף בצד הצפוני כדי לשפר את התאורה. האם זה ניתן ליישום
 - 🔊 CAD-האדריכל: כמובן, אני אעבור על הפרויקט כדי לכלול את החלון החדש ואשלח את התוכניות המעודכנות ב (BIM)".
 - 🔊 עם החלון הנוסף ואחרי CAD (BIM)-קיבלתי את הפרויקט החדש. אני מעדכן את מודל ה- CAD (BIM) מומחה "אספק את המיקום והמידות המדויקות של החלון החדש FEM אישור מהמהנדס ומעדכנים את כל D מנהל הבנייה: קיבלתי את הפרויקט החדש. אנו מתאימים את לוחות הזמנים של ההתקנה ב-4 "הקבלנים המשניים הרלוונטיים
 - 🔊 לניהול עידי של CAFM על החלון החדש למערכת D אני אכניס את נתוני ה-6: (CAFM) מהנדס המתקנים "המתקן ותכנון תחזוקה
 - 🔊 "מנהל הלוגיסטיקה": אני זקוק למידות ולמשקל של החלון החדש כדי לארגן את המשלוח שלו לאתר
 - 🔊 ERP-במערכת ה D אני זקוק לטבלאות כמויות ולסוג המדויק של החלון כדי לעדכן את התקציב ב-5: ERP אנליסט "שלנו, כדי לשקף את עלות החלון החדש בסך התקציב של הפרויקט
 - 🔊 מהנדס בקרת איכות: ברגע שהמפרטים של החלונות יהיו מוכנים, אני אודא שהם עומדים בסטנדרטים שלנו לאיכות "וחומרים
 - 🔊 מהנדס הבטיחות: אני אעריך את היבטי הבטיחות של החלון החדש, תוך מתן תשומת לב מיוחדת לעמידה בדרישות D ובתוכנית הפיננסי לפי שיטת 8
 - 🔊 כדי לשקף את D אנו נעדכן את ציר הזמן שלנו ב-4, ERP-מנהל הבקרה: בהתבסס על כמות העבודה המדויקת מה "התקנת החלון החדש, ונשמור את הנתונים החדשים במערכת ניהול התוכן של הפרויקט
 - 🔊 (עובד) מתקין: (אני זקוק להוראות התקנה, הרכבה ולוחות זמנים לביצוע העבודה. בנוסף, האם הוכנסו כללים "מיוחדים לבטיחות שעליי לעקוב אחריהם
 - 🔊 "מנהל הנדל"ן: לאחר ההתקנה אני אעדכן את המידע על האחריות והתחזוקה לניהול ארוך טווח
 - 🔊 "מנהל הנכסים": מהנדס הציוד, בבקשה שלח את הנתונים הסופיים למעקב נכסים ולניהול מחזור חיים
 - 🔊 "הלקוח": חכו, אולי אני ממחר, והחלון לא יהיה נחוץ. אולי כדאי לעשות מרפסת
- בתרחישים כאלה, המתרחשים לעיתים קרובות, אפילו שינוי קטן יוצר תגובת שרשרת בין מספר מערכות ותפקידים (4.41 איור) בשלב זה, כמעט כל התקשורת מתבצעת בצורה טקסטואלית: מכתבים, צ'אטים, פרוטוקולים של פגישות.
- במערכת כזו של תקשורת טקסטואלית עבור פרויקט בנייה, מערכת אישור משפטי ורישום של כל הפעולות בהחלפת נתונים וכל ההחלטות שהתקבלו היא מאוד חשובה. זה הכרחי כדי להבטיח את הכוח המשפטי ואת האפשרות לעקוב אחרי כל החלטה, הוראה או שינוי, מה שמפחית את הסיכון להיווצרות "אי הבנות" בעתיד

PROCESS EXTRA WINDOW ON THE NORTH SIDE

TEXTUAL DATA

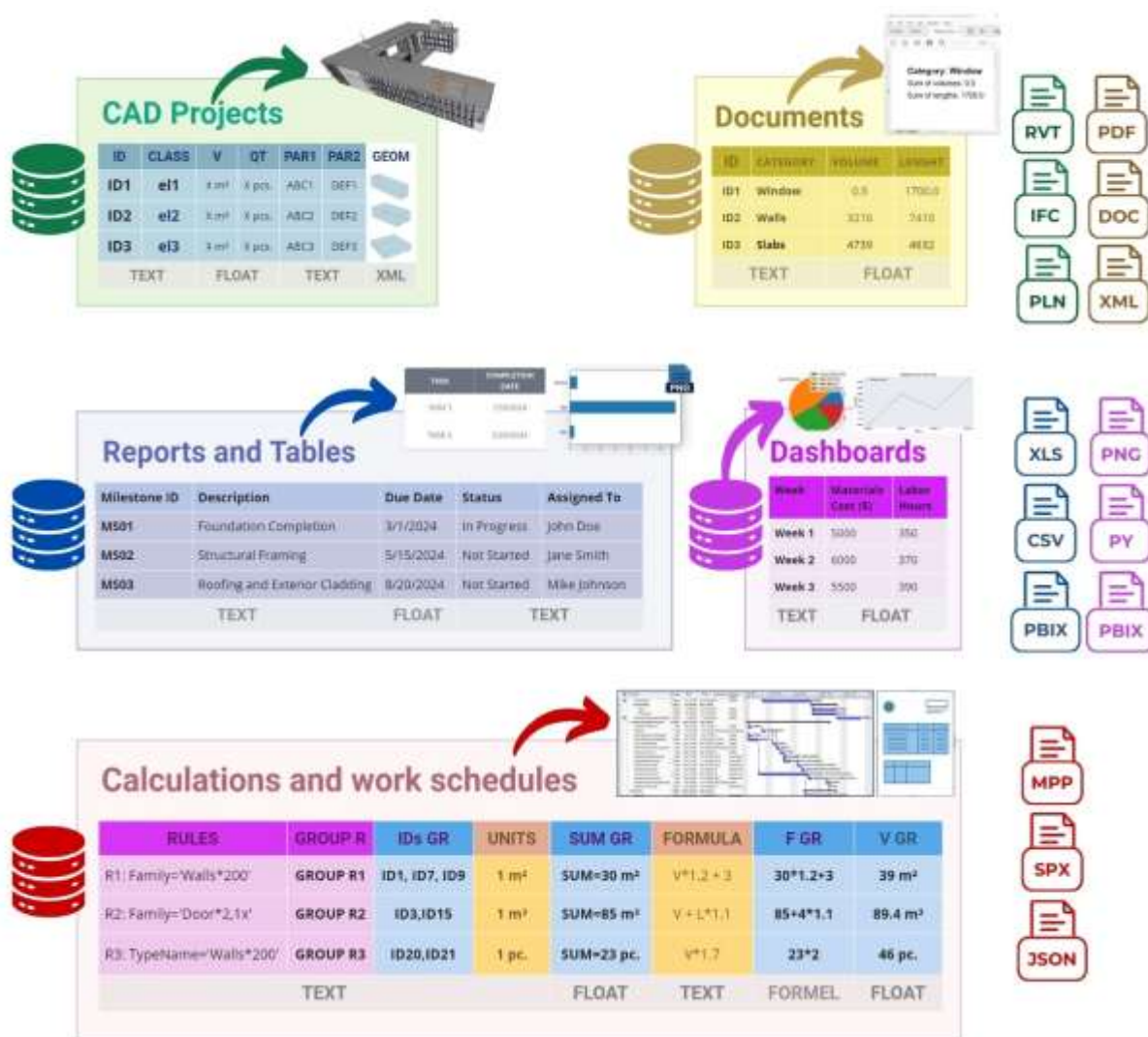


התקשורת בין הלקוח למבצע בשלב ההתחלתי של הפרויקט מכילה לעיתים קרובות נתונים טקסטואליים בפורמטים 4.41 איור שונים.

חוסר שליטה משפטית ואישור החלטות במערכות המתאימות לפרויקט הבנייה יכול להוביל לבעיות חמורות עבור כל המשתתפים בו. כל החלטה, הוראה או שינוי, שהתקבלו ללא תיעוד ואישור נאות, עשויים להוביל למחלוקות ולתביעות משפטיות).

ההבטחה של תיעוד משפטי לכל ההחלטות בתקשורת טקסטואלית יכולה להתבצע רק באמצעות כמות גדולה של מסמכים חתומים, אשר יוטלו על ההנהלה, המחויבת לתעד את כל העסקאות. בסופו של דבר, אם כל משתתף מחויב לחתום על מסמכים עבור כל פעולה, המערכת מאבדת את הגמישות שלה והופכת לבלבול בירוקרטי. חוסר אישור של עסקאות לא רק מעכב את יישום הפרויקט, אלא גם עלול להוביל להפסדים כספיים, כמו גם להחמרת היחסים בין המשתתפים, עד לבעיות משפטיות.

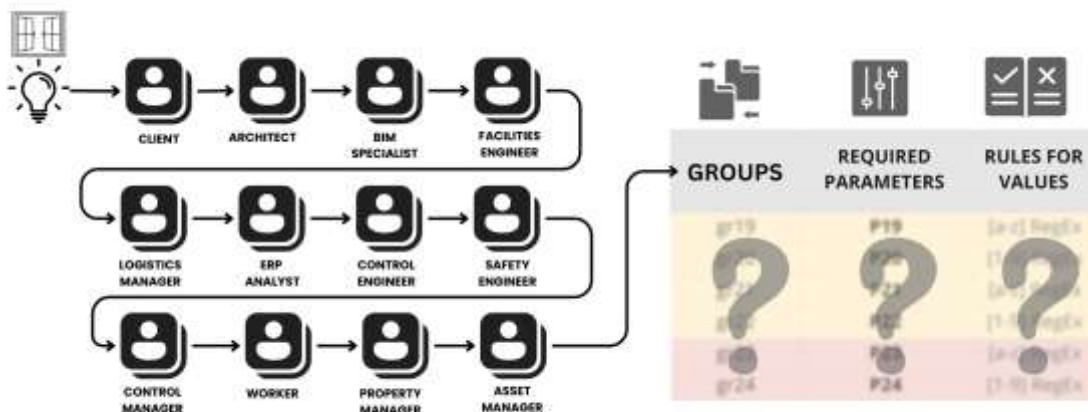
תהליך כזה של הסכמה ואישור עסקאות, שמתחיל בדרך כלל בדיונים טקסטואליים, בשלב הבא עובר בהדרגה לפורמט. מה שמסבך משמעותית את התקשורת, שהתרחשה רק דרך טקסט, (4.42 איור) של החלפת מסמכים בפורמטים שונים ללא דרישות מוגדרות בבירור, אוטומציה של תהליכים כאלה, המלאים בנתונים בפורמטים שונים ובכמות גדולה של דרישות טקסטואליות, הופכת כמעט לבלתי אפשרית.



כל מערכת בנוף של חברת הבנייה משמשת כמקור למסמכים משפטיים משמעותיים בפורמטים שונים 4.42 אור.

תקשורת טקסטואלית דורשת מכל מומחה או הכרה עם כל ההתכתבות, או השתתפות קבועה בכל הפגישות, כדי להבין את הסטטוס הנוכחי של הפרויקט.

כדי להתגבר על מגבלה זו, יש צורך לעבור מתקשורת טקסטואלית למודל דרישות מובנה. זה אפשרי רק באמצעות ניתוח בדיוק (4.43 אור) שיטתי, ויזואליזציה של התהליך ותיאור האינטראקציות בצורה של דיאגרמות זרימה ומודלי נתונים עברנו מרמת הקשר-רעיון לרמה קונספטואלית, והוספנו את המערכות והכלים שבהם (4.37 אור) כמו במידול נתונים. -משמשים המשתתפים, כמו גם את הקשרים ביניהם.



כדי ללמוד לנהל ולאוטומט את תהליך האימות, יש צורך לויזואליזציה של התהליכים ולמבנה את הדרישות **4.43** איור

הצעד הראשון במערכת הדרישות והקשרים הוא ויזואליזציה של כל הקשרים והיחסים באמצעות דיאגרמות קונספטואליות. הרמה הקונספטואלית לא רק שתקל על כל המשתתפים בתהליך להבין את כל שרשרת הטכנולוגיה, אלא גם תציג בצורה ברורה מדוע ולמי נדרשות הנתונים (והדרישות) (בכל שלב בתהליך

דיאגרמות תהליכים ויציאות של תכניות קונספטואליות

כדי לגשר על הפער בין הגישות המסורתיות לחדשניות בניהול נתונים, חברות צריכות לעבור באופן מודע מתיאורים טקסטואליים מפורקים לייצוג מסודר של תהליכים. האבולוציה של הנתונים - מלוחות חרס ועד מערכות אקולוגיות דיגיטליות - דורשת כלים חדשים לחשיבה. ואחד מהכלים הללו הוא המודל הקונספטואלי באמצעות דיאגרמות. יצירת דיאגרמות ויזואליות - דיאגרמות תהליכים, דיאגרמות אינטראקציה - מאפשרת למשתתפי הפרויקט להבין כיצד הפעולות וההחלטות שלהם משפיעות על כל מערכת קבלת ההחלטות.

אם התהליכים דורשים לא רק אחסון נתונים, אלא גם ניתוח או אוטומציה שלהם, אז יש להתחיל לעסוק בנושא יצירת הרמה הקונספטואלית-ויזואלית של הדרישות

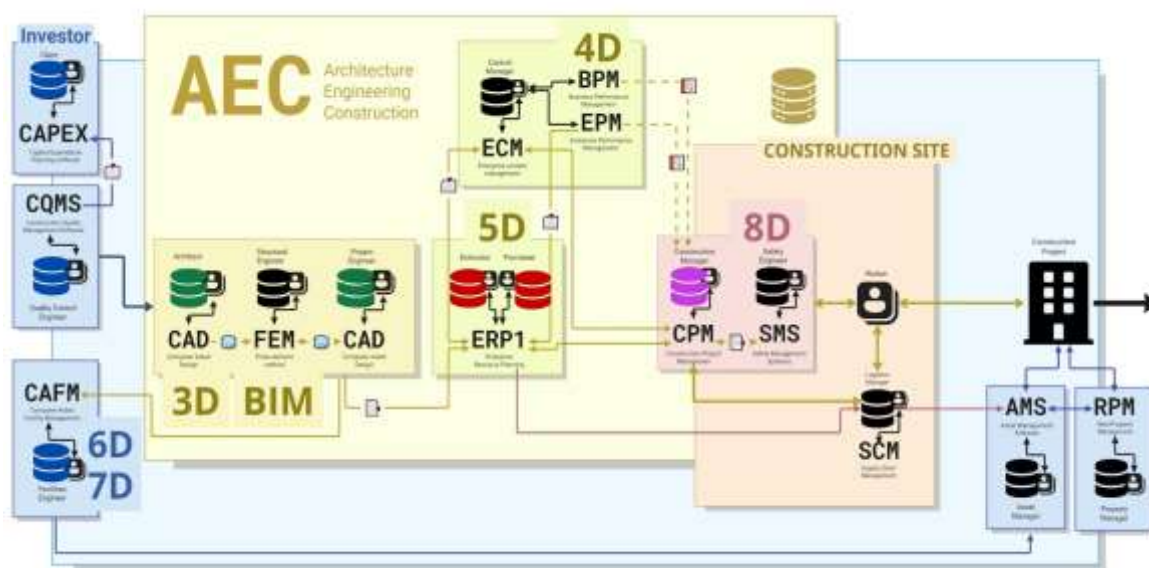
בדוגמה שלנו) איור 4.41 (כל מומחה יכול להיות חלק לא רק מקבוצה קטנה, אלא גם ממחלקה גדולה יותר, הכוללת עד ERP, CAD, עשרה מומחים תחת ניהולו של מנהל ראשי. כל מחלקה משתמשת בבסיס נתונים ייחודי של היישום) כגון וכו', (שמתעדכן באופן קבוע במידע הנכנס, הנדרש ליצירת מסמכים, רישום הסטטוס MEP, CDE, ECM, CPM המשפטי של ההחלטות וניהול התהליכים.

לפני 4000 שנה, כאשר לצורך אישור משפטי של **Ancient** תהליך העסקאות דומה לעבודה של מנהלים החלטות השתמשו בלוחות חרס ובפפירוסים. ההבדל בין המערכות המודרניות לבין קודמותיהן החרסיות והנייריות הוא שהשיטות המודרניות כוללות בנוסף את תהליך המרת המידע הטקסטואלי לצורה דיגיטלית לצורך עיבוד אוטומטי במערכות ובכלים אחרים.

יצירת ויזואליזציה של התהליך בצורה של דיאגרמות קונספטואליות תסייע לתאר כל שלב ואינטראקציה בין תפקידים שונים, מה שהופך את תהליך העבודה המורכב להבנה ופשוט.

ויזואליזציה של תהליכים מספקת שקיפות ונגישות ללוגיקה של התהליך לכל חברי הצוות.

אותו תהליך תקשורתי של הוספת חלון לפרויקט, שתואר בטקסט, בהודעות (איור 4.41) ובדיאגרמה, דומה למודל הקונספטואלי שנדון בפרק על מודל הנתונים (איור 4.44).



איור 4.44 דיאגרמה הקונספטואלית, משתתפי הפרויקט מוצגים כמשתמשי בסיס הנתונים, כאשר הבקשות שלהם מקשרות בין מערכות שונות.

believing that a visual scheme is sufficient for understanding processes. קל לנהל את הנתונים, רבות מהחברות מסתפקות רק ברמה זו יותר לתפוס את התמונה הכוללת, לראות את הקשרים בין המשתתפים ושילוב התהליך. עם זאת, סכמות כאלה אינן מספקות תמונה ברורה לגבי אילו נתונים נדרשים לכל משתתף, באיזה פורמט הם צריכים להיות מועברים ואילו פרמטרים ואטריבוטים הם הכרחיים ליישום האוטומציה. הסכמה הקונספטואלית דומה יותר למפת מסלול: היא מצביעה על כך מי מתקשר עם מי, אך אינה חושפת מה בדיוק מועבר באותם אינטראקציות.

גם אם התהליך מתואר בפירוט ברמה הקונספטואלית באמצעות סכמות, זה לא מבטיח את היעילות של הוויזואליזציה לעיתים מפשטת את העבודה של המנהלים, ומאפשרת להם לעקוב אחרי התהליך בצורה נוחה יותר באמצעות מערכת דיווח שלב אחר שלב. עם זאת, עבור מהנדסים המנהלים את בסיסי הנתונים הייצוג הקונספטואלי עשוי להישאר לא ברור מספיק ולא לספק הבנה ברורה כיצד ליישם את התהליך ברמת הפרמטרים והדרישות.

ככל שמתקדמים למערכות נתונים מורכבות יותר, ההטמעה הראשונית של כלים קונספטואליים וויזואליים הופכת להיות קריטית כדי שהתהליכים לעיבוד נתונים יהיו לא רק יעילים, אלא גם יתאימו למטרות האסטרטגיות של הארגון. כדי

להעביר את התהליך הזה של הוספת חלון לרמת הדרישות לנתונים, עלינו לרדת לרמה עמוקה יותר ולהעביר את הוויזואליזציה הקונספטואלית של התהליך לרמה הלוגית והפיזית של הנתונים, האטריבוטים הנדרשים והערכים הגבוליים שלהם.

RegEx דרישות מובנות וביטויים רגולריים

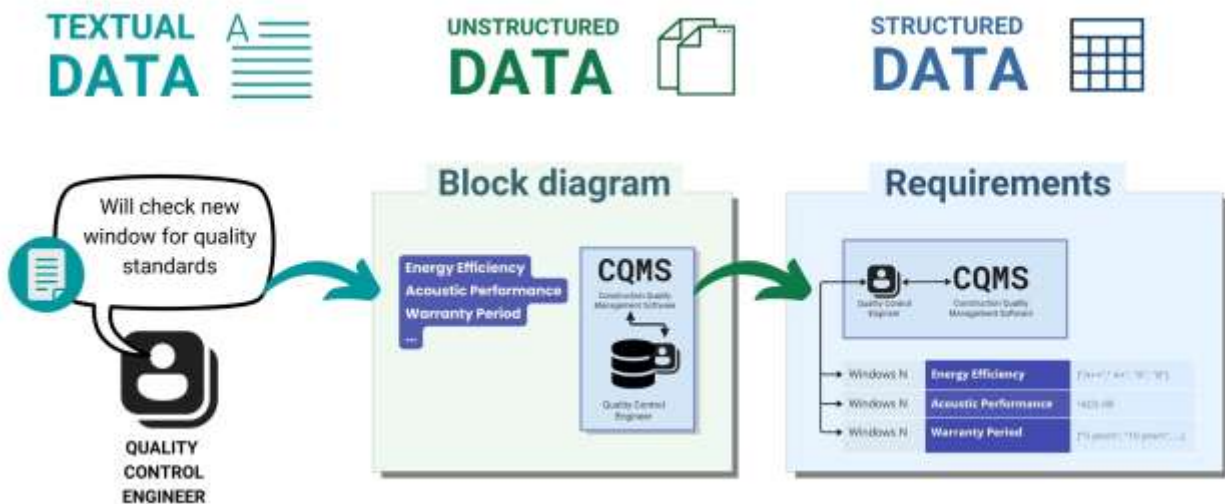
עד 80% מהנתונים הנוצרים בחברות מוצגים בפורמטים לא מובנים או חצי מובנים – טקסט, מסמכים, מכתבים, קבצי PDF. שיחות. נתונים כאלה קשים לניתוח, לבדוק, להעביר בין מערכות ולהשתמש בהם באוטומציה,

כדי להבטיח ניהול, שקיפות ואימות אוטומטי, יש צורך להמיר את הדרישות הטקסטואליות והחצי מובנות לפורמטים מובנים ומוגדרים בבירור. תהליך ההסדרה נוגע לא רק לנתונים (כפי שדנו בזה בפירוט בפרקים הראשונים של חלק זה של הספר, (אלא גם לדרישות עצמן, אשר משתתפי הפרויקטים בדרך כלל מנסחים בצורה חופשית של טקסט במהלך כל מחזור חיי הפרויקט, לעיתים מבלי לחשוב על כך שהתהליכים הללו ניתן לאוטומט

בדיוק כפי שכבר המנו את הנתונים מצורת טקסט לא מובנת לפורמט מובנה, במהלך העבודה על הדרישות "נמיר את הדרישות הטקסטואליות לפורמט מובנה של "רמה לוגית ופיזית

במסגרת הדוגמה של הוספת חלון, השלב הבא יהיה לתאר את הדרישות לנתונים בצורה טבלאית. נבנה את המידע עבור - כל מערכת, בה משתמשים משתתפי הפרויקט, ונציין את האטריבוטים המרכזיים ואת הערכים הגבוליים שלהם

אותה משתמש מהנדס בקרת איכות, (CQMS) נבחן, לדוגמה, אחת מהמערכות הללו – מערכת ניהול איכות הבנייה מטעם המזמין. בעזרתה הוא בודק אם האלמנט החדש של הפרויקט – במקרה זה "חלון חדש" – עומד בסטנדרטים ובדרישות שנקבעו



איור 4.45 המרת דרישות טקסט לפורמט טבלה עם תיאור של תכונות ישויות מקלה על ההבנה עבור אנשי מקצוע אחרים.

איור 4.46: (יעילות CQMS כדוגמה, נבחן כמה דרישות חשובות לתכונות של ישויות מסוג "מערכות חלון" במערכת אנרגטית, תכונות אקוסטיות ותקופת אחריות. כל קטגוריה כוללת תקנים ומפרטים מסוימים שיש לקחת בחשבון בעת -תכנון והתקנה של מערכות חלון



איור 4.46 מהנדס בקרת איכות צריך לבדוק את הרכיבים החדשים מסוג "חלון" לעמידה בתקני יעילות אנרגטית, בידוד רעש ושירות אחריות.

הדרישות לנתונים שהמהנדס לבקרת איכות קובע בפורמט טבלה כוללות, למשל, את הערכים הקיצוניים הבאים

■ הנחשבת, "B"-המייצגת את היעילות הגבוהה ביותר, ל, "A++" דרגת היעילות האנרגטית של החלונות נעה בין ["A++", "A+", "A", "A-", "B"]. לרמה המינימלית המותרת, ודירוגים אלו מיוצגים ברשימת ערכים מותרים

■ בידוד רעש של חלונות, הנמדד בדציבלים ומראה את יכולתם להפחית רעש חיצוני, מוגדר על ידי ביטוי רגולרי {2}dB.

■ תכונת "תקופת אחריות" עבור ישות "סוג חלון" מתחילה מחמש שנים, קובעת תקופה זו כמינימום המותר בבחירת המוצר; כמו כן, מצוינים ערכי תקופת האחריות, לדוגמה "5 שנים", "10 שנים" וכדומה [או תנאי "לוגי > 5" שנים).

חלון "עם דרגות נמוכות" class-בהתאם לדרישות שנאספו, במסגרת התכונות שנקבעו, רכיבים חדשים בקטגוריה או ב לא יעברו את הבדיקה ליעילות אנרגטית. בידוד רעש של חלונות בנתונים או במסמכים המגיעים, "D" או "C" כמו, "B"-מ או "40" B"לדוגמה, 35 ד, B"למהנדס בקרת איכות צריך להיות מסומן במספר דו-ספרתי, שאחריו מופיע הסיומת "ד (Regex-כיוון שלא יעברו את תבנית ה) או "100 דציבלים", לא יתקבלו B"והערכים מחוץ לפורמט זה, כמו "9 ד, B"ד תקופת האחריות צריכה להתחיל לפחות מ-"5 שנים", וחלונות עם תקופת אחריות קצרה יותר, כמו "3 שנים" או "4 שנים", לא יעמדו בדרישות שהמהנדס לבקרת איכות תיאר בפורמט טבלה

כדי לבדוק ערכים כאלה של תכונות-פרמטרים לעמידה בערכים הקיצוניים מהדרישות במהלך תהליך האימות, אנו (["A", "B", "C"], במילונים (["A": "H1", "H2"; "B": "W1", "W2"], משתמשים או ברשימת ערכים מותרים בפעולות לוגיות) כגון ">", "<", "<=", ">=", "==" ("ביטויים רגולריים הם כלי חשוב מאוד בעבודה עם ערכים מחרוזתיים וטקסטואליים כמו בתכונה "ביצועים אקוסטיים").

לחיפוש ושינוי (Re ספרייט) משמשים בשפות תכנות, כולל פייתון (RegEx) ביטויים רגולריים. הוא כמו בלש בעולם המחרוזות, המסוגל בדיוק לזהות תבניות טקסט בטקסט Regex. מחרוזות.

בביטויים רגולריים, אותיות מתוארות ישירות באמצעות הסימנים המתאימים של האלפבית, בעוד שמספרים יכולים שמתאים לכל ספרה בין 0 ל-9. סוגריים מרובעים משמשים לציון טווח של \d, להיות מיוצגים באמצעות הסימן המיוחד עבור סימנים שאינם \d-עבור כל אות קטנה של האלפבית הלטיני או [0-9], [a-z], אותיות או ספרות, לדוגמה. בהתאמה \W-וב \D-מספריים ואינם אותיים, משתמשים ב.

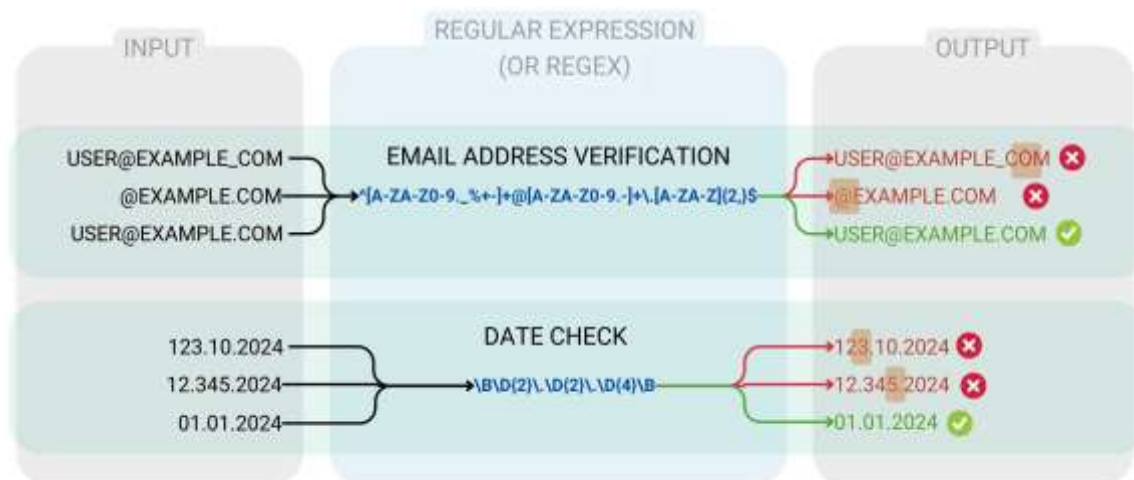
:- (4.47 איור) דוגמאות פופולריות לשימוש בביטויים רגולריים

■ "[a-zA-Z0-9._%+]+@[a-zA-Z0-9.]+\.[a-zA-Z]{2,}\$" בדיקת כתובת דוא"ל: כדי לבדוק אם מיתר הוא כתובת דוא"ל חוקית, ניתן להשתמש בתבנית

■ תבנית יכולה לשמש "^\b\d{2}\d{2}\d{2}\d{2}\d{2}\. \d{2}\. \d{4}\b" :חילוץ תאריך DD.MM.YYYY. לחילוץ תאריך מטקסט בפורמט

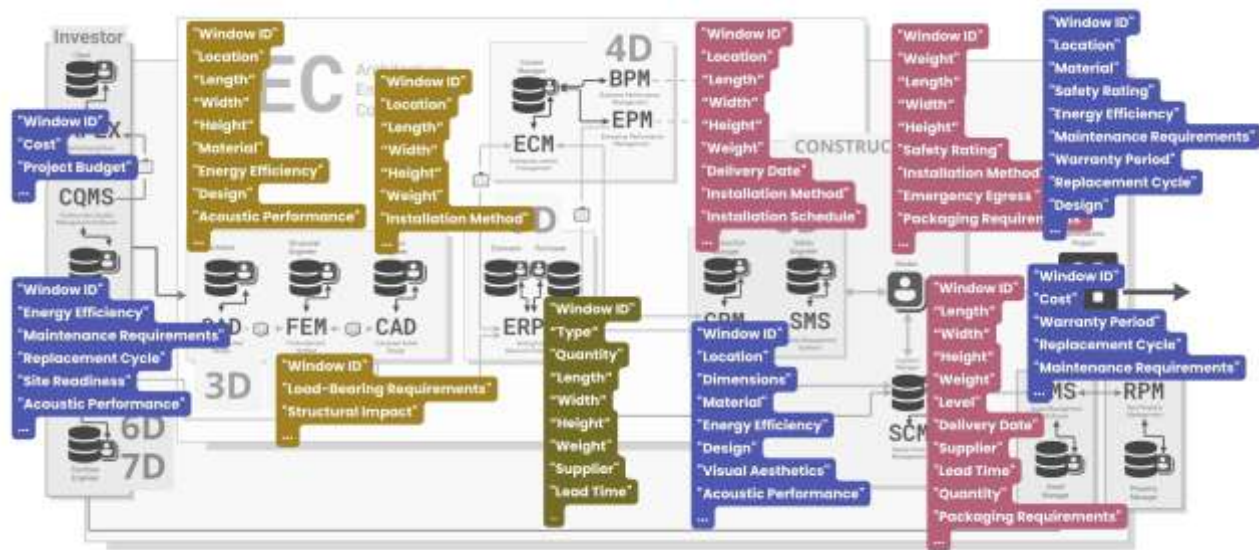
■ בדיקת מספרי טלפון: כדי לבדוק מספרי טלפון בפורמט +49(000-0000)000, התבנית תראה כך: "\+ \d{2} \(\d{3}\) \d{3}-\d{4}".

הפכנו אותן, (4.46 איור) על ידי תרגום הדרישות של המהנדס לבקרת איכות לפורמט של תכונות וערכי גבול מהפורמט הטקסטואלי המקורי) שיחות, מכתבים ומסמכים רגולטוריים (לטבלה מאורגנת ומסודרת, ובכך אפשרנו בדיקה אוטומטית וניתוח של כל הנתונים המתקבלים) למשל, רכיבים חדשים בקטגוריה "חלון". ("קיום הדרישות -מאפשר לסנן אוטומטית נתונים שלא עברו בדיקה, ונתונים שנבדקו מועברים אוטומטית למערכות לעיבוד נוסף



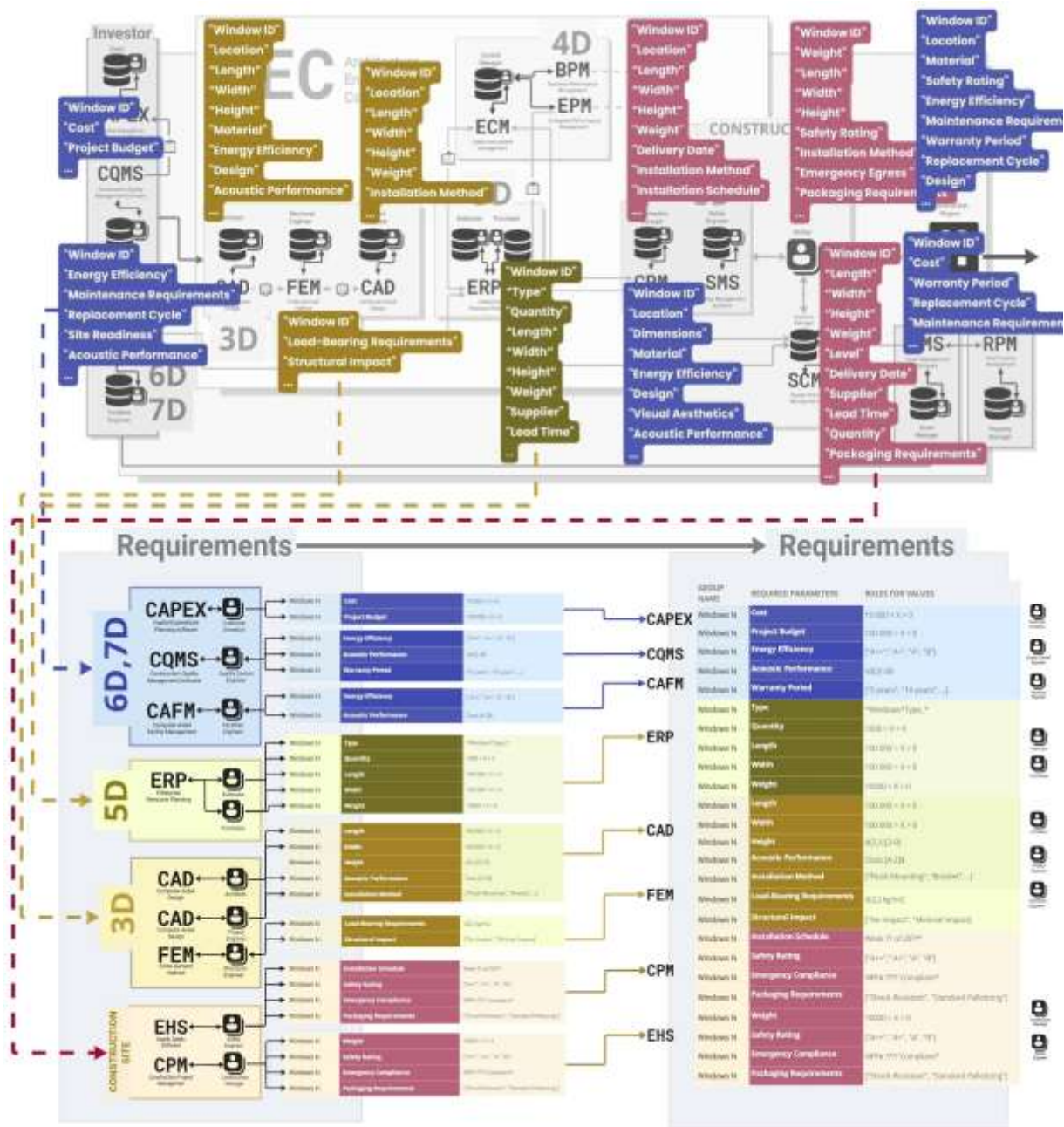
השימוש בביטויים רגולריים הוא כלי חשוב מאוד בתהליך בדיקות הנתונים הטקסטואליים 4.47 איור

עכשיו, כשאנחנו עוברים מהקונספטואלי לרמה הלוגית של העבודה עם הדרישות, נתרגם את כל הדרישות של כל רשימה מסודרת בפורמט של תכונות ונוסיף את הרשימות הללו (4.44 איור) המומחים בתהליך התקנת החלון החדש -- (4.48 איור) עם התכונות הנדרשות לדיאגרמת הזרימה שלנו עבור כל מומחה



ברמה הלוגית של התהליך, התכונות שמטופלות על ידי כל מומחה מתווספות למערכות המתאימות 4.48 איור.

לאחר שהוספנו את כל התכונות לטבלה כללית אחת של התהליך, הפכנו את המידע שהוצג קודם לכן בצורה של טקסט - (4.49 איור) לצורת טבלאות מאורגנת ומסודרת ברמה הפיזית (4.41 איור) ודיאלוג ברמה הקונספטואלית



המרת דיאלוג לא מובנה של המומחים לטבלאות מסודרות מסייעת להבין את הדרישות ברמה הפיזית 4.49 איור

עכשיו יש להעביר את הדרישות לנתונים למומחים שמייצרים מידע עבור מערכות ספציפיות. לדוגמה, אם אתה עובד לפני שתתחיל לדגם רכיבים, יש לאסוף את כל הפרמטרים הנדרשים, בהתבסס על תרחישי השימוש, CAD בבסיס נתונים, הסופיים של הנתונים הללו. בדרך כלל זה מתחיל בשלב ההפעלה, לאחר מכן מגיע לאתר הבנייה, מחלקת הלוגיסטיקה מחלקת האומדן, מחלקת החישובים המבניים וכן הלאה. רק לאחר שתקח בחשבון את הדרישות של כל אחד מהקשרים הללו, תוכל להתחיל ליצור את הנתונים - בהתבסס על הפרמטרים שנאספו. זה יאפשר בהמשך לאוטומט את הבדיקה והעברת הנתונים בשרשרת.

כאשר הנתונים החדשים תואמים לדרישות שנקבעו, הם משתלבים אוטומטית במערכת הנתונים של החברה, ומופנים

ישירות למשתמשים ולמערכות שאליהם הם מיועדים. אימות הנתונים לגבי קיום ותאימות של תכונות וערכיהן מבטיח שהמידע עומד בסטנדרטים הנדרשים של איכות ומוכן לשימוש בתרחישים של החברה.

דרישות הנתונים הוגדרו, ועכשיו, לפני שמתחילים באימות, יש ליצור, לקבל או לאסוף את הנתונים שצריכים להיות מאומתים, או לתעד את המצב הנוכחי של המידע בבסיסי הנתונים, כדי להשתמש בו בתהליך האימות.

איסוף נתונים לתהליך האימות

לפני שמתחילים באימות, חשוב לוודא שהנתונים זמינים בצורה המתאימה לתהליך האימות. זה לא אומר רק שהמידע קיים, אלא שהוא מוכן: יש לאסוף ולהמיר את הנתונים מצורות לא מובנות, חלשות מובנות, טקסטואליות וגיאומטריות לצורה מובנית. תהליך זה מתואר בפירוט בפרקים הקודמים, שבהם נדונו שיטות המרה של סוגי נתונים שונים. כתוצאה מכל ההמרות, הנתונים הנכנסים מקבלים צורת טבלאות מובנות פתוחות.

עם הדרישות ועם הטבלאות המובנות עם הפרמטרים והערכים הקצה הנדרשים, אנו יכולים להתחיל באימות הנתונים והן בפורמט של אימות שלב-שלב של כל מסמך נכנס (Pipeline) הן כהליך אוטומטי אחד.

כדי להתחיל את האימות, יש לקבל קובץ חדש או לתעד את המצב הנוכחי של הנתונים - ליצור צילום או לבצע ייצוא של המידע הנוכחי והנכנס, או להגדיר חיבור לבסיס נתונים חיצוני או פנימי. בדוגמה הנדונה, צילום כזה נוצר על ידי המרה לפורמט מובנה, שנשמר, נניה, בשעה 23:00:00 ביום שישי, 29 במרץ 2024, לאחר שכל CAD אוטומטית של נתוני המהנדסים עזבו הביתה.



המראה את המידע הנוכחי על תכונות עבור ישות חדשה מסוג "חלון" בגרסה הנוכחית של **CAD (BIM)** צילום של בסיס הנתונים מודל הפרויקט.

CAD לצורה מובנית, "מידע זה מכילים ומעורכי **CAD (BIM)** בזכות כלים של הנדסה לאחר, שנדונו בפרק "המרת נתונים שונים יכול להיות מאורגן בטבלאות נפרדות או מאוחד בטבלה אחת כללית, המאחדת את חלקי הפרויקט השונים (**BIM**)".

(**TypeName**) שמות טיפוסיים, (**ID** מזהה) בטבלה כזו - בבסיס הנתונים מוצגים מזהים ייחודיים של חלונות ודלתות. כמו גם מדדים של יעילות אנרגטית ואקוסטית, ותכונות נוספות, (**Material**) חומרים, (**Width, Length**) מידות נאספת על ידי המהנדס המהנדס ממחלקות וממסמכים שונים, ומבנה את המודל, **CAD (BIM)** טבלה כזו, המלאה בתוכנית.

המידע של הפרויקט.

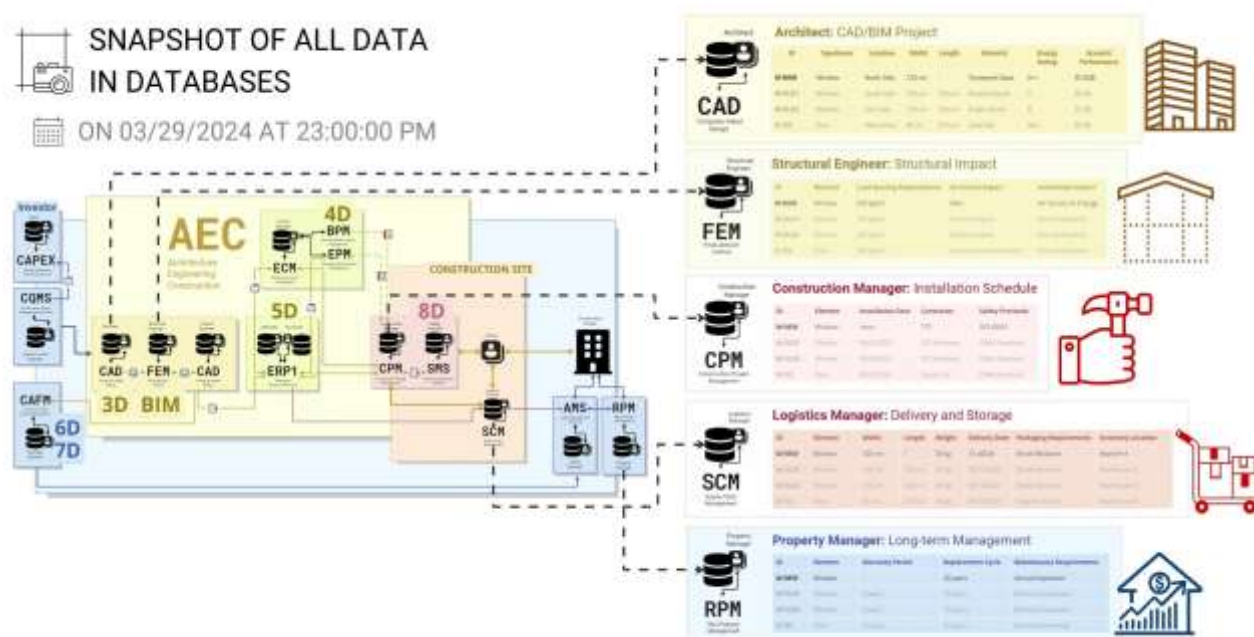


יכולים להציג טבלה דו-ממדית עם עמודות המייצגות את תכונות האלמנטים CAD נתונים מובנים ממערכות.

מסווגים CAD כוללים עשרות או מאות אלפי רכיבים. רכיבים בתוך פורמטי (BIM) CAD פרויקטים אמיתיים של אוטומטית לפי סוגים וקטגוריות - מחלונות ודלתות ועד לרצפות, תקרות וקירות. מזהים ייחודיים (כגון מזהה מקורי או תכונות מסוג) שם סוג, סוג, משפחה (מאפשרים לעקוב אחר אותו אובייקט CAD) המוקצה אוטומטית על ידי פתרון "W-NEW" במערכות שונות. כך, חלון חדש בקיר הצפוני של הבניין יכול להיות מסומן באופן חד משמעי באמצעות מזהה - בכל המערכות הרלוונטיות של הארגון.

למרות ששמות ומזהים של ישויות צריכים להיות אחידים בכל המערכות, קבוצת התכונות והערכים הקשורים לישויות אלו עשויה להשתנות באופן משמעותי בהתאם להקשר השימוש. אדריכלים, מהנדסי בניין, מומחים לבנייה, לוגיסטיקה ותפעול נכסים תופסים את אותם רכיבים בצורה שונה. כל אחד מהם מתבסס על מסווגים, תקנים ומטרות משלו: משיהו רואה את החלון רק מנקודת מבט אסתטית, מעריך את צורתו ופרופורציותיו, בעוד שמישהו אחר רואה אותו מנקודת מבט הנדסית או תפעולית, מנתח את מוליכות החום, שיטת ההתקנה, משקל או דרישות לתחזוקה. לכן, בעת מודליזציה של נתונים ותיאור רכיבים, חשוב לקחת בחשבון את הרב-גוניות של השימוש בהם ולהבטיח עקביות בנתונים תוך כדי התחשבות בהיבטים תעשייתיים.

עבור כל תפקיד בתהליכי החברה קיימות מסדי נתונים מיוחדים עם ממשק משתמש ייחודי - מעיצוב וחישובים ועד לוגיסטיקה, התקנה ותפעול של בניינים. כל מערכת כזו מנוהלת על ידי צוות מקצועי של מומחים דרך ממשק משתמש מיוחד או דרך בקשות למסד הנתונים, כאשר מאחורי סך כל ההחלטות המתקבלות על סמך הערכים שהוזנו בסוף השרשרת עומד מנהל המערכת או ראש מחלקה, האחראי על הלגיטימיות ואיכות הנתונים שהוזנו מול שותפיו, המשרתים מערכות אחרות.



- אותה ישות יש לה מזהה זהה במערכות שונות, אך תכונות שונות, החשובות רק במערכת זו

לאחר איסוף מאורגן של דרישות ונתונים מובנים ברמה הלוגית והפיזית, נותר לנו להגדיר את תהליך בדיקת הנתונים האוטומטית, המתקבלים ממסמכים שונים וממערכות שונות, על מנת לוודא שהם עומדים בדרישות שנאספו קודם לכן

אימות נתונים ותוצאות האימות

כל נתון חדש המתקבל במערכת - בין אם מדובר במסמכים, טבלאות או רשומות במסדי נתונים מהלקוח, האדריכל, המהנדס, המנהל, הלוגיסט או מנהל הנכסים - חייב לעבור בדיקה על מנת לוודא שהוא עומד בדרישות שהוגדרו קודם לכן. תהליך האימות הוא קריטי: כל טעות בנתונים עלולה להוביל לחישובים שגויים, עיכובים בלוח הזמנים ואפילו להפסדים כספיים. כדי למזער סיכונים כאלה, יש לארגן הליך בדיקה שיטתי וחוזר, איטרטיבי.

כדי לבדוק נתונים חדשים המתקבלים במערכת - לא מובנים, טקסטואליים או גיאומטריים - יש להמירם לפורמט חלש-מובנה או מובנה. לאחר מכן, במהלך תהליך הבדיקה, יש לבדוק את הנתונים על מנת לוודא שהם עומדים ברשימה המלאה של התכונות הנדרשות וערכיהן המותרים

לצורה מובנית נבחרה CAD (BIM) ונתונים מעורבים PDF המרה של סוגי נתונים שונים: טקסטים, תמונות, מסמכי "בפירוט בפרק" המרת נתונים לצורה מובנית

היא כוללת נתונים גיאומטריים חצי- (4.411 איור) CAD (BIM) כדוגמה ניתן להביא את הטבלה שהתקבלה מפרויקט -- "לדוגמה, רכיב מקטגוריית "חלונות" - (3.114 איור) מובנים ומידע אטריבוטי מובנה על ישויות הפרויקט

עם ערכי גבול ייחודיים שהוגדרו על ידי מומחים (4.411 איור) לצורך ביצוע בדיקה, אנו משווים את ערכי האטריבוטים

תאפשר להבין אילו ערכים מקובלים ואילו (4.413 איור) הטבלה הסופית להשוואה. (4.49 איור) בצורה של דרישות CAD (BIM).- זקוקים לתיקון, לפני שניתן יהיה להשתמש בנתונים מחוץ ליישומי



הטבלה הסופית לבדיקה מדגישה את ערכי האטריבוטים עבור ישות חדשה מקטגוריית "חלונות", שעליהם יש לשים 4.413 איור לב.

כלי חיוני לניתוח "Pandas: עליה דיברנו קודם בפרק Pandas, על מנת ליישם פתרון כזה, תוך שימוש בספריית (4.411 איור) CAD (BIM) (RVT, IFC, DWG, NWS, DGN) נתונים, "נבדוק את הנתונים מקובץ הטבלה שהופק מקובץ -- (4.49 איור) תוך שימוש בדרישות מקובץ טבלה אחר, (4.411

הסט המלא של נתונים) raw_data.xlsx מה יש לטעון מהקובץ LLM כדי לקבל את הקוד, עלינו לתאר בפרומפט עבור (4.413 איור) checked_data.xlsx לבדוק אותם ולשמור את התוצאה בקובץ חדש, (CAD (BIM) ממסד הנתונים

Pandas: מבלי להזכיר את ספריית LLM נקבל קוד בעזרת

ובדוק אותם לפי הכללים הבאים: ערכי raw_data.xlsx כתוב קוד לבדוק את הטבלה מקובץ 'Energy Rating', גדולים מאפס 'Length' ו-'Width' העמודות 'A+', 'A++', 'A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J', 'K', 'L', 'M', 'N', 'O', 'P', 'Q', 'R', 'S', 'T', 'U', 'V', 'W', 'X', 'Y', 'Z' כמשתנה שנוצרים מאוחר יותר - עם הוספת עמודות בדיקה 'Acoustic Performance', 'A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J', 'K', 'L', 'M', 'N', 'O', 'P', 'Q', 'R', 'S', 'T', 'U', 'V', 'W', 'X', 'Y', 'Z' checked_data.xlsx חדש Excel סופית, ושמור את הטבלה הסופית בקובץ

שניתן לדייק ולהשלים בפרומפטים הבאים, Python תתאר דוגמה קצרה לקוד LLM התגובה של ?



```

1 import pandas as pd
2
3 # Load raw data from Excel file
4 df_data = pd.read_excel('raw_data.xlsx', sheet_name=0)
5
6 # Adding a 'Requirements Met' column that checks all conditions
7 df_data['Requirements Met'] = (
8     (df_data['Width'] > 0) &
9     (df_data['Length'] > 0) &
10    (df_data['Energy Rating'].isin(['A++', 'A+', 'A', 'B'])) &
11    (df_data['Acoustic Performance'] >= min_acoustic_performance)
12 )
13
14 # Save the results to a new Excel file
15 df_data.to_excel('checked_data.xlsx', index=False)

```

המומר על פי דרישות האטריבוטים בצורה של ערכי גבול (BIM) CAD בודק את פרויקט LLM הקוד שנוצר על ידי מודל 4.414 איור.

PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse עם TOSK PyDev, Thonny, Wing IDE, IntelliJ IDEA עם TOSK Python, JupyterLab כמו Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker.

(CAD) ועוד רכיבים (ממסד הנתונים W-OLD1, W-OLD2, D-122 "יראה כי" הישויות (4.414 איור) ביצוע הקוד 'A++', עומדים בדרישות האטריבוטים: רוחב ואורך גדולים מאפס, וכיתה האנרגיה היא אחת מהערכים ברשימה (BIM) 'A', 'B', 'C' (4.415 איור).-

המייצג את הרכיב החדש מקטגוריית "חלון" בצד הצפוני, אינו עומד בדרישות) אטריבוט, W-NEW הרכיב הנדרש והחדש ואין בו 'Width' > 0 ערך "0.0" נחשב לבלתי מקובל לפי הכלל שלנו) דרישות מולאו, ("מכיוון שאורכו שווה לאפס" ציון של כיתה אנרגטית



CHECKED_DATA.XLSX (CSV)

VERIFIED DATA

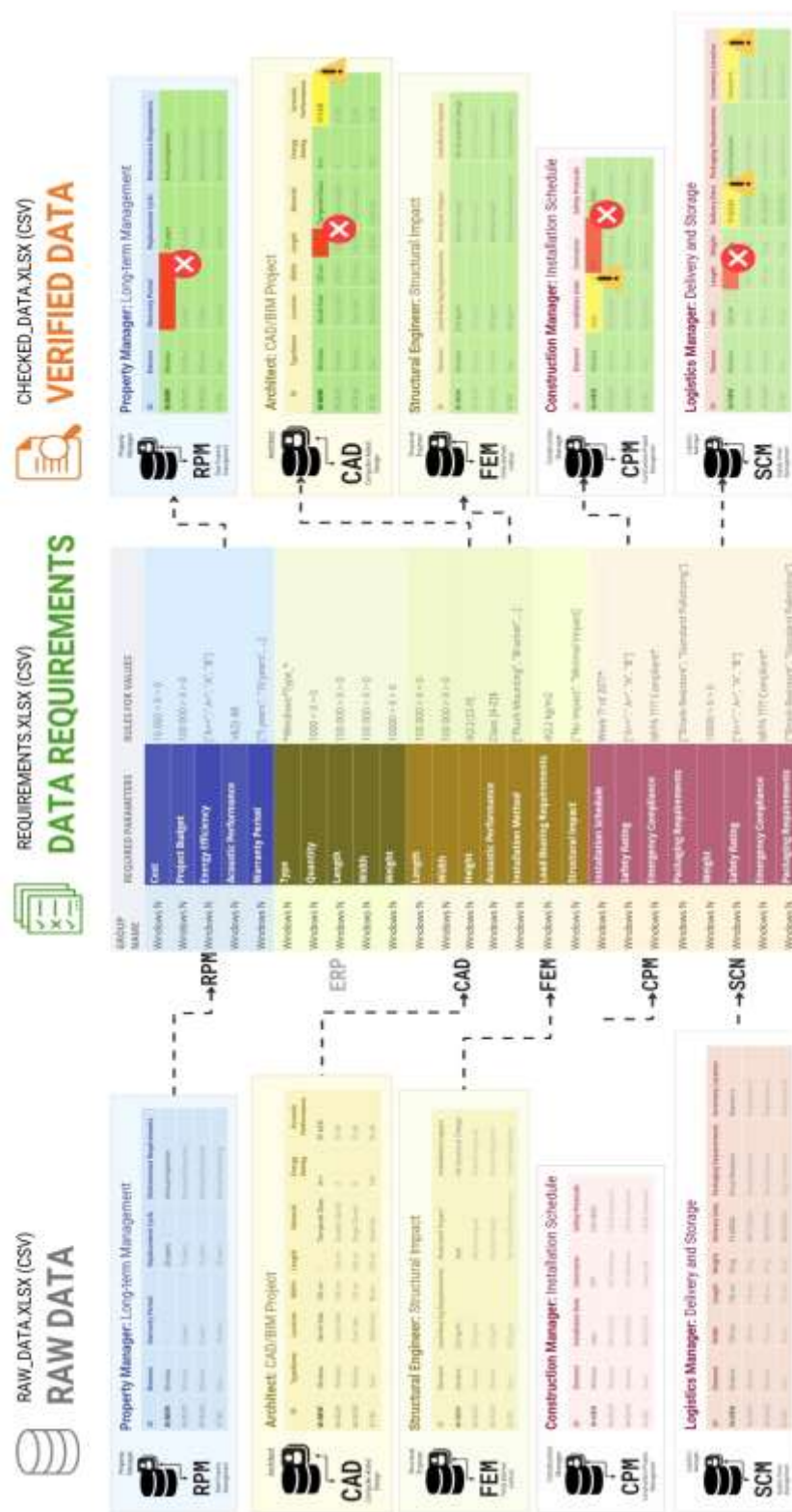
	ID	TypeName	Location	Width	Length	Material	Energy Rating	Acoustic Performance	Requirements Met
0	W-NEW	Window	North Side	120	0.0	Tempered Glass		35	False
1	W-OLD1	Window	South Side	100	140.0	Double Glazed	A++	30	True
2	W-OLD2	Window	East Side	110	160.0	Single Glazed	B	25	True
3	D-122	Door	Main Entry	90	210.0	Solid Oak	B	30	True

הבדיקה מזהה ישויות שלא עברו את תהליך האימות ומוסיפה לתוצאות מאפיין חדש עם ערכים 'שקרי' או 'נכון' 4.415 איור

באופן דומה, אנו בודקים את העקביות של כל רכיבי הפרויקט (ישויות) ואת המאפיינים הנדרשים עבור כל אחת במהלך תהליך הוספת (4.41 איור) מהמערכות, הטבלאות או מסדי הנתונים בכל הנתונים שאנו מקבלים ממומחים שונים חלון לפרויקט.

בטבלה הסופית, נוח להדגיש את תוצאות הבדיקה בצבעים: מאפיינים שעברו בהצלחה את הבדיקה מסומנים בירוק, ערכים (4.416 איור) עם סטיות לא קריטיות מסומנים בצהוב, ואילו אי התאמות קריטיות מסומנות באדום.

אנו מקבלים רשימה של רכיבים מהימנים ומאומתים עם מזהים שלהם, אשר נבדקו (4.416 איור) כתוצאה מהבדיקה על פי דרישות המאפיינים. רכיבים מאומתים מספקים ביטחון בכך שהרכיבים הללו עומדים בסטנדרטים ובמפרטים המוצהרים עבור כל המערכות המעורבות בתהליך הוספת רכיבי הקטגוריה "חלון" או כל קטגוריה אחרת (נרחיב על ("ובדיקות נתונים ETL בפרק "אוטומציה של ETL אוטומציה של בדיקות נתונים ויצירת תהליך אוטומטי



תוצאת הבדיקה שנערכה עבור כל המערכות מאפשרת לקבוע אילו נתונים אינם עומדים בדרישות החברה **4.416** איור.

יישויות שעברו בהצלחה את הבדיקה בדרך כלל אינן דורשות תשומת לב מוגברת. הן עוברות בקלות לשלב הבא של עיבוד

ואינטגרציה במערכות אחרות. בניגוד לרכיבים "איכותיים", רכיבים שלא עברו את הבדיקה מהווים את העניין המרכזי המידע על סטיות כאלה הוא קריטי: יש להעביר אותו לא רק בצורה של דוחות טבלאיים, אלא גם באמצעות אמצעי ויזואליזציה שונים. הצגת תוצאות הבדיקה בצורה גרפית מסייעת להעריך במהירות את מצב איכות הנתונים, לזהות אזורים בעייתיים ולנקוט צעדים מהירים לתיקון או הבהרת פרמטרים.

ויזואליזציה של תוצאות האימות

ויזואליזציה היא כלי מרכזי בפרשנות תוצאות הבדיקה. מעבר לטבלאות סיכום מוכרות, היא יכולה לכלול לוחות מידע המיוצרים אוטומטית, שבהם רכיבי הפרויקט מקובצים לפי מצב הבדיקה. קידוד צבעים יכול PDF תרשימים ומסמכי לשחק כאן תפקיד עזר: ירוק יכול לסמן אובייקטים שנבדקו בהצלחה, צהוב יכול לסמן רכיבים שדורשים תשומת לב נוספת, ואדום יכול לסמן את אלה שבהם זוהו שגיאות קריטיות או חסרים נתונים חיוניים.

ונייהול נכסים ועד (BIM) CAD-אנו מנתחים שלב אחר שלב את הנתונים מכל מערכת: מ (4.41 איור) בדוגמה שלנו בעקבות הביקורת, עבור כל מומחה נוצרים אוטומטית התראות או (4.416 איור) לוגיסטיקה ולוחות זמנים של התקנה אם הנתונים נכונים, המומחה מקבל הודעה קצרה: תודה על (4.417 איור) PDF מסמכי דוח אישיים, לדוגמה בפורמט שיתוף הפעולה. "במקרה של גילוי אי התאמות, נשלח דוח מפורט עם הניסוח: במסמך זה מפורטים רכיבים, מזהים --"שלהם, מאפיינים וערכים שלא עברו את הבדיקה על פי דרישות



אימות ויצירה אוטומטית של מסמכי דוח מאיצים את תהליך החיפוש והבנת החסרונות בנתונים עבור המומחה 4.417 איור שמייצר את הנתונים.

בזכות תהליך האימות האוטומטי - ברגע שמתגלה שגיאה או חוסר בנתונים, נשלחת מיד הודעה בצורת הודעת צ'אט לאדם האחראי על יצירה או עיבוד של ישויות רלוונטיות ואטריבוטים שלהן (איור 4.418, (עם PDF דוא"ל או מסמך - רשימה של פריטים ותיאורי אטריבוטים שלא עברו את האימות



איור 4.418 דוחות אוטומטיים על תוצאות האימות מקלים על הבנת השגיאות ומאיצים את העבודה על מילוי נתוני הפרויקט.

לדוגמה, אם למסלול ניהול הנדל"ן נכנס (לאחר מבנה) מסמך שבו אטריבוט "תקופת אחריות" מולא בצורה שגויה, מנהל הנדל"ן מקבל התראה עם רשימת האטריבוטים שצריך לבדוק ולתקן.

באופן דומה, כל פגם בלוח הזמנים של ההתקנה או בנתוני הלוגיסטיקה מביא ליצירת דוח אוטומטי ולשליחת הודעה בצ'אט או דוא"ל עם תוצאות האימות למומחה הרלוונטי.

ולגופים עם התוצאות, ניתן ליצור לוחות מחוונים (דשבורדים) (ודגמים אינטראקטיביים בתלת PDF בנוסף למסמכי מימד) (איור 7.16, איור 7.212) (עם הדגשה של פריטים עם אטריבוטים חסרים, מה שמאפשר למשתמשים להשתמש בצורה ויזואלית בגיאומטריות תלת מימדיות של פריטים כדי לסנן ולהעריך את איכות ושלמות הנתונים של הפריטים -- בפרויקט.

הוויזואליזציה של תוצאות האימות בצורה של מסמכים, גרפים או לוחות מחוונים המיוצרים אוטומטית מקלה מאוד על פרשנות הנתונים ותורמת לאינטראקציה יעילה בין משתתפי הפרויקט.

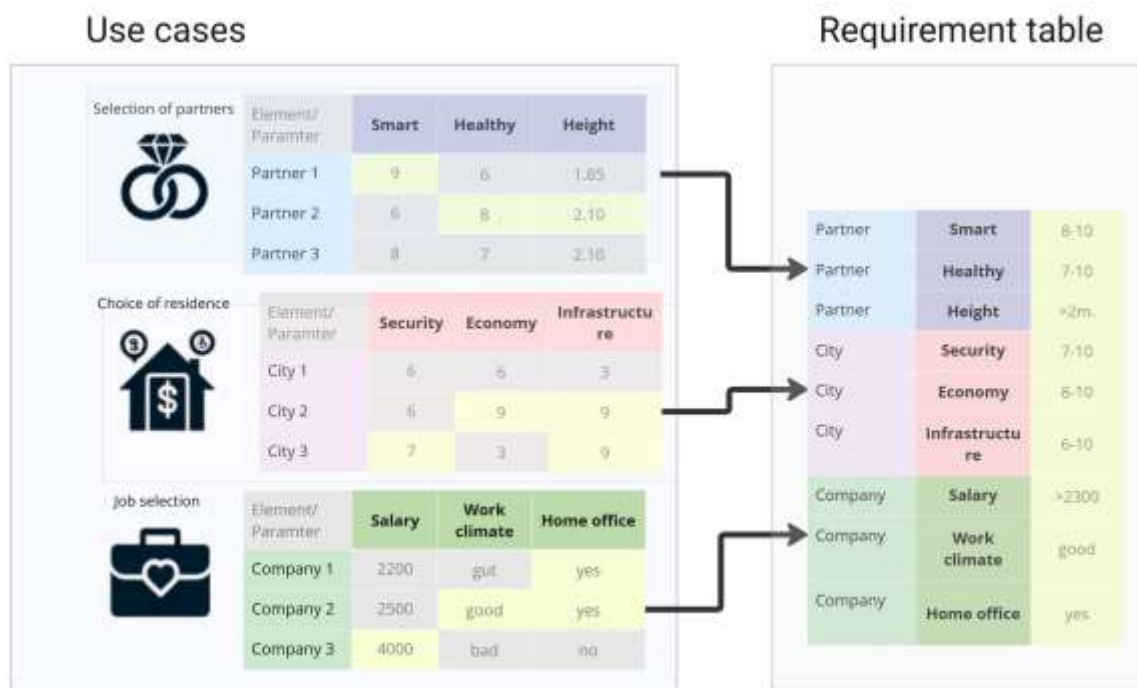
תהליך האימות האוטומטי של נתונים ממערכות ומקורות מידע שונים ניתן להשוואה לקבלת החלטות מודעת בחיי היומיום. בדומה לכך, חברות בתעשיית הבנייה לוקחות בחשבון מגוון משתנים - מהאמינות של הנתונים המקוריים ועד השפעתם על לוחות הזמנים, העלויות ואיכות ביצוע הפרויקט - כך גם אדם, כאשר הוא מקבל החלטות חשובות, כמו בחירת מקום מגורים, שוקל שורה של גורמים: נגישות תחבורתית, תשתיות, עלות, ביטחון, איכות חיים. כל השיקולים הללו יוצרים מערכת קריטריונים שעל בסיסם מתקבלות ההחלטות הסופיות שמרכיבות את חיינו.

השוואת בדיקות איכות נתונים לצרכים חיים של בני אדם

למרות ההתפתחות המתמדת של שיטות וכלים לבקרת איכות הנתונים, העיקרון הבסיסי - התאמת המידע לדרישות שנקבעו. נשאר ללא שינוי. עיקרון זה משולב בליבת מערכת ניהול בשלה, בין אם בעסקים ובין אם בחיי היומיום -

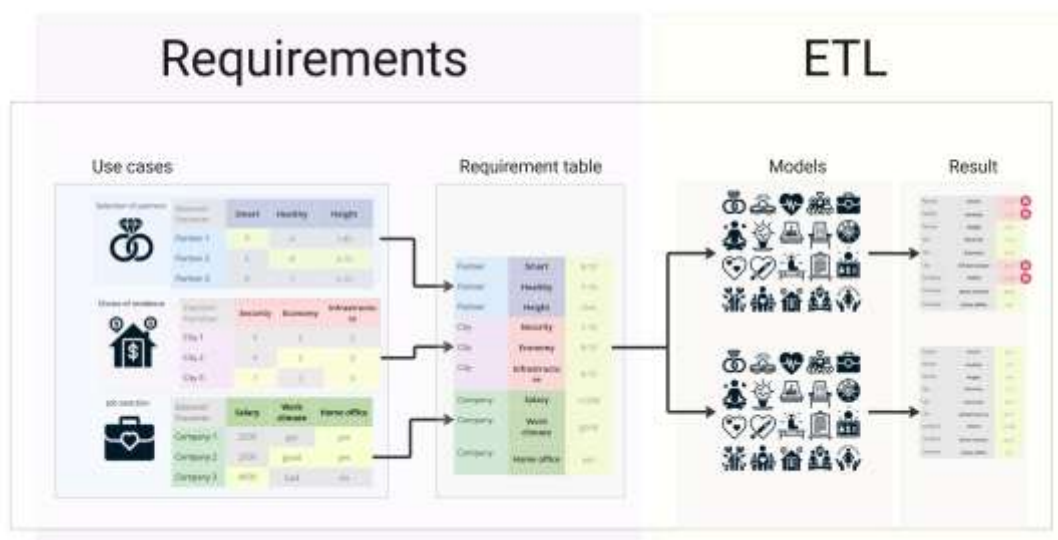
תהליך האימות האיטרטיבי של נתונים דומה במידה רבה לתהליך קבלת ההחלטות שבו נתקל כל אדם מדי יום. בשני המקרים אנו נשענים על ניסיון מצטבר, נתונים קיימים ומידע חדש שמתקבל. יתרה מכך, יותר ויותר החלטות חיים ומקצועיות - מהחלטות אסטרטגיות ועד החלטות יומיומיות - מתקבלות על בסיס נתונים.

כך, לדוגמה, כאשר אנו בוחרים מקום מגורים או בן זוג לחיים, אנו באופן אינטואיטיבי יוצרים במחשבה שלנו טבלה מסוימת עם קריטריונים ותכונות, על פיהן אנו משווים (חלופות) איור 4.419. (תכונות אלו - בין אם מדובר בתכונות אישיות של אדם או בפרמטרים של נכס נדל"ן - מהוות אטריבוטים המשפיעים על ההחלטה הסופית.



הבחירה של מקום מגורים, עבודה או שותפות מתבססת על דרישות אישיות לאטריבוטים 4.419 איור.

תורם לבחירה מבוססת יותר ומודעת יותר (איור 4.420) השימוש בנתונים מובנים ובגישה פורמלית לתיאור דרישות הן בפעילות מקצועית והן בחיים האישיים.



הפורמליזציה של הדרישות מאפשרת לסדר את התפיסה של החלטות חיים ועסקיות 4.420 איור.

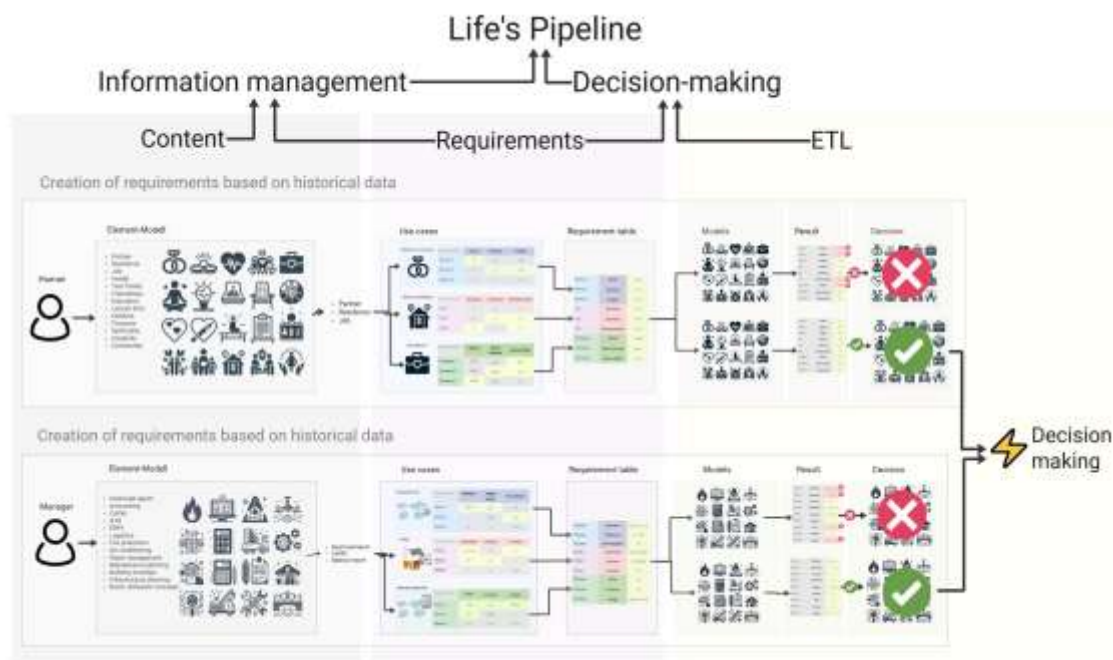
הגישה לקבלת החלטות מבוססת נתונים אינה כלי עסקי בלבד. היא משתלבת באופן אורגני בחיים היומיומיים, בעקבות שכבר דנו, ETL (Extract, Transform, Load) הדומים לתהליך, (4.421 איור) שלבים כלליים של עיבוד נתונים -בו בתחילת חלק זה בעת סידור הנתונים ונעסוק בו בפירוט בהקשר של אוטומציה של משימות בחלק השביעי של הספר

■ בכל תחום - בין אם בעבודה או בחיים האישיים - אנו אוספים מידע. בעסק זה: (Extract) נתונים כבסיס יכול להיות דוחות, מדדים, נתוני שוק; בחיים האישיים - ניסיון אישי, עצות של קרובים, ביקורות, תצפיות

■ - המידע שנאסף מפורש על בסיס קריטריונים שנקבעו מראש. בעבודה: (Transform) קריטריונים להערכה, מגבלות תקציב וסטנדרטים; בחיים האישיים - פרמטרים כמו מחיר, נוחות מיקום, (KPI) אלו מדדי ביצוע 'אמינות, כריזמה וכו

■ בשלב הסופי מתקבלת החלטה, המבוססת על ניתוח הנתונים המומרות: (Load) חיזוי וניתוח סיכונים והשוואת ההשלכות האפשריות. זה דומה לתהליכים עסקיים, שבהם הנתונים עוברים דרך מסנן של לוגיקה עסקית וסיכונים

ההחלטות שאנו מקבלים - מהעדפות פשוטות כמו מה לאכול לארוחת בוקר ועד אירועים חשובים בחיים כמו בחירת קריירה או בן זוג - הן במהותן תוצאה של עיבוד והערכה של נתונים



עסקים וחיים בכלל - זו סדרה של החלטות מבוססות נתונים, שבהן איכות הנתונים המשמשים לקבלת החלטות היא **4.421** איור גורם מפתח.

הכל בחיינו קשור זה לזה, וכמו אורגניזמים חיים, כולל האדם, פועלים לפי חוקי הטבע, מתפתחים ומסתגלים לתנאים משתנים, כך גם התהליכים האנושיים, כולל שיטות איסוף וניתוח נתונים, משקפים את העקרונות הטבעיים הללו. הקשר, ההדוק בין הטבע לפעילות האנושית מאשר לא רק את התלות שלנו בטבע, אלא גם את הרצון שלנו ליישם את החוקים שנחצבו במשך מיליוני שנות אבולוציה, ביצירת ארכיטקטורות נתונים, פיתוח תהליכים ומערכות לקבלת החלטות.

שוב ושוב מהטבע כדי **Inspired** טכנולוגיות חדשות, במיוחד בתחום הבנייה, הן דוגמה בולטת לכך שהאנושות מת ליצור פתרונות הטובים ביותר, יותר בני קיימא ויעילים.

צעדים נוספים: הפיכת נתונים לחישובים מדויקים ותוכניות

בחלק זה דנו כיצד להמיר נתונים לא מובנים לפורמט מובנה, לפתח מודלים של נתונים ולארגן תהליכי בדיקת איכות מידע בפרויקטים בנייה. ניהול נתונים, סטנדרטיזציה וסיווג שלהם הם תהליכים בסיסיים, הדורשים גישה מערכתית והבנה ברורה של דרישות העסק. השיטות והכלים שנדונו בחלק זה מאפשרים להבטיח אינטגרציה אמינה בין מערכות שונות במהלך כל מחזור חיי הפרויקט.

לסיכום חלק זה, את הצעדים המעשיים העיקריים שסייעו ליישם את הגישות שנדונו במשימות היומיומיות שלכם:

■ התחילו בסיסטמטיזציה של הדרישות

☐ צרו רישום של תכונות ופרמטרים עבור רכיבי המפתח של הפרויקטים והתהליכים שלכם

☐ תעדו את הערכים הקיצוניים עבור כל תכונה

☐ ויזואליזו את התהליכים והקשרים בין הקטגוריות, המערכות והתכונות באמצעות דיאגרמות (למשל ב

Miro, Canva, Visio)

■ אוטומטו את המרת הנתונים

☐ בדקו אילו מהמסמכים שלכם, הנמצאים בשימוש תדיר בתהליכים, ניתן לדיגיטציה באמצעות ספריות ולהמירם לפורמט טבלאי OCR

☐ CAD (BIM)-הכירו את הכלים להנדסה הפוכה כדי לחלץ נתונים מ

☐ נסו להגדיר קבלת נתונים אוטומטית מהמסמכים או הפורמטים שאתם משתמשים בהם לעיתים קרובות בעבודתכם, לפורמט טבלאי

☐ הגדרו המרות אוטומטיות בין פורמטים שונים של נתונים

■ צרו בסיס ידע לסיווג

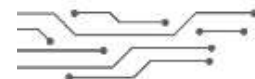
☐ פתחו מסווג פנימי או השתמשו במסווג קיים, המוסכם עם הסטנדרטים התעשייתיים

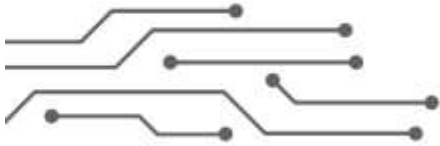
☐ תעדו את הקשרים בין מערכות הסיווג השונות

☐ דונו עם הצוות שלכם בנושא השימוש במערכת זיהוי אחידה וסיווג חד משמעי של רכיבים

☐ התחילו לבנות תהליך של בדיקה אוטומטית של נתונים - הן של הנתונים שאתם עובדים איתם בתוך הצוות והן של הנתונים המועברים למערכות חיצוניות

השימוש בגישות כאלה יאפשר לכם לשפר באופן משמעותי את איכות הנתונים ולפשט את העיבוד וההמרה שלהם בהמשך. בחלקים הבאים של הספר נבחן כיצד ליישם נתונים שכבר הובנו ומוכנים לחישובים אוטומטיים, הערכת עלויות, תכנון לוח זמנים וניהול פרויקטי בנייה





V חלק בנייה

חשובי עלות וזמן: הטמעת נתונים בתהליכי

החלק החמישי מוקדש להיבטים מעשיים של שימוש בנתונים לצורך אופטימיזציה של חישובי עלויות ותכנון פרויקטים בנייה. ניתוח מעמיק של שיטת המשאבים להכנת אומדנים ואוטומציה של ממודלים (Quantity Take-Off) תהליכי חישוב. נבחנות שיטות להשגת כמויות אוטומטיות D-5 ו-D-8 ואינטגרציה שלהם עם מערכות חישוב. נבדקות טכנולוגיות של דימוי 4 (BIM) CAD לתכנון פרמטרים זמניים וניהול עלויות הבנייה, עם דוגמאות קונקרטיות לשימושן. מוצג ניתוח של המבטיות גישה כוללת להערכת עמידות, תפעול ובטיחות של D-8D שכבות מידע מורחבות 6 של פרויקטים בנייה ESG-נכסי גדל"ן. נבחנות בשקיפות שיטות חישוב של טביעת רגל פחמנית ו המסורתיות נבחנות PMIS ו-ERP בהקשר של דרישות וסטנדרטים סביבתיים מודרניים. מערכות באופן ביקורתי בנוגע ליכולות ולמגבלות שלהן בניהול תהליכי הבנייה, עם ניתוח השפעתן על שקיפות תהליך קביעת המחירים. ניבאות מגמות המעבר מפתרונות סגורים לסטנדרטים פתוחים וכלים גמישים לניתוח נתונים, המסוגלים להבטיח יעילות רבה יותר בתהליכי הבנייה.

פרק 5.1

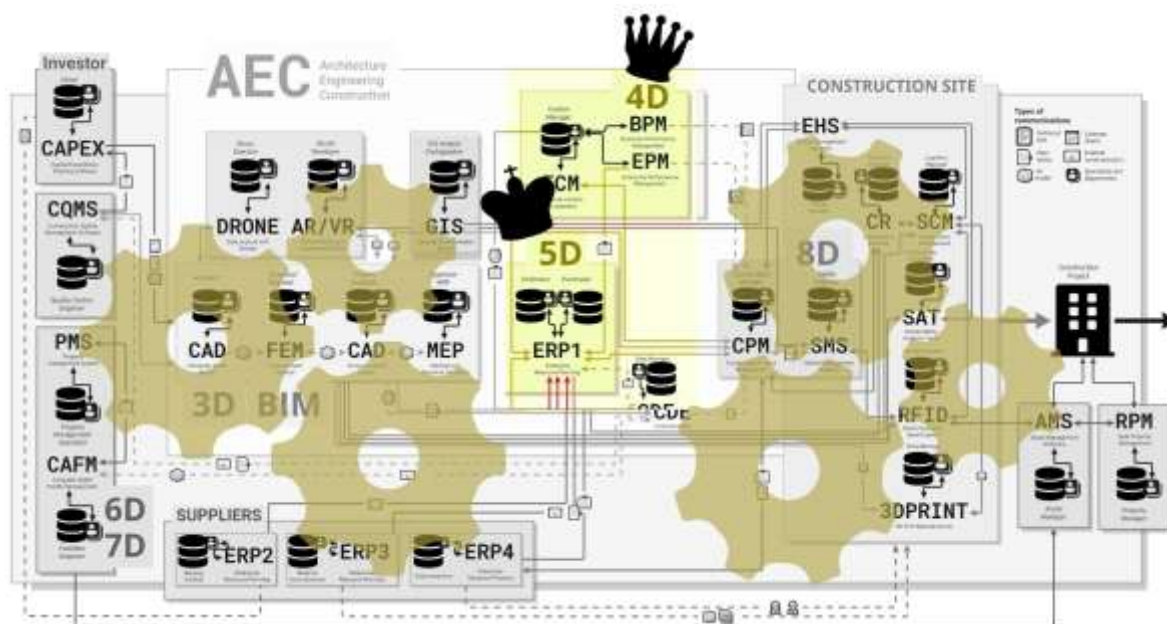
חישובי עלות והערכות פרויקטים בבנייה

יסודות הבנייה: הערכת כמות, עלות וזמן

- בין שלל תהליכי העסקים המגדירים את עמידות החברה בתעשיית הבנייה, יש חשיבות רבה - כמו לפני אלפי שנים לתהליכים של הערכת כמות מדויקת של רכיבים, עלות הפרויקט ולוחות זמנים לביצוע (איור 5.11).

התפתחות הכתיבה הייתה תוצאה של מכלול גורמים, כולל הצורך במעקב אחר פעולות כלכליות, התפתחות המסחר וניהול - משאבים בחברות מוקדמות. המסמכים המשפטיים הראשונים - לוחות חימר עם חישובי עלויות חומרים ותשלומי שכר שימשו בהקשר של מסחר ובנייה. לוחות אלו תיעדו את התחייבויות הצדדים בעת בניית מבנים ונשמרו כראיות להסכמות ולמערכות יחסים של סחורות וכסף.

במשך אלפי שנים הגישה להערכה כמעט ולא השתנתה: החישובים בוצעו ידנית, בהתבסס על ניסיון הגישה, CAD, מודולריות וכלי ERP ואינטואיציה של המהנדס המעריך. עם זאת, עם הופעת מערכות המסורתיות להערכת כמות, עלות וזמן החלה לעבור שינוי מהיר. הטכנולוגיות הדיגיטליות המודרניות מאפשרות אוטומציה מלאה של חישובים מרכזיים של זמן ועלות, מה שמעלה את הדיוק, המהירות וההשקיפות של תכנון המשאבים בפרויקטי בנייה.



איור 5.11 מתוך שלל מערכות שונות, הכלים החשובים ביותר בעסקים הם הכלים העוסקים במדדים של כמויות, עלויות וזמן.

תשומת הלב העיקרית של חברות הבנייה ממוקדת בנתונים מדויקים לגבי לוחות זמנים ועלויות העבודה. מדדים אלו בתורם, תלויים בכמויות החומרים בשימוש ובמאמצי העבודה, ושקיפותם משפיעה על הרווחיות. עם זאת, המורכבות של

תהליכי החישוב וחוסר השקיפות שלהם לעיתים קרובות מביאים לעלויות גבוהות יותר של פרויקטים, לעיכובים בלוחות הזמנים ואפילו לפשיטת רגל של חברות.

בעיות מוכרות - גישות חדשות" (רק 50% מפרויקטי הבנייה מסתיימים בזמן, ו-KPMG לפי דוח מהחברות מדווחות על עלייה בבקרת הכלכלה של פרויקטים קפיטליים. הבעיות העיקריות קשורות 87% למחסור בכוח אדם מיומן ולמורכבות חיזוי הסיכונים.

נתונים היסטוריים על חישובי עלות וזמן ביצוע תהליכים נאספים במהלך בניית פרויקטים קודמים לאורך כל חיי החברה ('וכו ERP, PMIS, BPM, EPM) הבנייתית ומוזנים לבסיסי נתונים של מערכות שונות.

קיום נתונים היסטוריים איכותיים על חישובים הוא יתרון תחרותי מרכזי של הארגון הבנייתי, המשפיע ישירות על הישרדותו.

מחלקות אומדן וחישובים בחברות בנייה והנדסה מוקמות לצורך איסוף, אחסון ועדכון נתונים היסטוריים על חישובי פרויקטים. הפונקציה העיקרית שלהן היא צבירה וסיסטמטיזציה של הניסיון של החברה, מה שמאפשר עם הזמן לשפר, את דיוק ההערכה של כמות, מועדים ועלויות של פרויקטים חדשים. גישה זו מסייעת למזער טעויות בחישובים עתידיים. בהתבסס על פרקטיקה ותוצאות של פרויקטים שכבר בוצעו.

שיטות חישוב עלות פרויקטים

בעבודת המומחים לחישובים נעשה שימוש בשיטות הערכה שונות, כל אחת מהן מכוונת לסוג נתונים ספציפי, זמינות המידע ורמת הפירוט של הפרויקט. השיטות הנפוצות ביותר כוללות:

- שיטת המשאבים: הערכת עלות הפרויקט על בסיס ניתוח מפורט של כל המשאבים הנדרשים, כגון חומרים, ציוד ועבודה. שיטה זו דורשת רשימה מפורטת של כל המשימות והמשאבים הנדרשים לביצוע כל משימה, עם חישוב עלותם לאחר מכן. שיטה זו מתאפיינת בדיוק גבוה ומיושמת באופן נרחב בהכנת אומדנים.
- השיטה הפרמטרית: עושה שימוש במודלים סטטיסטיים להערכת עלות על בסיס פרמטרים של הפרויקט. זה יכול לכלול ניתוח עלות ליחידת מידה, כגון שטח הבנייה או נפח העבודה, והתאמת עלויות אלו לתנאים הספציפיים של הפרויקט. השיטה יעילה במיוחד בשלבים המוקדמים, כאשר מידע מפורט עדיין אינו זמין.
- שיטת המדדים היחידניים) שיטת עלות ליחידה: (מחשבת את עלות הפרויקט על בסיס עלות ליחידת מידה) למשל, למטר רבוע או למטר קוב. (זה מספק השוואה נוחה ומהירה וניתוח עלויות של פרויקטים שונים או חלקיהם.
- הערכות מומחים) שיטת דלפי: (מבוססות על דעות של מומחים, המשתמשים בניסיון ובידע שלהם להערכת עלות הפרויקט. גישה זו מועילה כאשר אין נתונים מדויקים או כאשר הפרויקט ייחודי.

יש לציין בנפרד כי השיטה הפרמטרית והערכות המומחים יכולות להיות מותאמות למודלים של למידת מכונה. זה מאפשר לבנות באופן אוטומטי תחזיות על עלויות וזמנים של ביצוע הפרויקט על בסיס דוגמאות לימוד. דוגמאות לשימוש במודלים - (9.35 איור) "כאלה נדונות בפירוט בפרק "דוגמת שימוש בלמידת מכונה למציאת עלות וזמנים של פרויקט

עם זאת, השיטה הפופולרית והנמצאת בשימוש נרחב בעולם היא שיטת המשאבים. היא מספקת לא רק הערכה מדויקת של עלות האומדן, אלא גם מאפשרת לחשב את משך הזמן של תהליכים ספציפיים באתר הבנייה ושל הפרויקט כולו ("D"פירוט נוסף בפרק "גרפים של בנייה ונתוני פרויקט 4).

שיטת משאבים להכנת הערכות וחישובים בבנייה

חישוב על בסיס משאבים הוא שיטה של ניהול חשבונאות, שבה עלות הפרויקט נקבעת על סמך רישום ישיר של כל המשאבים המעורבים. בתחום הבנייה, גישה זו כוללת ניתוח מפורט והערכה של כל המשאבים החומריים, העבודה והטכנולוגיים הנדרשים לביצוע העבודות.

השיטה המשאבית מספקת שקיפות גבוהה ודייקנות בתכנון התקציב, מכיוון שהיא מתמקדת במחירים בפועל של המשאבים בזמן הכנת ההצעה. זה חשוב במיוחד בתנאים של מצב כלכלי לא יציב, כאשר תנודות מחירים יכולות להשפיע באופן משמעותי על העלות הכוללת של הפרויקט.

בפרקים הבאים נבחן בפירוט את תהליך החישוב בשיטה המשאבית. כדי להבין טוב יותר את העקרונות שלה בתחום הבנייה, נעשה אנלוגיה לחישוב עלות ארוחה במסעדה. מנהל המסעדה, המתכנן ערב, מכין רשימה של המוצרים הנדרשים לוקח בחשבון את זמן ההכנה של כל מנה, ולאחר מכן מכפיל את העלויות במספר האורחים. בתחום הבנייה, התהליך דומה: עבור כל קטגוריה של רכיבי הפרויקט (אובייקטים) (נערכות הצעות מפורטות, והעלות הכוללת של הפרויקט נקבעת על ידי סיכום כל העלויות בחשבון הסופי - הצעת המחיר הסופית לפי קטגוריות).

השלב המרכזי והראשוני בגישה המשאבית הוא יצירת בסיס נתונים ראשוני של החברה. בשלב הראשון של החישוב נערך רשימה מסודרת של כל הפריטים, החומרים, סוגי העבודות והמשאבים שברשות החברה במסגרת הפרויקטים הבנייתיים שלה - מהמסמר במחסן ועד תיאור העובדים לפי הכשרתם ושיעור השכר השעתי שלהם. מידע זה מאורגן לבסיס נתונים אחד - "בסיס נתונים משאבים וחומרים לבנייה" - רישום טבלאי המכיל נתונים על שמות, תכונות, יחידות מידה ומחירים נוכחיים. בסיס זה הופך למקור המידע העיקרי לכל החישובים הבאים של המשאבים - הן העלות והן זמני הביצוע.

בסיס נתונים של משאבי בנייה: קטלוג חומרים ועבודות בנייה

- בסיס הנתונים או טבלת המשאבים והחומרים לבנייה כוללת מידע מפורט על כל רכיב שיכול לשמש בפרויקט הבנייה מוצר, פריט, חומר או שירות, כולל שמו, תיאורו, יחידת המידה ועלות היחידה, המוקלטים בצורה מסודרת. בטבלה זו ניתן למצוא הכל: מסוגי דלק וחומרים שונים בשימוש בפרויקטים ועד רשימות מפורטות של מומחים בקטגוריות שונות עם תיאור של תשלום שעתי (איור 5.12).

Database of resources

	1st grade potatoes 1 kg \$2,99		Sand lime bricks 1 pcs \$1
	Black Angus marble beef 1 kg \$26,99		JCB 3CX backhoe loader 1 h \$150
	Broccoli 1 pcs \$1,99		Laborer of the 1st category 1 h \$30

איור 5.12 טבלת המשאבים - זו רשימת מרכיבים, המתארת את החומר והשירות עם ציון עלות היחידה.

בסיס הנתונים של המשאבים "דומה לקטלוג מוצרים של חנות אינטרנטית, שבה לכל מוצר יש תיאור" מפורט של מאפייניו. זה מקל על המעריכים לבחור את המשאבים הנדרשים (כמו בחירת מוצרים בעת הוספתם לעגלת הקניות), (הנדרשים לחישוב תהליכי הבנייה הספציפיים בצורה של חישובים) (ההזמנה הסופית בחנות האינטרנטית).

בסיס הנתונים של המשאבים ניתן גם להציג כרשימה של כל המרכיבים בספר מתכונים של מסעדה. כל חומר בניין, ציוד ושירות דומים למרכיבים בשימוש במתכונים". בסיס המשאבים "הוא רשימה מפורטת של כל המרכיבים - חומרים וישרותים לבנייה, כולל עלותם ליחידה: יחידה, מטר, שעה, ליטר וכו'.

ניתן להוסיף אלמנטים חדשים לישות בטבלת "מאגרי נתונים של משאבי בנייה" (בשני דרכים - ידנית) איור 5.13 (א) או אוטומטית, באמצעות אינטגרציה עם מערכות ניהול המלאי של החברה או מאגרי נתונים של ספקים.

Database of construction resources

Code	Description	Unit	Unit cost
dos_p	Delivery of the mainland to the site	total	500
dos_m	Delivery to the place of installation	total	300
metprofil	Metallic profile	Kg	10.30

 - Manual input

איור 5.13 מאגר הנתונים של המשאבים מתמלא ידנית או מעביר נתונים ממאגרים אחרים באופן אוטומטי.

חברה לבנייה בגודל ממוצע משתמשת במאגר נתונים המכיל אלפי, ולעיתים עשרות אלפי אלמנטים עם תיאור מפורט שניתן להשתמש בהם בפרויקטי בנייה. נתונים אלה משמשים לאחר מכן באופן אוטומטי בחוזים ובמסמכי תכנון לתיאור מדויק של הרכב העבודות וההליכים.

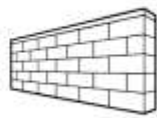
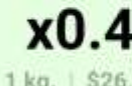
כדי לא להישאר מאחור בשינויים בתנאי השוק, כגון אינפלציה, מאפיין "עלות ליחידה" עבור כל מוצר (סחורה או שירות) (במאגר הנתונים של המשאבים) איור 5.13 (מעודכן באופן קבוע ידנית או באמצעות טעינת מחירים עדכניים ממערכות אחרות או פלטפורמות מקוונות).

עדכון עלות היחידה של המשאב יכול להתבצע מדי חודש, מדי רבעון או מדי שנה - בהתאם לאופי המשאב, לאינפלציה ולמצב הכלכלי החיצוני. עדכון זה הכרחי לשמירה על דיוק החישובים וההערכות, שכן אלמנטים בסיסיים אלה משמשים כנקודת מוצא לעבודת המעריכים. על בסיס הנתונים העדכניים נוצרים אומדנים, תקציבים ולוחות זמנים, המייצגים את התנאים האמיתיים בשוק ומפחיתים את הסיכון לטעויות בחישובים הפרויקטיים הבאים.

הכנת חישובים וחישוב עלות עבודות על בסיס בסיס המשאבים

לאחר שמילאנו את "מאגר הנתונים של משאבי הבנייה" איור 5.13 (באלמנטים-יחידות מינימליות, ניתן להתחיל ביצירת חישובים, המבוצעים עבור כל תהליך או עבודה באתר הבנייה לפי יחידות מידה מסוימות: לדוגמה, עבור מטר מעוקב של בטון, מטר רבוע של קיר גבס, מטר של שפתון או התקנת חלון אחד).

לדוגמה, עבור בניית קיר לבנים בשטח של 1 מ"ר) איור 5.14, (בהתבסס על ניסיון מפרויקטים קודמים, דרושים כ-65 לבנים) אלמנט "לבני סליקט", ("בעלות של 1 דולר ליחידה) מאפיין "עלות ליחידה", ("מה שמצטבר ל-65 דולר. בנוסף במשך 10 דקות, שימקם את הלבנים ליד ("JCB 3CX אלמנט "מעמס) על פי ניסיון, יש צורך להשתמש בציוד בנייה אזור העבודה. מכיוון שעלות השכרת הציוד היא 150 דולר לשעה, 6 דקות שימוש יעלו כ-15 דולר. בנוסף, יידרשו 2 שעות עבודה של קבלן להנחת הלבנים, בשכר שעתי של 30 דולר, והסכום הכולל יהיה 60 דולר.

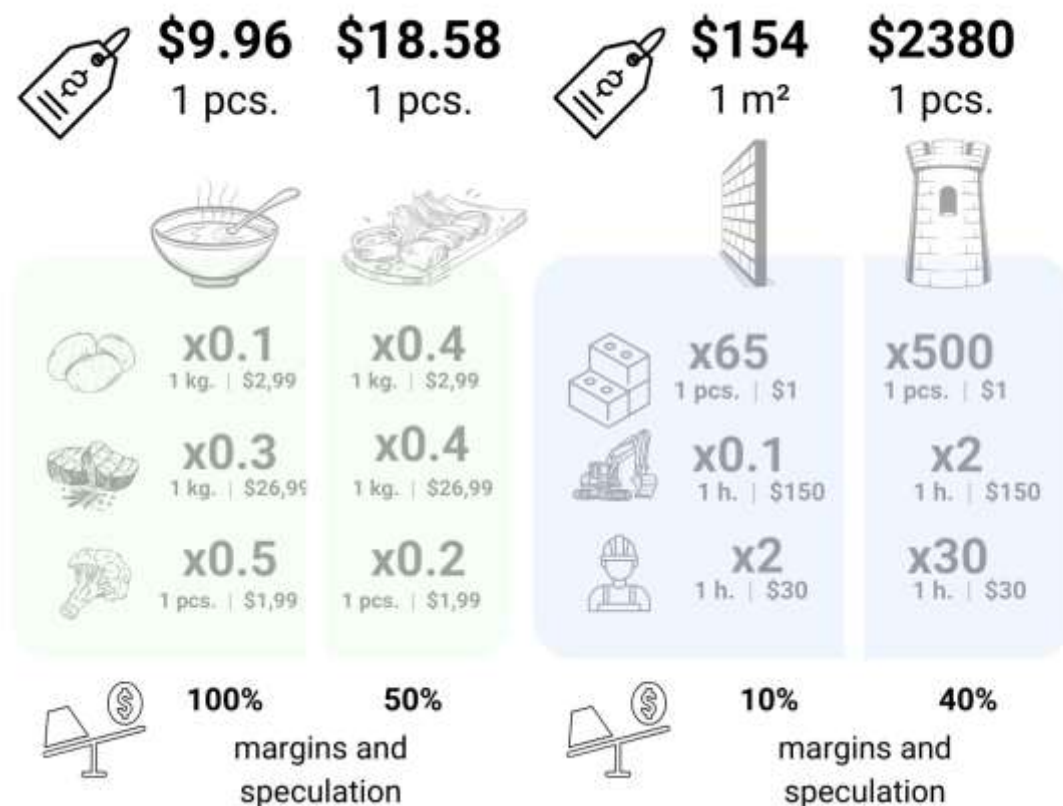
			
 x0.1 1 kg \$2,99	 x0.4 1 kg. \$2,99	 x65 1 pcs \$1	 x500 1 pcs \$1
 x0.3 1 kg \$26,99	 x0.4 1 kg. \$26,99	 x0.1 1 h \$150	 x2 1 h \$150
 x0.5 1 pcs \$1,99	 x0.2 1 pcs. \$1,99	 x2 1 h \$30	 x30 1 h \$30
 + margins and speculation	+ margins and speculation	 + margins and speculation	+ margins and speculation

איור 5.14 חישובי העלויות מכילים רשימה מפורטת של חומרי בניין ושירותים הנדרשים לביצוע העבודות וההליכים.

הרכב החישובים (המכונים "מתכונים" על בסיס הניסיון ההיסטורי שצברה החברה במהלך ביצוע כמות גדולה של עבודות דומות. ניסיון מעשי זה, בדרך כלל, מצטבר דרך משוב מהאתר הבנייה. בפרט, המנהל האחראי אוסף מידע ישירות באתר הבנייה, מתעד את שעות העבודה בפועל, את צריכת החומרים ואת הדקויות של ביצוע הפעולות הטכנולוגיות. לאחר מכן, בשיתוף פעולה עם מחלקת החישובים, מידע זה עובר תהליך של שיפור איטרטיבי: תיאור התהליכים מתעדכן, הרכב המשאבים מתוקן, והחישובים מעודכנים בהתאם לנתונים הפנימיים של הפרויקטים האחרונים.

כפי שמתארים במתכון את המרכיבים הנדרשים וכמותם להכנת מנה, כך ברשימת החישובים מופיע פירוט מדויק של כל החומרים הבנייתיים, המשאבים והשירותים הנדרשים לביצוע עבודה או תהליך ספציפי.

עבודות המבוצעות באופן קבוע מאפשרות לעובדים, למנהלי הקבוצות ולמחשבי החישובים להתמצא בכמות הנדרשת של משאבים: חומרים, דלק, שעות עבודה ופרמטרים נוספים הנדרשים לביצוע יחידת עבודה (איור 5.15). (נתונים אלה מוזנים למערכות החישובים בצורה של טבלאות, כאשר כל משימה ופעולה מתוארת דרך רכיבי בסיס המשאבים) עם מחירים שמתעדכנים באופן קבוע, (מה שמבטיח דיוק בחישובים).



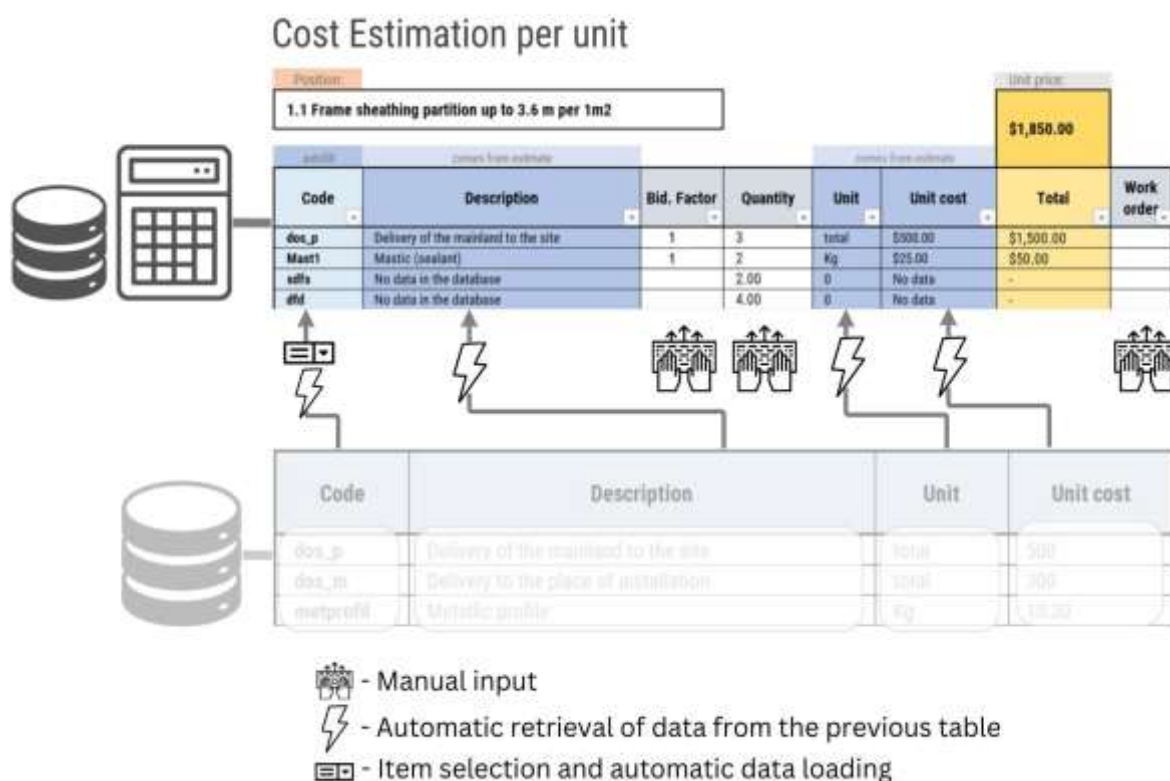
- איור 5.15 עבור כל עבודה נאספים מחירים יחידניים, כאשר מאפיין נפח הישות מוכפל בכמות שלה עם הוספת אחוז רווח

כדי לקבל את העלות הכוללת של כל תהליך או עבודה (אובייקט החישוב, (מאפיין העלות מוכפל בכמות שלו ובמקדם המקדם יכול לכלול גורמים שונים, כגון מורכבות העבודה, מאפיינים אזוריים, רמת אינפלציה, סיכונים פוטנציאליים (אחוז צפוי של הוצאות כלליות (או ספקולציות) מקדם רווח נוסף)

מחשב החישובים, כאנליסט, ממיר את הניסיון וההמלצות של המנהל האחראי לחישובים סטנדרטיים מתאר את התהליכים הבנייתיים דרך ישויות משאבים בצורה טבלאית. למעשה, משימת מחשב החישובים היא לאסוף ולארגן את המידע המגיע מהאתר הבנייה דרך פרמטרים ומקדים

כך, העלות הסופית של יחידת עבודה (למשל, מטר רבוע או מטר קוב, או התקנה אחת (כוללת לא רק את ההוצאות הישירות על חומרים וכוח אדם, אלא גם את הרווחים של החברה, הוצאות כלליות, ביטוח וגורמים נוספים) איור 5.16)

בשלב זה, אין צורך לדאוג לגבי עדכניות המחירים בחישובים, מכיוון שהמחירים האמיתיים תמיד משתקפים ב"מאגר המשאבים" (טבלת המרכיבים). (בשלב החישובים, הנתונים נטענים אוטומטית) (למשל, לפי קוד רכיב או מזהה ייחודי שלו (מהמאגר, שמספק תיאור ועלות עדכנית ליחידה, שיכולה להיות גם נטענת אוטומטית מפלטפורמות מקוונות או חנויות אינטרנטיות לחומרים בנייתיים. למחשב החישובים בשלב החישובים נותר רק לתאר את העבודה או התהליך דרך מאפיין "כמות המשאבים" וגורמים נוספים



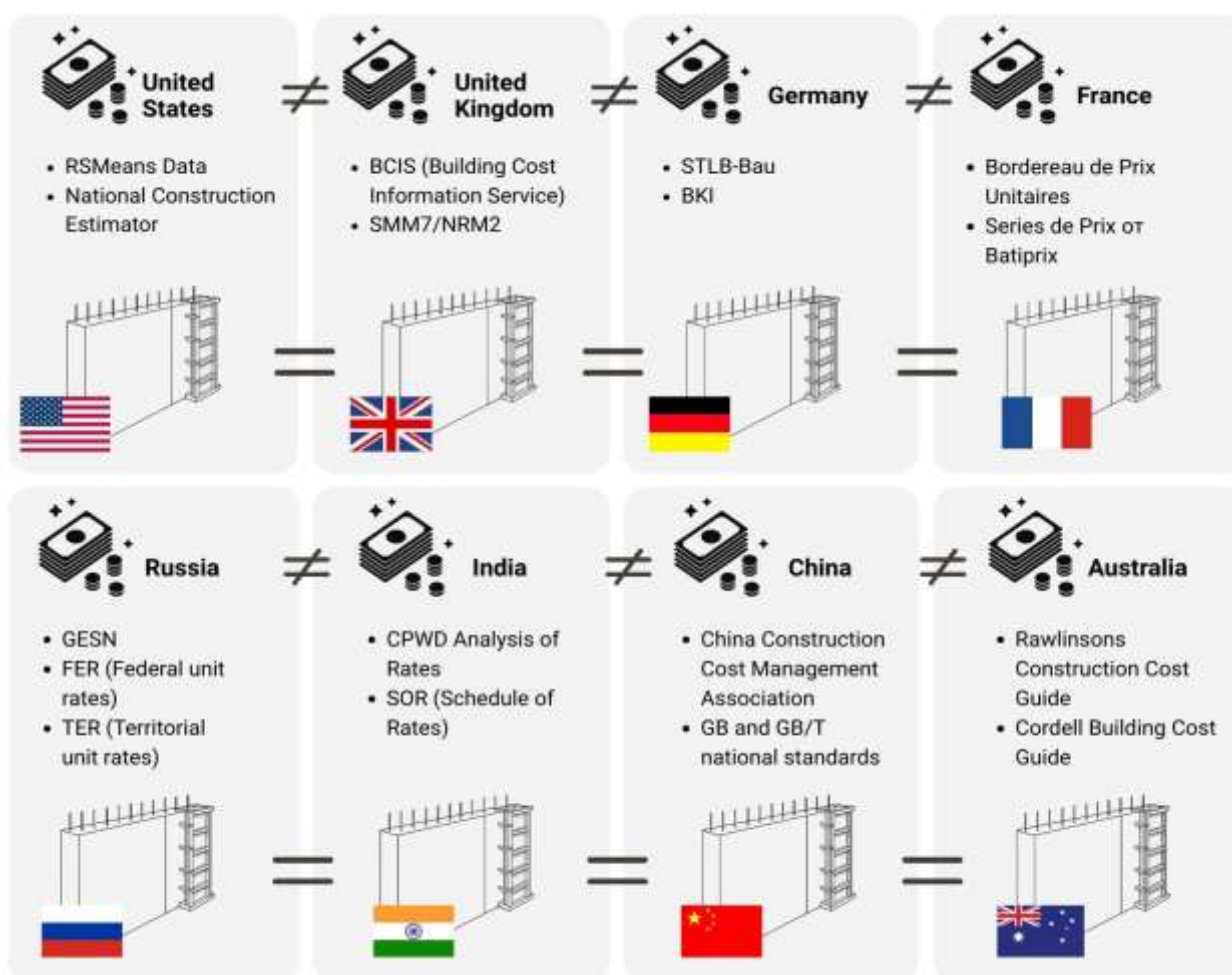
איור 5.16 בשלב חישוב עלות יחידת העבודה, ממלאים רק את מאפייני כמות המשאבים הנדרשים, וכל השאר נטען אוטומטית מהמאגר.

חישובי עלות העבודה שנוצרו נשמרים בטבלאות תבנית של פרויקטים טיפוסיים, הקשורים ישירות למסד הנתונים של משאבים וחומרים בנייתיים. תבניות אלו מהוות מתודולוגיות סטנדרטיות לביצוע סוגי עבודות חוזרות, עבור פרויקטים עתידיים, ומבטיחות אחידות בחישובים ברמה של כל החברה.

כאשר משתנה מחיר של כל משאב במסד הנתונים (איור 5.13) (- בין אם ידנית או אוטומטית דרך טעינת מחירים שוק עדכניים) למשל, בתנאי אינפלציה (-) העדכונים משתקפים מיד בכל החישובים הקשורים (איור 5.16). (זה אומר שדי לבצע שינויים רק במסד הנתונים של המשאבים, בעוד שתבניות החישוב וההערכות נשארות ללא שינוי לאורך זמן. גישה זו מבטיחה יציבות וחזרתיות בחישובים בכל תנודות המחירים, אשר נלקחות בחשבון רק בטבלת המשאבים הפשוטה -- יחסית) (איור 5.13).

עבור כל פרויקט חדש נוצר עותק של תבנית החישוב הסטנדרטית, מה שמאפשר לבצע שינויים ולתאם פעילויות בהתאם לדרישות מיוחדות מבלי לשנות את התבנית המקורית, שהתקבלה בחברה. גישה זו מספקת גמישות בהתאמת החישובים: ניתן לקחת בחשבון את מאפייני אתר הבנייה, את רצונות הלקוח, להכניס מקדמים של סיכון או רווחיות (ספקולציות) (- וכל זאת מבלי להפר את הסטנדרטים של החברה. זה עוזר לחברה למצוא את האיזון בין מקסום הרווחים, דרישות הלקוחות ושמירה על התחרותיות שלה.

במדינות מסוימות תבניות חישוב אלו, שנצברו במשך עשורים, מתועדות ברמה הלאומית והופכות לחלק מהסטנדרטים הממשלתיים של מערכת חישוב עלות עבודות הבנייה (איור 5.17).



איור 5.17 במדינות שונות בעולם לחישובים של אותו רכיב קיימות כללים משלהן עם) מתודולוגיות (אוספים ותקנים לעבודות בנייה.

בסיסי המשאבים הסטנדרטיים של ההערכות) איור 5.17 (מחויבים לשימוש על ידי כל המשתתפים בשוק, בראש ובראשונה, בעת ביצוע פרויקטים עם מימון ממשלתי. תהליך הסטנדרטיזציה הזה מבטיח ללקוח שקיפות, השוואתיות והוגנות בהקניית מחירים והתחייבויות חוזיות.

חישוב סופי של עלות הפרויקט: מהערכות לתקציב

התקנים הממשלתיים והענפיים של ההערכות ממלאים תפקידים שונים בפרקטיקת הבנייה במדינות שונות. בעוד שכמה מדינות מחייבות הקפדה על תקנים אחידים, רוב הכלכלות המפותחות מאמצות גישה גמישה יותר. במדינות עם כלכלה שוק, התקנים הממשלתיים לבנייה בדרך כלל משמשים רק כנקודת מוצא בסיסית. חברות הבנייה מתאימות את הסטנדרטים הללו למודלים התפעוליים שלהן או מעבדות אותם לחלוטין, ומוסיפות להם מקדמים משלהן, המתחשבים במאפייני הפעילות שלהן. התאמות אלו משקפות את הניסיון התאגידי, את היעילות בניהול המשאבים ולעיתים קרובות גם גורמים, שבהם ניתן לקחת בחשבון את הרווח הספקולטיבי של החברה.

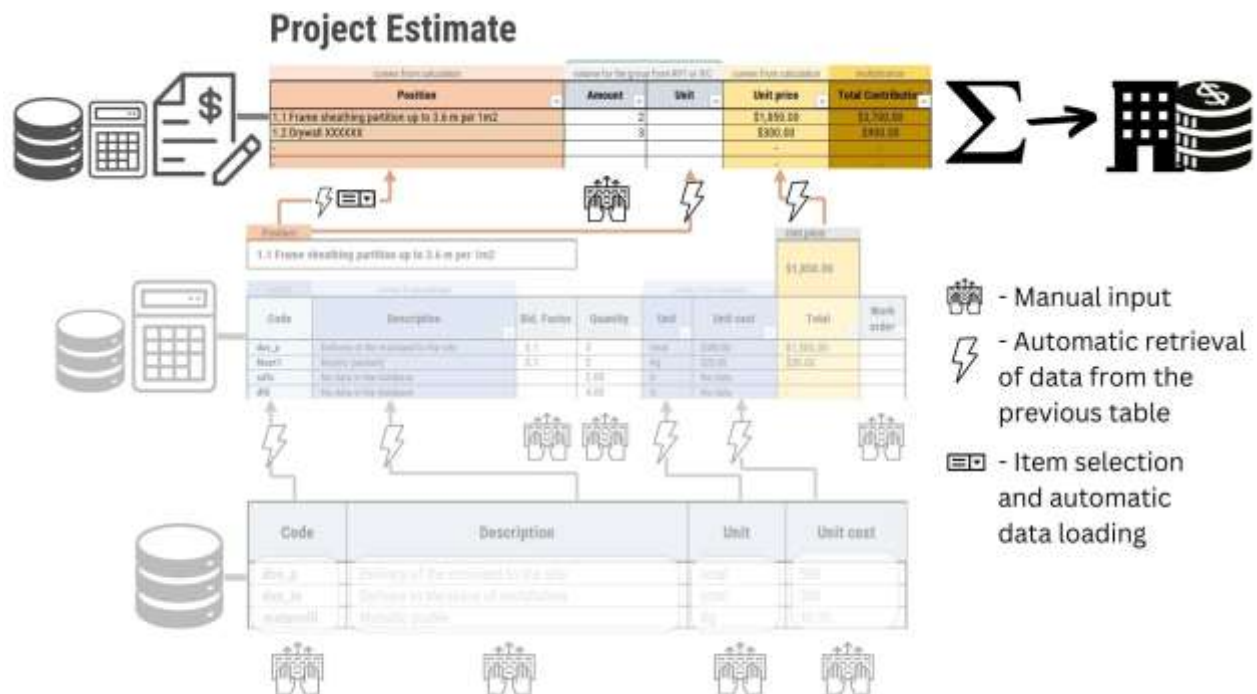
בסופו של דבר, רמת התחרות, הביקוש בשוק, שיעור הרווח המיועד ואפילו הקשרים עם לקוחות ספציפיים יכולים להוביל לסטיות משמעותיות מהנורמות הסטנדרטיות. פרקטיקה זו מספקת גמישות בשוק, אך בו בזמן מקשה על השוואה שקופה של הצעות קבלנים שונים, תוך הכנסת אלמנט של תמחור ספקולטיבי לתחום הבנייה בשלב זה של החישובים.

לאחר שכבר הוכנו תבניות חישוב עבור סוגי עבודות ותהליכים שונים - או, מה שקורה לעיתים קרובות יותר, פשוט עם מקדמים המייצגים את "המאפיינים" של החברה (5.17 איור) הועתקו מתוך אומדנים סטנדרטיים ממשלתיים הספציפית - בשלב הסופי נותר רק להכפיל את עלות כל פריט באטריבוט המתאים של היקף העבודות או התהליכים בפרויקט החדש.

בחישוב העלות הכוללת של פרויקט בנייה חדש, שלב מפתח הוא סיכום ההוצאות בכל סעיפי החישובים, מוכפל בהיקף של פריטים-עבודות בפרויקט.

כדי ליצור את העלות הכוללת של הפרויקט, בדוגמה הפשוטה שלנו, נתחיל בחישוב עלות הבנייה של מטר רבוע אחד של קיר ונכפיל את עלות החישוב שלו (למשל, עבודה "1מ² התקנה סטנדרטית של אלמנטים קיר" (במספר הכולל של מטרים של ישות מסוג "אלמנטים קיר" מתוך (5.18 איור) "רבועים של קירות בפרויקט) (למשל, אטריבוט "שטח" או "כמות (או חישובי מנהל העבודה CAD פרויקט).

לוקחים את עלות יחידת העבודה: (5.18 איור) באופן דומה, אנו מחשבים את העלות עבור כל האלמנטים בפרויקט, ומכפילים אותה בהיקף של אלמנט מסוים או קבוצתו בפרויקט זה. למעריך עלויות נותר רק להזין את כמות האלמנטים העבודות או התהליכים בפרויקט בצורה של היקף או כמות. זה מאפשר ליצור אוטומטית את אומדן הבנייה המלא.



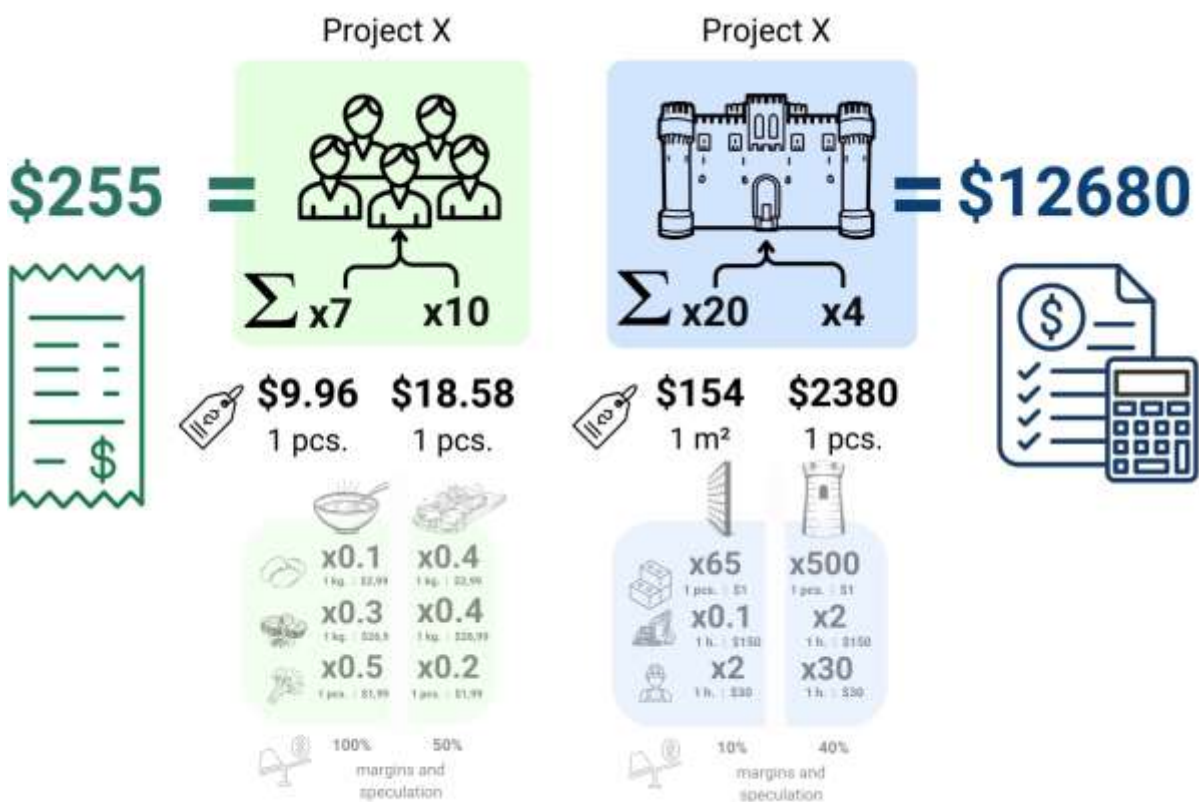
בשלב יצירת האומדן או מזינים רק את היקף העבודות 5.18 אזור.

כמו במקרה של חישובים, בשלב זה אנו טוענים אוטומטית פריטים מחושבים מוכנים) מתבנית החישובים או חדשים שהועתקו מתבנית וערכו מחדש, (אשר מביאים אוטומטית את העלות העדכנית ליחידת עבודה) שמתעדכנת אוטומטית בהתאם, בכל שינוי בנתונים במאגר המשאבים או בטבלאות. (טבלה תחתונה 5.18 אזור) ממאגר הנתונים של המשאבים החישובים - הנתונים באומדן יתעדכנו אוטומטית ליום הנוכחי, ללא צורך לשנות את החישובים או את האומדן עצמו.

בהקשר של מסעדה, העלות הסופית של האירוע מחושבת באופן דומה והיא שווה לעלות הסופית של כל הארוחה, כאשר וכמו בבנייה, מתכוני הכנת. (5.19 אזור) עלות כל מנה, מוכפלת במספר האורחים, מצטברת לעלות הכוללת של החשבון המנות במסעדה עשויים לא להשתנות במשך עשורים. בניגוד למחירים, שבהם עלות המרכיבים עשויה להשתנות בכל שעה.

כפי שבעל מסעדה מכפיל את עלות כל מנה במספר המנות ואנשים, כדי לקבוע את העלות הכוללת של האירוע, כך גם מנהל האומדנים מסכם את עלות כל רכיבי הפרויקט, כדי לקבל את האומדן המלא של הבנייה.

כך, עבור כל עבודה בפרויקט נקבעת עלותה הסופית) אזור 5.19, (אשר, כאשר היא מוכפלת בנפח האטריבוטיבי של הישות המתאימה לעבודה זו - נותנת את עלות קבוצת העבודות, ממנה מתקבלת העלות הסופית של כל הפרויקט.



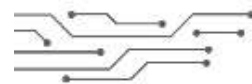
איור 5.19 התקציב הסופי מחושב על ידי סיכום האטריבוט של עלות העבודות של כל רכיב על האטריבוט של נפח.

העלות הסופית של הפרויקט (איור 5.18) (מהווה תמונה פיננסית של הפרויקט, המאפשרת ללקוחות, משקיעים או ארגונים פיננסיים להבין את התקציב הכולל ואת המשאבים הפיננסיים הנדרשים לביצוע הפרויקט, בכל יום, בהתחשב במחירים העדכניים).

ואם תהליכי הרכבת בסיסי המשאבים, חישובים ותקציבים) מתכונים לתהליכים (כבר פותחו, חצי-אוטומטיים ומדויקים במשך עשרות אלפי שנים ונרשמו ברמה הממשלתית, אז קבלת מידע איכותי אוטומטית על הנפח וכמות הרכיבים לשלב האחרון של התקציב הסופי - היום נשאר צוואר בקבוק בתהליכי כל החישובים של האטריבוטים העלותיים והזמניים של הפרויקט, וגם באופן כללי של התקציב הכולל של הפרויקט.

במשך אלפי שנים, השיטה המסורתית לחישוב הנפחים הייתה שיטות מדידה ידניות של מאפיינים נפחיים וכמותיים באמצעות שרטוטים שטוחים. עם הגעת העידן הדיגיטלי, חברות גילו כי ניתן להפיק אוטומטית מה שהביא למהפכה בדרכים, CAD מידע על נפחים וכמויות מנתונים גיאומטריים הכלולים במודלים אלפי שנים של קבלת נתונים כמותיים.

גישות מודרניות להערכת תהליכים ועבודות תקציביות הפקת אוטומטית של אטריבוטים נפחיים וכמותיים מבסיסי שניתן לייצא ולחבר לתהליך החישובים, כדי לקבל נפחים עדכניים של קבוצות פרויקטים בכל שלב של CAD, נתונים התכנון ועד לניצול.



פרק 5.2. חישוב כמות ואוטומציה של הכנת הערכות ותוכניות לוח זמנים

שימוש בפרמטרים נפחיים וכמותיים: D-5 ל-D-4 המעבר מ-3

עם טבלאות החישובים המיועדות עם התהליכים המתוארים דרך המשאבים (איור 5.18), השלב הבא הוא קבלת אוטומטית של פרמטרי הנפח או הכמות עבור קבוצת רכיבים, הנדרשים לחישובים ולעריכת התקציב הסופי.

המאפיינים הנפחיים של רכיבי הפרויקט - לדוגמה, קירות או תקרות - יכולים להיות מופקים אוטומטית מבסיס נתונים מומרות באמצעות הליבה הגיאומטרית לערכים מספריים של CAD אובייקטים פרמטריים שנוצרו בתוכניות CAD. פרמטרים כמו אורך, רוחב, שטח, נפח ועוד. תהליך קבלת הנפחים על בסיס גיאומטריה תלת-ממדית יידון בהרחבה בחלק בנוסף לנפחים, כמות של רכיבים דומים יכולה גם להיות. CAD (BIM) השישי הבא (איור 6.33), (המוקדש לעבודה עם על ידי סינון וקיבוץ אובייקטים לפי קטגוריות ותכונות. פרמטרים אלו CAD מופקת מבסיס הנתונים של מודל המאפשרים קיבוץ - הופכים לבסיס לקישור רכיבי הפרויקט דרך חישובי המשאבים עם החישובים, התקציב הסופי והתקציב הכולל של הפרויקט.

D-5 מתמלא בשכבות חדשות של פרמטרים, המוגדרות כ-4, (CAD) כך, המודל הנתוני, המופק ממודל תלת-ממדי משמשים כמקור D) עלות (-) הנתונים הגיאומטריים של D-3) זמן (D-5) בשכבות החדשות של אטריבוטים של ישויות - 4. לערכי האטריבוטים של נפחי הישויות.

- שכבת מידע של פרמטרים, המוסיפה לפרמטרים תלת-ממדיים של רכיבים מידע על משך ביצוע פעולות - 4D בנייה. נתונים אלה חיוניים לתכנון לוחות זמנים ולניהול מועדי ביצוע הפרויקט
- השלב הבא בהרחבת מודל הנתונים, שבו רכיבים מתווספים לתכונות עלות. כך, למידע הגיאומטרי - 5D מתווסף היבט פיננסי: עלות חומרים, עבודות וציוד, מה שמאפשר לבצע חישובי תקציב, לנתח רווחיות ולנהל הוצאות במהלך הבנייה.

ERP של קבוצות ישויות בפרויקט מתוארים בדומה לחישובים במערכות D-5 נתוני עלות ופרמטרים תלת-ממדיים, 4 ומשמשים לחישוב אוטומטי של עלות ותכנון תקציב עבור (Excel-או כלים דומים ל) PIMS מודולריות, מערכות קבוצות נפרדות כמו גם עבור התקציב הכולל של הפרויקט

CAD-והפקת נפחים של תכונות מ D תכונות 5

בעת הכנת אומדן סופי של פרויקט הבנייה, שהכנתו נדונה בפרקים הקודמים (איור 5.18), (תכונות הנפח עבור כל CAD. קטגוריית רכיבים בפרויקט נאספות ידנית או מופקות ממפרטי תכונות הנפח, המוצעים על ידי תוכנות

השיטה המסורתית לחישוב נפחים ידנית מחייבת את המנהל והאומדן לנתח שרטוטים, שהוצגו במשך אלפי או PDF (PLT) שנים בצורה של קווים על נייר, וב-30 השנים האחרונות - בפורמטים דיגיטליים כמו בהתבסס על ניסיון מקצועי, הם מודדים את נפחי העבודות והחומרים הנדרשים, לעיתים קרובות DWG. בעזרת סרגל ומעגל. שיטה זו דורשת מאמצים וזמן משמעותיים, כמו גם תשומת לב מיוחדת לפרטים.

הגדרת תכונות נפח העבודות בדרך זו יכולה לקחת מספר ימים עד מספר חודשים, בהתאם לגודל הפרויקט. בנוסף, מכיוון

שכל המדידות והחישובים מתבצעים ידנית, קיים סיכון לשגיאות אנושיות, שעלולות להוביל לנתונים לא מדויקים, מה שישפיע לאחר מכן על שגיאות בהערכה של זמן ועלות הפרויקט, שעליהן תישא החברה כולה באחריות.

מפשטות באופן משמעותי את חישוב הנפחים. במודלים, CAD שיטות מודרניות, המבוססות על שימוש בבסיסי נתונים של הגיאומטריה של רכיבים כבר כוללת תכונות נפח, שניתן לחשב אוטומטית (דרך הליבה הגיאומטרית) איור, CAD של - ולהציג או לייצא בצורה טבלאית (6.33).

נתונים על תכונות כמותיות ונפחיות של רכיבי הפרויקט. נתונים CAD-בתרחיש כזה, מחלקת האומדן מבקשת מהמעצב ב או במערכות ERP במערכות, Excel-אלה מיוצאים בצורה של טבלאות או משולבים ישירות בבסיסי החישוב – בין אם ב תהליך זה לעיתים קרובות מתחיל לא מבקשה פורמלית, אלא משיחה קצרה בין המזמין (לדאריכל ולמנהל. PMIS האומדן מצד החברה הבנייתית או המתכננת. להלן דוגמה פשוטה, המדגימה כיצד מתקבלת טבלה מובנית לחישובים מתוך תקשורת יומיומית (QTO) אוטומטיים:

- המזמין – רוצה להוסיף עוד קומה לבניין, באותה קונפיגורציה כמו הקומה השנייה
- מוסיפים קומה שלישית, הקונפיגורציה זהה לזו שבקומה השנייה. ולאחר ההודעה הזו – (CAD) האדריכל למנהל האומדן CAD-שולח גרסה חדשה של פרויקט ה
- אעביר את הפרויקט דרך טבלת" – (ERP, PMIS, Excel) המעריך מבצע באופן אוטומטי קיבוץ וחישוב אקבל את הנפחים לפי קטגוריות עבור הקומה החדשה ואכין את, QTO (ERP, PMIS) עם כללי Excel "ההצעה".

בסופו של דבר, הדיאלוג הטקסטואלי מתממשק למבנה טבלה עם כללי קיבוץ

קומה	קטגוריה	רכיב
3	OST_Floors	תקרה
3	OST_StructuralColumns	עמוד
3	OST_Stairs	מדרגות

של המעריך וכפל הנפחים בחישובי QTO מהמעצב לפי כללי CAD-לאחר תהליך הקיבוץ האוטומטי של מודל ה -אנו מקבלים את התוצאות הבאות שנשלחות ללקוח, (5.18 איור) משאבים

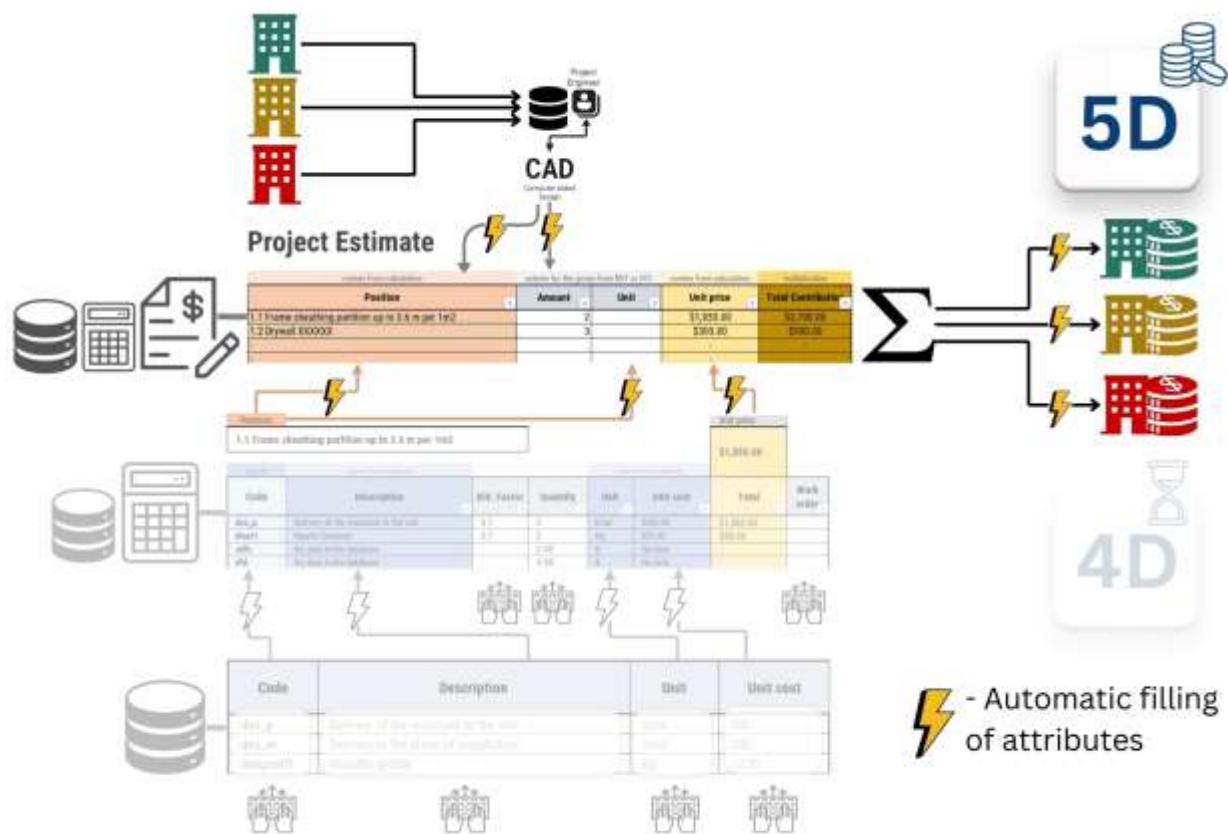
עלות סופית	מחיר ליחידה	קומה	נפח	רכיב
63 000 €	150 €/מ ²	3	מ ² 420	תקרה
9 800 €	2450 €/יח'	3	יח' 4	עמוד
8 600 €	4300 €/יח'	3	יח' 2	מדרגות
81 400 €	—	—	—	סה"כ:

הלקוח – "תודה, זה הרבה, צריך לקצץ כמה חדרים." והמחזור חוזר על עצמו פעמים רבות.

תרחיש כזה יכול לחזור על עצמו פעמים רבות, במיוחד בשלב ההסכמות, שבו הלקוח מצפה לתגובה מיידית. עם זאת, בפועל תהליכים כאלה יכולים להימשך ימים או אפילו שבועות. היום, בזכות יישום כללי קיבוץ וחישובים אוטומטיים, פעולות שלפני כן לקחו זמן רב, צריכות להסתיים בתוך דקות ספורות. קבלת הנפחים באופן אוטומטי, דרך כללי קיבוץ, לא רק מזרזת את החישובים והכנת ההצעות, אלא גם מפחיתה את הסיכוי לטעויות על ידי צמצום הגורם האנושי, ומספקת הערכת עלות שקופה ומדויקת של הפרויקט.

נלקחו בחשבון דרישות מחלקת ההערכה (שמהן בפועל CAD אם בעת יצירת מודל תלת-ממדי במערכת זה עדיין קורה לעיתים רחוקות, (ושמות, מזהים של קבוצות רכיבים ותכונות סיווג שלהם נקבעו כפרמטרים, התואמים למבנים של קבוצות וסוגי ההערכה, אז ניתן להעביר את התכונות הנפחיות באופן אוטומטי למערכות ההערכה ללא המרות נוספות.

בצורה של טבלאות-מפרט מאפשר לקבל במהירות נתונים עדכניים על עלות CAD-חילוץ אוטומטי של תכונות נפחיות מ עם הנפחים הפרויקטיים במהלך CAD-על ידי עדכון רק קובץ ה. (5.21 איור) עבודות ספציפיות ועלות הפרויקט כולו החישוב או במערכת החישוב, החברה יכולה לחשב במהירות את ההצעה בהתחשב בשינויים האחרונים, מה שמבטיח דיוק גבוה ועקביות בכל החישובים הבאים.



מוזנונת אוטומטית להצעה, מה שמאפשר לחשב מיד את העלות CAD תכונות הנפחים מטבלאות או ממסדי נתונים של 5.21 איור הכוללת של הפרויקט.

בתנאים של מורכבות גוברת של פרויקטים קפיטליים, חישוב התקציב המלא וניתוח העלות הכוללת של פרויקטים -הופך לכלי מפתח לקבלת החלטות מבוססות - (5.21 איור) בתרחיש כזה

יצירת ערך רב יותר באמצעות פרויקטים הוניים) "2024] (20, [חברות" Accenture בהתאם למחקר של מובילות משולבות באופן פעיל ניתוח נתונים ביוזמות דיגיטליות, תוך שימוש במידע היסטורי כדי לחזות ולייעל תוצאות. המחקר מראה שיותר ויותר בעלי-מפעילים משתמשים בניתוח נתונים גדולים כדי לחזות מגמות שוק ולהעריך את הכדאיות המסחרית עוד לפני תחילת התכנון. זה מושג באמצעות ניתוח מאגרי נתונים מתוך תיק הפרויקטים הקיים. בנוסף, 79% מבעלי-המפעילים מיישמים ניתוח חיזוי "מהימן" כדי להעריך את היעילות של פרויקטים ולתמוך בקבלת החלטות תפעוליות בזמן אמת

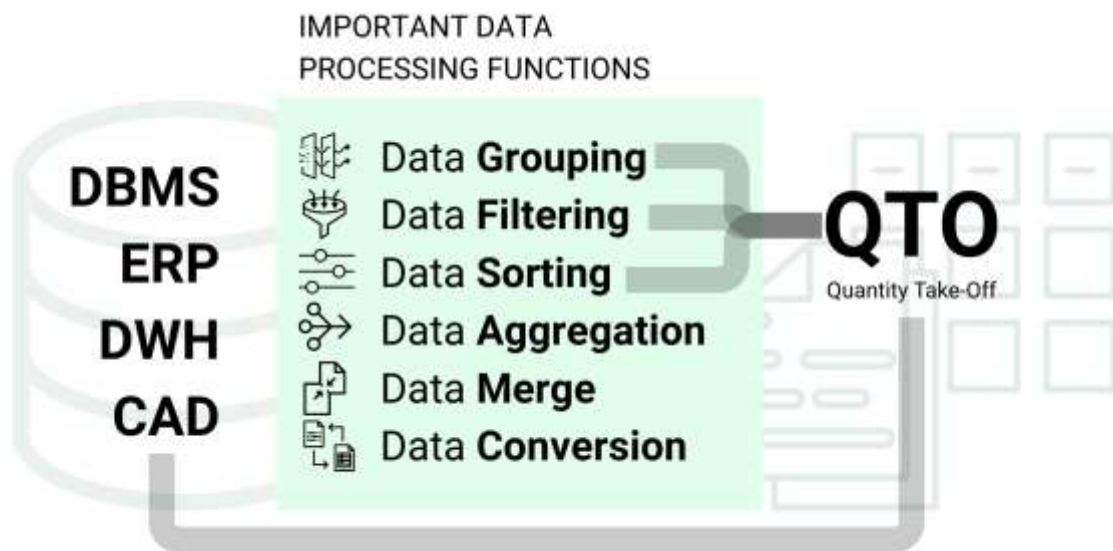
ניהול פרויקטים בנייה יעיל מודרני קשור באופן בלתי נפרד לעיבוד ולניתוח כמויות גדולות של מידע בכל שלב של התכנון ובתהליכים המקדימים את התכנון. השימוש במאגרי נתונים, חישובי משאבים, מודלים חיזוי ולמידת מכונה מאפשר לא רק למזער סיכונים בחישובים, אלא גם לקבל החלטות אסטרטגיות לגבי מימון הפרויקט בשלב מוקדם של התכנון. על מאגרי הנתונים ומודלי החיזוי, שימשיכו להשלים את החישובים, נדבר בפירוט בחלק התשיעי של הספר

הנדרשים להכנת אומדנים, מתבצעת באמצעות כלים, CAD קבלת פרמטרים כמותיים אוטומטית של רכיבים מפרויקטי כלים אלו פועלים על ידי קיבוץ כל האובייקטים של הפרויקט לפי מזהי רכיבים. QTO (Quantity Take-Off) לקיבוץ CAD מיוחדים או פרמטרים של תכונות רכיבים, תוך שימוש בפרמטרים ובטבלאות שנוצרו במאגר הנתונים של

קיבוץ נתוני פרויקט לפי תכונות (QTO): חישוב כמות

בבנייה מהווה תהליך של חילוף מאפיינים (QTO - Quantity Take-Off) חישוב פרמטרים כמותיים וכמות חומרים לעיתים קרובות נשאר תהליך חצי-ידני, הכולל איסוף QTO, כמותיים של רכיבים, הנדרשים לביצוע הפרויקט. בפועל CAD, ודגמים דיגיטליים של DWG שרטוטים בפורמט PDF, נתונים ממקורות שונים: מסמכי

מתבצע כסדרה של פעולות סינון (QTO) תהליך ההערכה הכמותית, CAD, בעבודה עם נתונים שהופקו ממאגרי נתוני מיון, קיבוץ ואגירה. רכיבי המודל נבחרים לפי פרמטרים של קטגוריות, סוגים ומעמדות, ולאחר מכן מאפייניהם הכמותיים - (5.22 איור) כמו נפח, שטח, אורך או כמות - מסוכמים בהתאם ללוגיקת החישוב -



קיבוץ וסינון נתונים - אלו הפונקציות הפופולריות ביותר המיועדות למאגרי נתונים ולמאגרי נתונים 5.22 איור.

(סינון וקיבוץ (מאפשר לסדר את הנתונים, ליצור מפרטים ולהכין מידע ראשוני לחישוב אומדנים, רכישות QTO תהליך הוא סיווג רכיבים לפי סוג המאפיינים הכמותיים הנמדדים. עבור כל QTO ולוחות זמנים לביצוע עבודות. הבסיס של רכיב או קבוצת רכיבים נבחר פרמטר מתאים למדידה כמותית. לדוגמה

- מאפיין אורך (אבן שפה - במטרים)
- מאפיין שטח (עבודות גבס - במטרים רבועים)
- מאפיין נפח (עבודות בטון - במטרים מעוקבים)
- מאפיין כמות (חלונות - לפי יחידות)

לעיתים קרובות נעשה, QTO בנוסף למאפיינים הכמותיים המתקבלים מתמטית על בסיס הגיאומטריה, לאחר קיבוץ שימוש בקואפיציוני עורך (ראה איור 5.212 לדוגמה, 1.1 כדי לקחת בחשבון 10% בלוגיסטיקה ובהתקנה (- ערכים מתוקנים, המתחשבים בהפסדים, מאפייני ההתקנה, האחסון או ההובלה. זה מאפשר לחזות בצורה מדויקת יותר את הצריכה בפועל של חומרים ולהימנע גם מחוסר וגם מעודף במלאי באתר הבנייה.

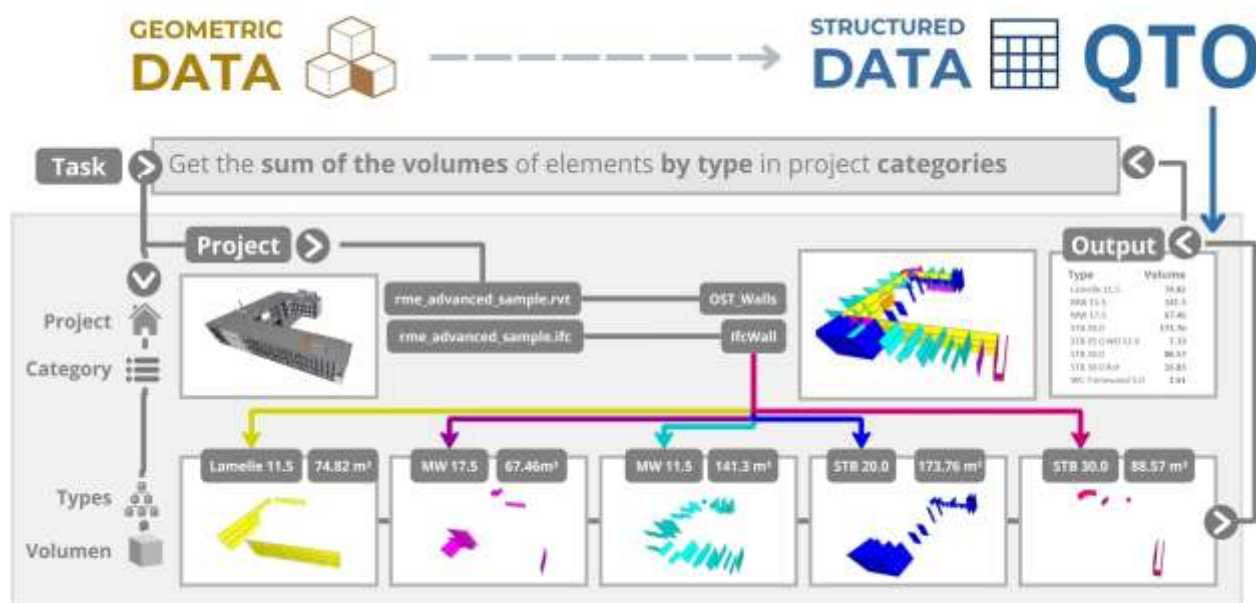
הכרחי להכנת חישובים מדויקים והצעות מחיר, מפחית את (QTO) תהליך הכימות הכמותי האוטומטי הגורם האנושי בתהליכים של מציאת מאפיינים כמותיים ומונע עודף או חוסר בהזמנת חומרים

טבלת מפרט של כמויות לפי סוגי CAD נבחן מקרה נפוץ שבו יש להציג מתוך מסד הנתונים QTO, כדוגמה לתהליך רכיבים עבור קטגוריה מסוימת, סוגי רכיבים. נקבץ את כל רכיבי הפרויקט לפי סוגים מתוך קטגוריית הקירות של פרויקט (ראה איור 5.23) QTO ונחבר את מאפייני הכמות עבור כל סוג, כדי להציג את התוצאה בטבלת כמויות CAD.

מקובצים לפי CAD טיפוס (ראה איור 5.23) (כל רכיבי קטגוריית הקירות בתוך מסד הנתונים CAD בדוגמת פרויקט ויש להם מאפייני כמות ברורים המוצגים, "STB 20.0"-ו "MW 11.5", "Lamelle 11.5", סוגי קירות, לדוגמה במטרים מעוקבים.

המטרה של המנהל, הנמצא בין המעצבים למומחי החישובים, היא לקבל טבלת כמויות אוטומטית לפי סוגי רכיבים בקטגוריה הנבחרת. ולא רק עבור פרויקט ספציפי, אלא גם בצורה אוניברסלית, שניתן ליישם על פרויקטים אחרים עם מבנה מודל דומה. זה מאפשר להרחיב את הגישה ולספק שימוש חוזר בנתונים מבלי לשכפל מאמצים

חלפו הימים שבהם מנהלי עבודה ומעריכים היו מצוידים בסרגל, מודדים בקפדנות כל קו על גבי נייר או מסורת שלא השתנתה במשך אלפי שנים. עם התפתחות המידול התלת-ממדי, שבו - PDF תכניות הגיאומטריה של כל רכיב קשורה ישירות למאפייני הכמות המחושבים אוטומטית, תהליך קביעת הכמויות הפך לאוטומטי QTO והכמות



מתוך הפרויקט QTO איור 5.23 קבלת מאפייני הכמות והכמות

בדוגמה שלנו, המשימה היא "לבחור את קטגוריית הקירות בפרויקט, לקבץ את כל הרכיבים לפי סוג ולהציג את המידע, על מאפייני הכמות בפורמט מסודר, טבלאי, כך שטבלה זו תוכל לשמש עשרות מומחים אחרים לחישובי הצעות מחיר

לוגיסטיקה, לוחות זמנים ועוד מקרים עסקיים) ראה איור 6.13).

על הסיבות) CAD לא כל מומחה יכול כיום להשתמש בגישה ישירה למסד הנתונים, CAD בשל האופי הסגור של נתוני המבוססים על, BIM ופתרונות לבעיה זו מפורט בפרק השישי של הספר. (לכן רבים נאלצים לפנות לכלים מיוחדים של ניתן, CAD או ישירות בסביבת תוכנת BIM בעבודה עם כלים מיוחדים של [63] open BIM ו-closed BIM מושגי בדרכים שונות - בהתאם אם נעשה שימוש בממשק ידני QTO (Quantity Take-Off) ליצור את הטבלה עם תוצאות או באוטומציה תוכניתית

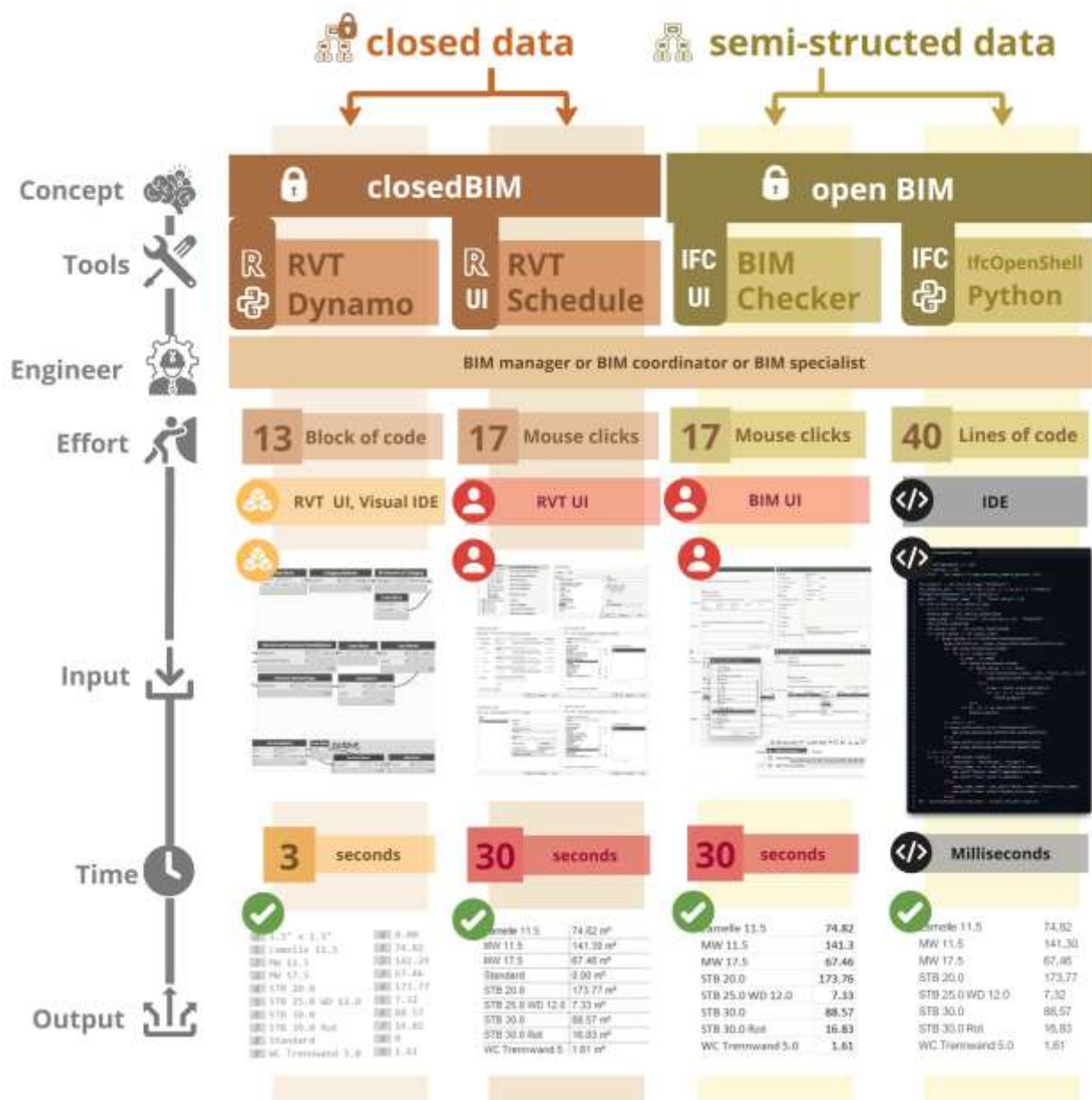
מספיק לבצע כ-17 פעולות) לחיצות כפתורים (כדי לקבל, CAD (BIM) לדוגמה, באמצעות ממשק המשתמש של תוכנת CAD טבלת כמויות מוכנה) איור 5.24. (עם זאת, על המשתמש להבין היטב את מבנה המודל ואת הפונקציות של תוכנת (BIM).

מספר הפעולות הידניות, CAD בתוך תוכנות API אם נעשה שימוש באוטומציה דרך קוד תוכנה או דרך תוספים וכלים: לקבלת טבלאות כמויות מצטמצם, אך יהיה צורך לכתוב בין 40 ל-150 שורות קוד, בהתאם לספרייה או הכלי בשימוש

■ QTO מאפשרים לקבל טבלת – Dynamo IronPython (closed BIM) או IfcOpSh (open BIM)
תוך כ-40 שורות קוד CAD או תוכנת CAD מתוך פורמט

■ IFC דורש כ-150 שורות קוד כדי לחלץ את המאפיינים הכמותיים מתוך מודל – IFC_js (open BIM)

■ מאפשרים להשיג את אותו התוצאה ידנית, תוך 17 לחיצות עכבר – CAD (BIM) כלים ממשקיים של



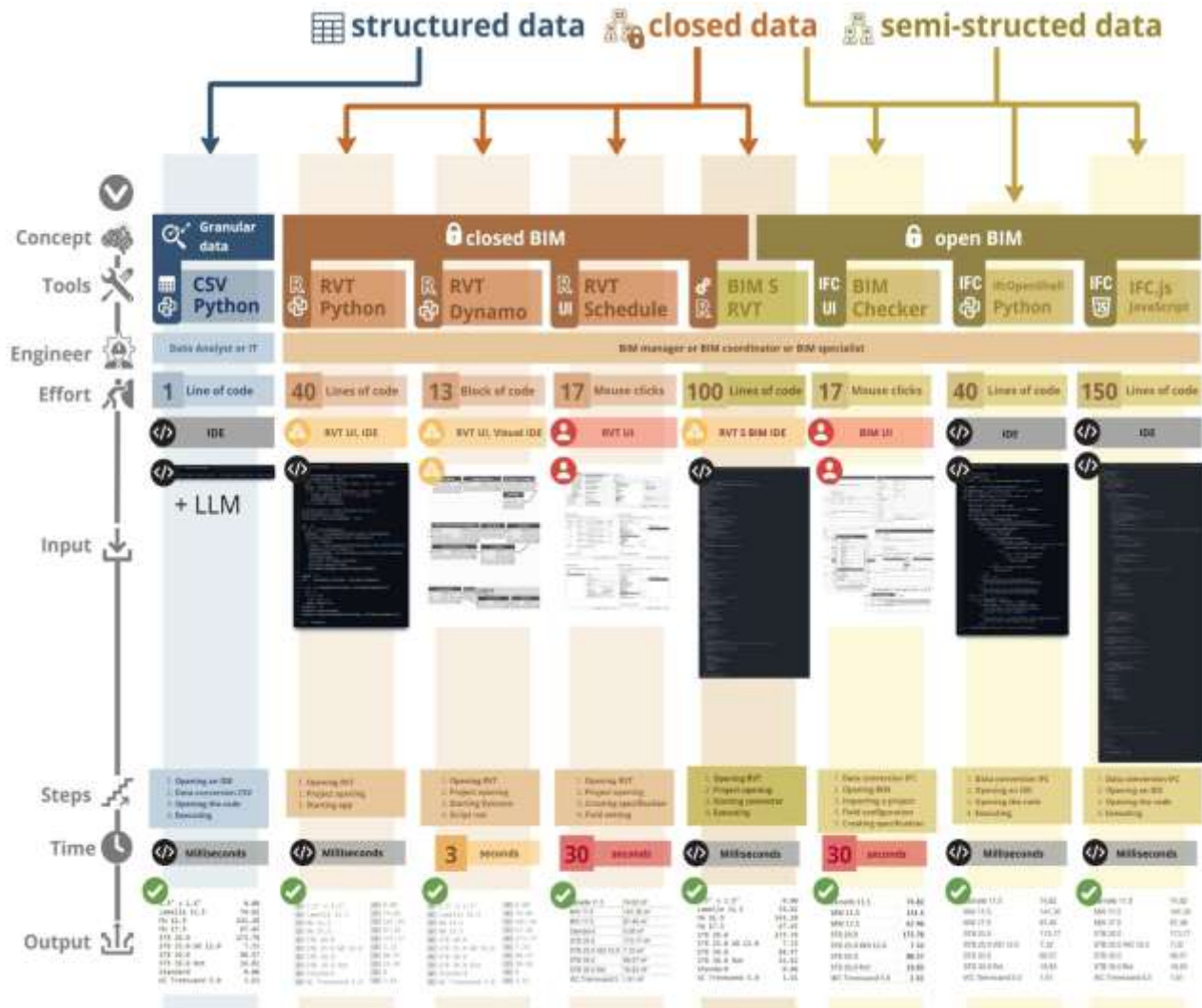
משתמשים בין 40 ל-150 שורות קוד או עשרה לחיצות כפתורים כדי ליצור טבלאות CAD (BIM) איור 5.24 מתכננים ומנהלי QTO.

כתוצאה מכך, התוצאה הסופית היא אותה תוצאה – טבלה מסודרת עם מאפיינים כמותיים עבור קבוצת אלמנטים ההבדל הוא רק במאמץ הנדרש וברמת ההכשרה הטכנית הנדרשת מהמשתמש (איור 5.24). (הכלים המודרניים, ביחס ומפחיתים את הסיכון לטעויות. הם מאפשרים QTO-לאיסוף הכמויות הידני, מאיצים באופן משמעותי את תהליך ה לחלץ נתונים ישירות מתוך מודל הפרויקט, תוך כדי הוצאת הצורך לחשב את הכמויות ידנית על פי השרטוטים, כפי שנעשה בעבר.

זהו עם QTO ניתן להשיג טבלת – closed BIM או open BIM ללא קשר לשיטה בה נעשה שימוש – בין אם (BIM-) CAD- כמויות של אלמנטים בפרויקט (איור 5.24). (עם זאת, בעבודה עם נתוני פרויקט במושגים של

המוצעים על ידי ספקים) איור 3.213. (זה יוצר רמות נוספות של תלות API-המשתמשים תלויים בכלים מיוחדים וב--ודורש לימוד של סכמות נתונים ייחודיות, תוך כדי הגבלת הגישה הישירה לנתונים

ומאפיינים אחרים מקשה על אוטומציה של חישובים ואינטגרציה QTO השגת טבלאות, CAD בשל הסגירות של נתוני של הפרויקט באמצעות כלים של CAD עם מערכות חיצוניות. בעזרת כלים לגישה ישירה לבסיסי נתונים והמרת נתוני זהה תוך שורת קוד QTO) איור 4.113 (ניתן להשיג טבלת DataFrame הנדסה הפוכה לפורמט מובנה פתוח של --אחת בלבד) איור 5.25 – אפשרות עם נתונים גרנולריים)



איור 5.25 כלים שונים נותנים תוצאות זהות בצורת טבלאות מאפיינים של ישויות הפרויקט, אך עם מאמצים שונים.

CAD (BIM) כפי שהוזכר בפרק "המרת נתוני CAD, כאשר נעשה שימוש בנתונים מובנים פתוחים מתוך פרויקטי מתפשט באופן משמעותי, QTO, לצורת מבנה, "תהליך הקיבוץ

פטורות, CAD גישות המבוססות על שימוש בנתונים מובנים פתוחים או גישה ישירה לבסיסי נתונים של מודלי הן מתבססות על כלים מוכחים, אשר כבר בשימוש בתעשיות BIM. מהמגבלות השיווקיות הקשורות בראשי התיבות -ETL איור 7.310 תהליך) אחרות

בהתאם למחקר של מקינזי "נתונים פתוחים: חשיפת חדשנות ופרודוקטיביות באמצעות מידע זורם", שנערך בשנת 2013, השימוש בנתונים פתוחים יכול ליצור הזדמנויות לחיסכון של בין 30 ל-50 מיליארד דולר בשנה בתכנון, הנדסה, רכש ובניית מתקני אנרגיה. משמעות הדבר היא חיסכון של 15% בעלויות ההון לבנייה.

BIM עבודה עם נתונים פתוחים ומסודרים (גרנולריים) (מקלה על חיפוש ועיבוד מידע, מפחיתה את התלות בפלטפורמות מיוחדות ופותחת את הדרך לאוטומציה מבלי צורך להשתמש במערכות קנייניות או במודלים נתונים פרמטריים ומורכבים CAD. בפורמטים

נתונים מובנים LLM באמצעות QTO אוטומציה של

המרת נתונים לא מסודרים לצורת נתונים מסודרת מעלה באופן משמעותי את היעילות של תהליכים שונים: היא מקלה על עיבוד הנתונים ומאיצה את תהליך האימות, מה שהופך את הדרישות לברורות ושקופות, כפי שכבר דיברנו בפרקים לצורת נתונים פתוחה ומסודרת מקלה על תהליך קבוצת האטריבוטים (BIM) CAD הקודמים. באופן דומה, המרת נתוני QTO ותהליך

אנו, CAD נמצאת בצורה מסודרת, ולכן כאשר אנו משתמשים בנתונים מסודרים של QTO טבלת האטריבוטים של עובדים עם מודל נתונים אחיד, מה שמפנה אותנו מהצורך לעסוק בהמרה ובתרגום של מודלי נתונים של פרויקטים וכללי קבוצות לאותו מכנה משותף. זה מאפשר לקבץ נתונים לפי אטריבוט אחד או יותר באמצעות שורת קוד אחת בלבד שבהם הנתונים מאוחסנים בפורמטים חצי-מסודרים, פרמטריים או, BIM-closed וב BIM-open בניגוד לכך, ב, כדי לתקשר עם הגיאומטריה והמידע API-סגורים, העיבוד דורש עשרות או אפילו מאות שורות קוד, כמו גם שימוש ב-האטריבוטי

LLM של פרויקט מסודר לפי אטריבוט אחד. בקשה טקסטואלית בכל צ'אט QTO דוגמה לקיבוץ (או כל אחד אחר ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN):

אנא סנן את נתוני הפרויקט כך שיתקבלו אלמנטים – DataFrame בצורת CAD יש לי פרויקט "Type 1" מכיל רק את הערך "Type" שבהם הפרמטר

Pandas-תוך שימוש ב Python-תהיה בסבירות גבוהה בצורת קוד ב LLM-התשובה מ



ולקבל את קבוצת האלמנטים "Type" לפי האטריבוט CAD-מאפשרת לקבץ את כל פרויקט ה LLM, שורת קוד אחת, שנכתבה בעזרת הנדרשת.

את הסכימה ומודל הנתונים, מה שמקצר LLM-דו-ממדי, אין צורך להסביר ל DataFrame בזכות המבנה הפשוט של את שלבי הפרשנות ומאיץ את יצירת הפתרונות הסופיים. בעבר, כדי לכתוב אפילו קוד פשוט היה צורך ללמוד שפות מאפשרים להמיר אוטומטית את הלוגיקה של התהליך לקוד בעבודה (LLM) תכנות, אך כיום מודלים לשוניים מודרניים עם נתונים מסודרים באמצעות בקשות טקסטואליות.

יכולים לחלוטין לשחרר את המומחים העובדים עם קיבוץ ועיבוד נתוני LLM אוטומציה ומודלים לשוניים ומספקים אפשרות לפתור בעיות BIM, מהצורך ללמוד שפות תכנות או כלים של CAD (BIM) באמצעות בקשות טקסטואליות.

– אותה בקשה – קיבוץ כל האלמנטים של הפרויקט מקטגוריית "קירות" עם חישוב הנפחים לפי כל סוג (ריש 5.25) דורשת 17 לחיצות בממשק או כתיבת 40 שורות קוד, בכלים לעיבוד נתונים פתוחים) כגון CAD (BIM) אשר בסביבת נראית כמו בקשה פשוטה ואינטואיטיבית (Pandas או SQL).

Pandas-בעזרת שורה אחת ב

```
df[df['Category'].isin(['OST_Walls'])].groupby('Type')['Volume'].sum()
```

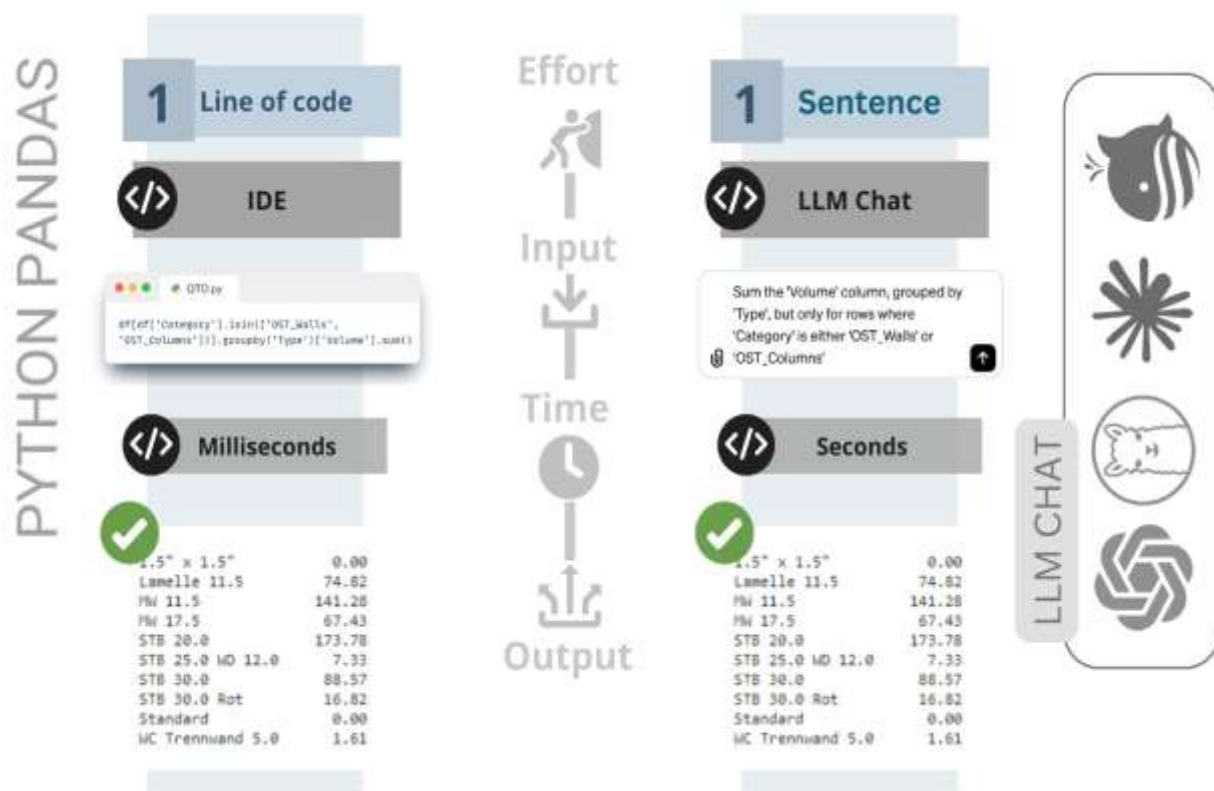
"OST_Walls", את האלמנטים שבהם העמודה "קטגוריה" מכילה את הערך df (DataFrame) פענוח הקוד: קח מתוך "וסכם עבור הקבוצה שהתקבלה את העמודה "נפח "Type" קבץ את כל האלמנטים שהתקבלו לפי העמודה

SQL: בעזרת CAD-קיבוץ פרויקט מובנה שהתקבל מ

```
SELECT Type, SUM(Volume) AS TotalVolume FROM elements WHERE Category = 'OST_Walls' GROUP BY Type;
```


ניתן לנסח את הבקשה לקיבוץ בבסיס הנתונים של הפרויקט בצורה של פנייה טקסטואלית פשוטה LLM בעזרת - (5.27 איור) פרומפט -

אך רק עבור 'Type' של הפרויקט, קבץ את האלמנטים לפי הפרמטר DataFrame עבור
 וסכם 'OST_Columns' או 'OST_Walls'-שווה ל 'Category' האלמנטים שבהם הפרמטר
 ↪ עבור הקבוצה שהתקבלה 'Volume' בבקשה את העמודה

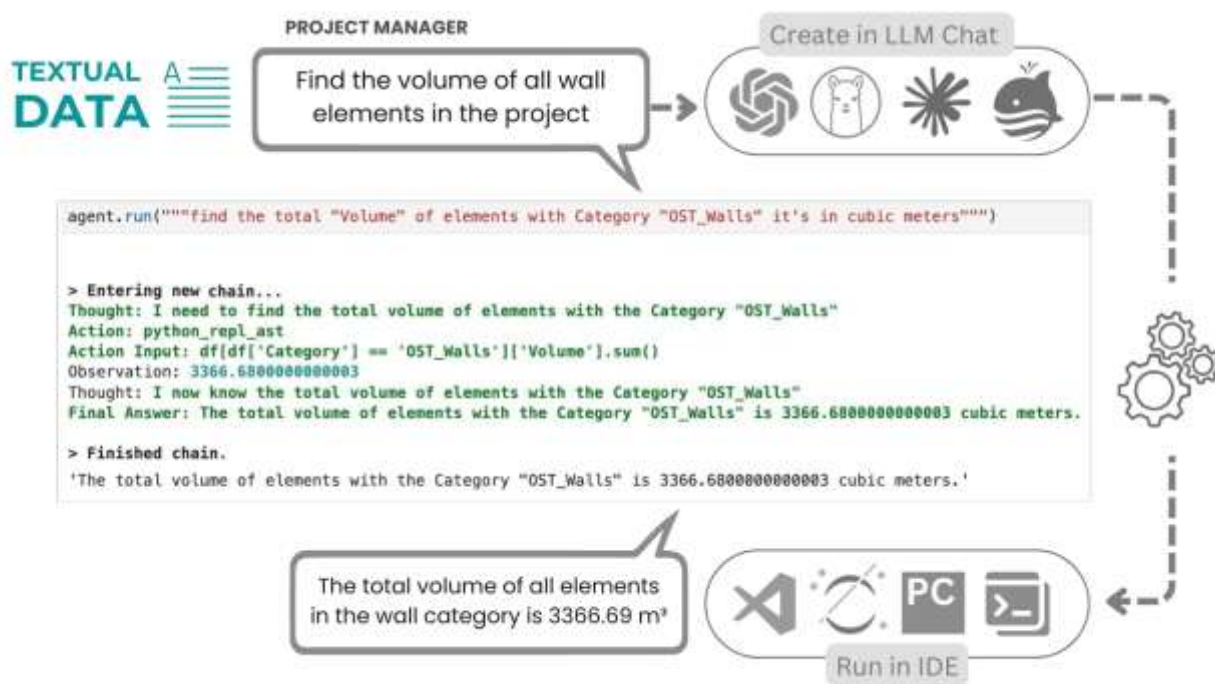


אוטומציה של עיבוד נתונים אפשרית כעת בעזרת מספר שורות קוד LLM ו-Pandas-SQL בזכות השימוש ב 5.27 איור ובקשות טקסטואליות.

משנה באופן דרמטי את השיטות המסורתיות להפקת מידע אטריביוטיבי, נתונים כמותיים ונפחיים עבור (QTO, QWEN, Grok, Claude, DeepSeek, Mistral, LLaMa, ChatGPT) LLM באמצעות כלים של CAD מנתוני QTO קבלת אובייקטים בודדים וקבוצותיהם.

כעת גם מנהלי פרויקטים, מומחים בהערכות או לוגיסטיקה, שאינם בעלי ידע מעמיק בתכנון ואינם לקבל תוך CAD ייחודיות, יכולים, לאחר גישה לבסיס הנתונים של CAD (BIM) מחזיקים בתוכנות שניות ספורות את הנפח הכולל של אלמנטים מקטגוריית הקירות או אובייקטים אחרים, פשוט על ידי כתיבה או דיבור של הבקשה.

מעבד את הפנייה של המשתמש כדי ליישם פונקציה מסוימת על אחד LLM מודל (5.28 איור) בבקשות טקסטואליות מקבל או עמודה-פרמטר חדשה עם LLM או יותר מהפרמטרים – עמודות הטבלה. כתוצאה מכך, המשתמש בשיח עם ערכים חדשים, או ערך ספציפי אחד לאחר הקיבוץ.



בעבודתו עם נתונים מובנים, מבין מההקשר של הבקשה הטקסטואלית על איזו קיבוץ ואטריבוטים שואל, LLM מודל 5.28 איור המשתמש.

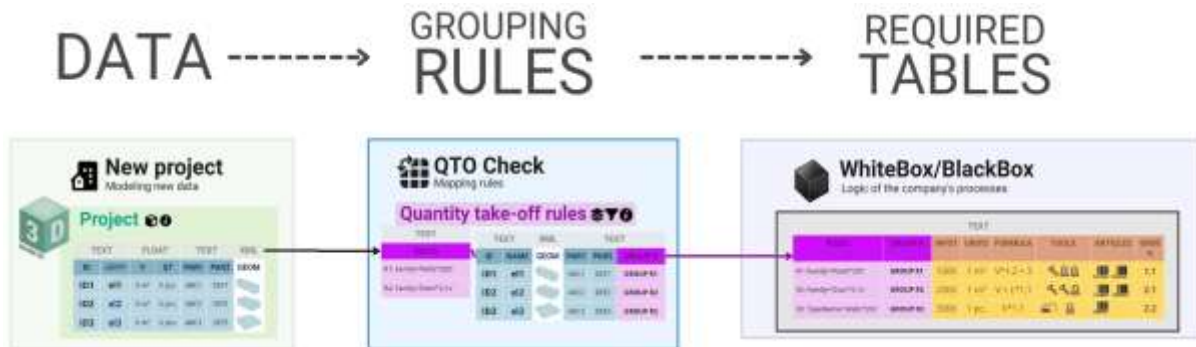
לנתוני (5.27 איור) פשוטה QTO אם יש צורך לקבל נתוני נפח רק עבור קבוצה אחת של אלמנטים, מספיק לבצע בקשת עם זאת, בעת חישוב התקציב או ההערכה עבור כל הפרויקט, המורכב ממספר קבוצות של אלמנטים, לעיתים CAD מודל קרובות יש צורך בהפקת מאפיינים כמותיים עבור כל סוגי האלמנטים (קטגוריות, (כאשר כל קטגוריה של אלמנטים מעובדת בנפרד – עם קיבוץ לפי האטריבוטים המתאימים.

בפרקטיקה של מערכי עלויות והערכות, נעשה שימוש בכללים אישיים לקיבוץ וחישוב עבור סוגים שונים של (Level, Rooms פרמטר הקיבוץ – אטריבוט) אובייקטים. לדוגמה, חלונות בדרך כלל מקובצים לפי קומות או אזורים כדי לאוטומט את תהליך הקיבוץ, כללים אלו. (Material, Type פרמטר) ואילו קירות – לפי חומר או סוג בנייה מתוארים מראש בצורה של טבלאות כללים לקיבוץ. טבלאות אלו פועלות כ, המגדירים אילו אטריבוטים יש להשתמש בחישובים עבור כל קבוצת אלמנטים בפרויקט.

Excel של כל הפרויקט באמצעות כללים לקבוצות מטבלת QTO חישוב

בפרויקטים בנייה אמיתיים לעיתים קרובות עולה הצורך לבצע אגרגציות על מספר מאפיינים בו זמנית בתוך קבוצה אחת OST_Windows של אלמנטים. לדוגמה, בעבודה עם הקטגוריה "חלונות" (כאשר המאפיין קטגוריה מכיל ערכים כמו ניתן לקבץ את האלמנטים לא רק לפי סוג – לדוגמה, לפי הערך בשדה שם הסוג או סוג, – אלא, IfcWindows או

גם לפי מאפיינים נוספים, כמו רמת הולכת חום, המצוינת במאפיין המתאים. קיבוץ רב-ממדי כזה מאפשר לקבל תוצאות מדויקות יותר עבור קבוצה ספציפית. באופן דומה, בעת חישובים בקטגוריות של קירות או תקרות ניתן להשתמש בשילובים שרירותיים של מאפיינים – לדוגמה, חומר, רמה, קומה, עמידות באש ופרמטרים נוספים – כמסננים או (ריש. 5.29) קריטריוני קיבוץ.



עבור כל קבוצה או קטגוריה של ישויות בפרויקט קיימת נוסחת קיבוץ משלה, המורכבת מאחד או יותר קריטריונים (ריש. 5.29).

תהליך הגדרת כללי הקיבוץ הללו דומה לתהליך יצירת דרישות נתונים, המתואר בפרק "יצירת דרישות ובדיקת איכות שבו דנו בפירוט בעבודה עם מודלים של נתונים. כללי קיבוץ וחשוב כאלה מבטיחים את הדיוק, (ריש. 4.45) "נתונים והרלוונטיות של התוצאות לחישוב אוטומטי של מאפיינים סיכומיים של כמות או נפח של קטגוריית ישויות, תוך התחשבות בכל התנאים הנדרשים שיש לקחת בחשבון במהלך החישובים והקלטות.

הדוגמה הבאה של קוד מסננת את טבלת הפרויקטים כך שהקבוצה הנתונים המתקבלת מכילה רק את ובו "IfcWindows" או "OST_Windows" הישויות שבהן המאפיין-עמודה "קטגוריה" מכיל ערכים "זמנית המאפיין-עמודה" סוג "מכילה את הערך" סוג 1:

של הפרויקט - סנן את הנתונים כך שבקבוצת הנתונים יישארו רק אלמנטים DataFrame יש לי ובו זמנית "IfcWindows" או "OST_Windows" שבהם המאפיין "קטגוריה" מכיל ערכים "המאפיין סוג מכיל את הערך" סוג 1

LLM: תשובת



מאפשרת לקבץ את כל הישגיות של הפרויקט לפי מספר קריטריונים. **Excel** שורת קוד אחת. הדומה לנוסחת **5.210**. ריש

ניתן להרצה (4.113. ריש) לפורמטים פתוחים ומסודרים CAD לאחר המרת נתוני (5.210. ריש) הקוד המתקבל
(PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse עם תוסף
PyDev, Thonny, Wing IDE, IntelliJ IDEA עם Python, JupyterLab פופולריים
Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker.-

בקטגוריה "חלונות" רק עם ערך QTO DataFrame כדי לקבל את הישויות של הפרויקט בצורה של LLM-מסוים של הולכת חום. אנו יכולים להשתמש בשאילתא הבאה ל

של הפרויקט - סנן את הנתונים כך שבקבוצת הנתונים יישארו רק רשומות עם DataFrame יש לי ובו זמנית העמודה, "IfcWindows" או "OST_Windows" קטגוריה, "המכילה ערכים"

LLM: תשובת



עבור מספר פרויקטים בו זמנית QTO מאפשרת לבצע Pandas Python שפת השאילתות הפשוטה של 5.211. ריש

נעשה שימוש בתנאי לוגי "&" כדי לשלב בין שני קריטריונים: ערך הולכת, (5.211. ריש) LLM-בתשובה שהתקבלה מ בודקת אם הערך של המאפיין-עמודה "קטגוריה" נמצא "isin" החום והשתייכות לאחת משתי הקטגוריות. השיטה -ברשימה שניתנה

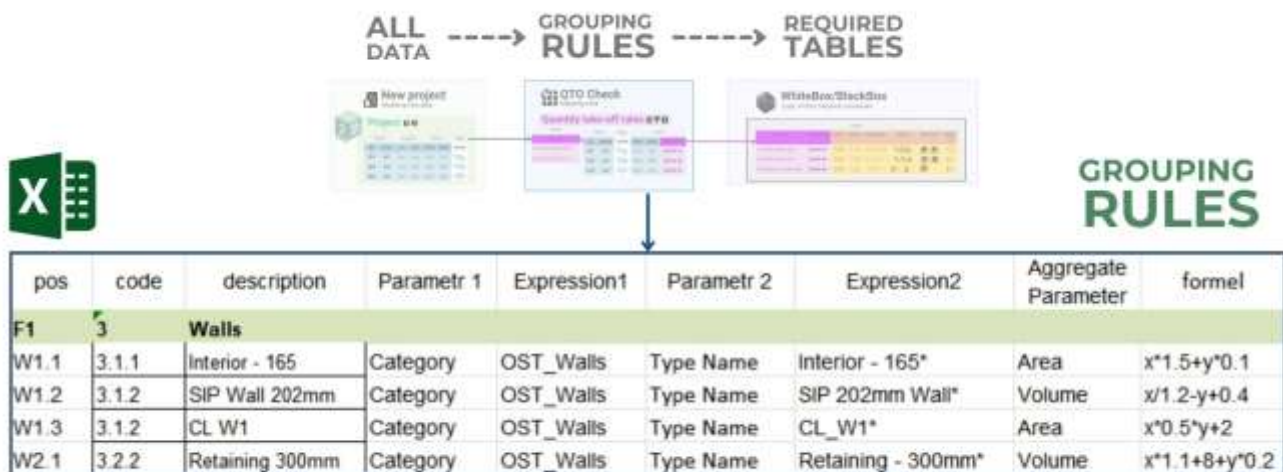
בפרויקטים עם מספר רב של קבוצות רכיבים, עם לוגיקה שונה של קיבוץ - עבור כל קטגוריית ישויות בפרויקט (למשל חלונות, דלתות, תקרות (יש לקבוע כללים קיבוץ אישיים, שיכולים לכלול גורמים נוספים או נוסחאות סופיות לחישוב והגורמים (של נפח הקבוצה y של כמות וערך x למשל ערך, "formel", מאפיין 5.212 איור) מאפיינים. נוסחאות אלו -מתחשבים בתכונות הייחודיות של כל קבוצה, למשל

■ אחוזים נוספים לנפח החומר לצורך התחשבות בהוצאות יתר

■ כמות חומר נוספת קבועה

■ תיקונים הקשורים לסיכונים אפשריים ולשגיאות חישוב בצורה של נוסחאות

לאחר שנוסחו כללי הסינון והקיבוץ בצורה של נוסחאות פרמטרים עבור כל קטגוריית רכיבים, ניתן לשמור אותם בצורה שמירת כללים אלו בצורה מסודרת מאפשרת אוטומציה. (5.212 איור) Excel של טבלה שורתית - למשל, בפורמט מלאה של תהליך חילוף, סינון וקיבוץ נתוני הפרויקט. במקום לכתוב ידנית מספר רב של שאילתות נפרדות, המערכת פשוט קוראת את טבלת הפרמטרים ומיישמת את הכללים המתאימים על המודל (הדאטה פריים הכללי של הפרויקט) איור -סופיות עבור כל קטגוריית רכיבים בפרויקט QTO ומביאה ליצירת טבלאות, (4.113)



קובעת את כללי קיבוץ רכיבי הפרויקט, ומבטיחה כמות ונפח מדויקים עבור כל קטגוריה QTO טבלת קיבוץ מאפייני 5.212 איור

הכללים שנאספו יאפשרו לקבץ את כל הפרויקט ולבצע את כל החישובים הנדרשים, כולל תיקון מאפייני הנפח. כתוצאה מכך, הנפחים מתואמים ל"נפח אמיתי", אשר משמש לחישובים ולחשבונות, ולא אלו שהיו בתחילת שלב התכנון במודל CAD.

אוטומטיות עבור כל הפרויקט, האפליקציה צריכה לעבור על כל הקטגוריות QTO במהלך יצירת טבלאות של טבלת כללי הקיבוץ, לקחת את מאפייני הקיבוץ, לקבץ על פיהם את כל רכיבי הפרויקט ולאגד את מאפייני הנפח עבור קבוצה זו, בנוסף להכפיל אותו בגורם או גורם מבהיר

וטבלת (5.212 איור) לכתוב עבורנו קוד לפתרון כזה, שבו הקוד יטען שתי טבלאות - טבלת כללי הקיבוץ LLM-נבקש מ ולאחר מכן יישם את כללי הקיבוץ, יקבץ את הרכיבים בהתאם לכללים שנקבעו, (4.113 איור) נתוני הפרויקט עצמו --חדש Excel יחשב ערכים מצטברים וישמור את התוצאות בקובץ

LLM: שלחו בקשה טקסטואלית בצ'אט של

ולאחר מכן, "basic_sample_project.xlsx" אני צריך קוד לקריאת נתוני הפרויקט מקובץ 'Parameter 1' על פיהם לקבץ את כל הנתונים לפי "Grouping_rules_QTO.xlsx"-כללים מ לבצע, 'Expression2' לסנן לפי, 'Aggregate Parameter' לאגד, 'Parameter 2'-ו 'Formel1'-חשובים מ 'QTQ_table2.xlsx'-ב QTQ ולשמור את טבלת 'Formel1'-חשובים מ

LLM: תשובת

```

1 import pandas as pd
2
3 # Reading CSV and Excel files with project data and grouping rules respectively
4 df = pd.read_csv('../input/rac_basic_sample_project.xlsx', low_memory=False)
5 df_qto = pd.read_excel('../input/Grouping_rules_QTO.xlsx', header=0, skiprows=[1])
6
7 # Cleaning numerical columns in the dataframe by removing non-numeric characters
8 numeric_cols = ['Area', 'Volume', 'Depth', 'Width', 'Length', 'Perimeter']
9 df[numeric_cols] = df[numeric_cols].apply(lambda col: pd.to_numeric(col.astype(str).str.extract('(\d*\.\d*)')[0], fillna(0)))
10
11 # Define a function to process each row in the grouping rules dataframe
12 def process_row(row):
13     # Group by specified parameters and calculate sum and count for the aggregate parameter
14     group = df.groupby([row["Parameter 1"], row["Parameter 2"]])
15     [[row["Aggregate Parameter"], group.agg(['sum', 'count'])]]
16     group = group[group.index.get_level_values(1).str.contains(row["Expression2"], regex=True)]
17
18     # Perform calculation defined in the rules
19     group["After Calculation"] = eval(row["formel"], {'x': group['sum'], 'y': group['count']})
20     return group.reset_index()
21
22 # Apply the function to each row in the grouping rules dataframe and concatenate the results
23 dfgroups = pd.concat([process_row(row) for _, row in df_qto.iterrows()]).reset_index(drop=True)
24
25 # Save the processed data to an Excel file using the OpenPyXL engine
26 dfgroups.to_excel('QTQ_table2.xlsx', engine='openpyxl', index=False)
27

```

Excel-מקבץ אותם בהתאם לכללים שנקבעו ב, CAD מעבד את נתוני הפרויקט מתוך פרויקט Python Pandas קוד 5.213 איור

תהיה טבלת קבוצות-ישויות, אשר מכילה לא רק מאפיינים מצטברים (5.213 איור) התוצאה הסופית של ביצוע הקוד אלא גם מאפיין חדש של נפח אמיתי, המתחשב בכל הדרישות ליצירת חישובים, CAD- (BIM) של נפח מהמודל המקורי -(5.214 דוגמה איור) והערכות נכונים

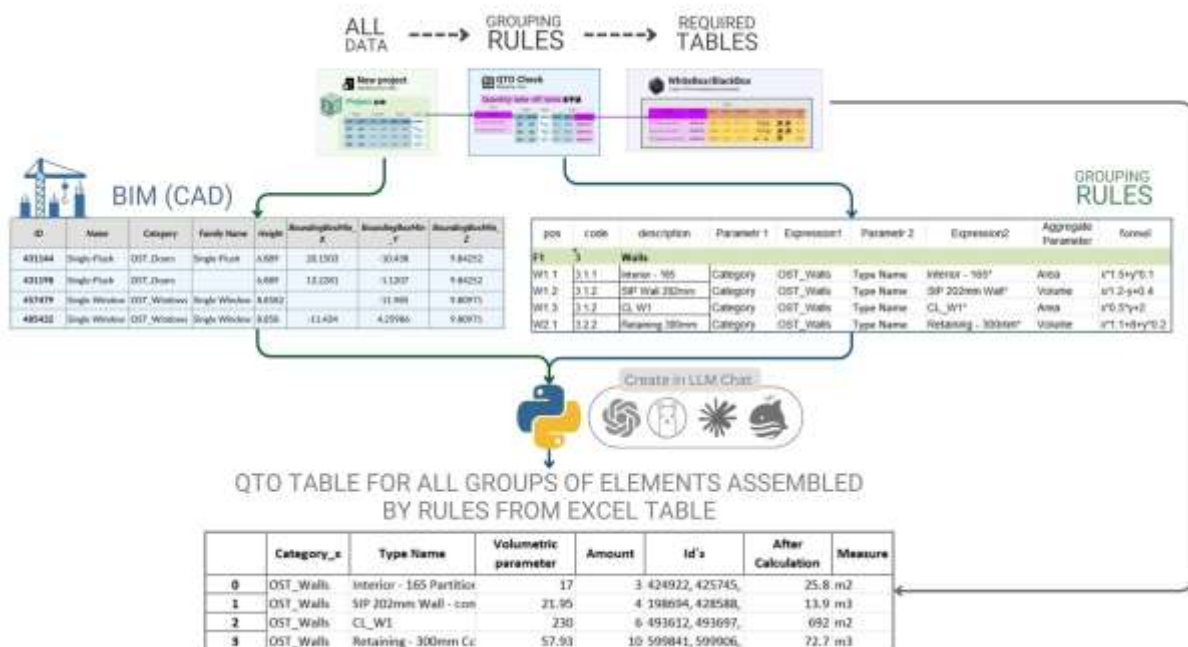


QTO TABLE FOR ALL GROUPS OF ELEMENTS ASSEMBLED BY RULES FROM EXCEL TABLE

	Category_x	Type Name	Volumetric parameter	Amount	Id's	After Calculation	Measure
0	OST_Walls	Interior - 165 Partitor	17	3 424922, 425745,		25.8	m2
1	OST_Walls	SIP 202mm Wall - con	21.95	4 198694, 428588,		13.9	m3
2	OST_Walls	CL_W1	230	6 493612, 493697,		692	m2
3	OST_Walls	Retaining - 300mm Cc	57.93	10 599841, 599906,		72.7	m3

מאפיין "לאחר חישוב" מתוסף לטבלה הסיכומית לאחר ביצוע הקוד, אשר יחשב אוטומטית את הנפח האמיתי 5.214 איור.

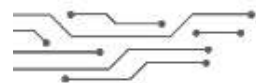
הפופולריות) עליהם דיברנו קודם (וניתן ליישם קוד זה על IDE-ניתן להרצה באחת מה (5.213 איור) הקוד שהתקבל בין אם מדובר בכמה פרויקטים, (ועוד RVT, IFC, DWG, NWS, DGN) כל כמות של פרויקטים קיימים או חדשים (5.215 איור) או, אולי, מאות פרויקטים בפורמטים שונים, המובאים בצורה מסודרת.



דרך כללים מתוך QTO עם טבלאות BIM (CAD)-תהליך הארגון האוטומטי של נתוני הבנייה מקשר את נתוני ה 5.215 איור Excel. גיליון אלקטרוני של

מאפשר אוטומציה מלאה של איסוף נתונים על (5.215 איור) תהליך האיסוף המותאם והמוגדר של נתונים כמותיים מאפיינים כמותיים ונפחים של רכיבי הפרויקט לצורך עבודה נוספת איתם, כולל הערכת עלויות, לוגיסטיקה, לוחות זמנים וחישוב טביעת רגל פחמנית ומשימות אנליטיות אחרות.

לאחר שלמדנו על הכלים המאפשרים לארגן ולמייין קבוצות של רכיבי פרויקט לפי קריטריונים מסוימים, אנו מוכנים כעת לשלב את הפרויקטים הממיינים והמנוטרים עם חישובים שונים ותסריטי עסק של החברה.



פרק 5.3. CO₂ וחישוב פליטות פחמן דו-חמצני 4D, 6D-8D

שילוב זמן בהערכות בנייה: Dמודל 4

— מעבר לחישוב העלויות, אחד הכיוונים המרכזיים של שימוש בנתוני פרויקטים בבנייה הוא קביעת פרמטרים זמניים הן עבור פעולות בנייה בודדות והן עבור הפרויקט כולו. לצורך חישוב אוטומטי של מועדים והכנת לוח זמנים לביצוע עבודות, לעיתים קרובות נעשה שימוש בשיטת הערכת משאבים ובסיס הנתונים הקשור לכך, אשר נבחן בפירוט בפרק "הקודם" חישובים והערכות עבור פרויקטים בנייה.

במסגרת הגישה המשאבית, נלקחים בחשבון לא רק העלויות של חומרים, אלא גם משאבים זמניים. בעת הכנת חישובים, וכן לציין את כמות הזמן, ("פרמטר" סדר עבודה – 5.31 איור) ניתן להקצות לכל תהליך מאפיין של סדר ביצוע עבודות והעלות הקשורה לביצוע תהליך זה. פרמטרים אלה חשובים במיוחד לתיאור פעולות שאין להן מחיר שוק קבוע ואינן ניתנות לרכישה ישירה — כמו שימוש בציוד בנייה, תעסוקת עובדים או תהליכי לוגיסטיקה (שבדרך כלל מתבטאים בשעות. (במקרים כאלה, העלות נקבעת לא על ידי מחלקת הרכישות, אלא ישירות על ידי החברה המבצעת על בסיס 5.31 איור) תקנים פנימיים או תעריפים ייצוריים.

Position: Concrete Foundation Block per 1 piece.				Unit price: \$11,934.00			
Code	Description	Bid. Factor	Quantity	Unit	Unit cost	Total	Work order
LabPr	Preparation Work Labor	1	16.00	hr	€ 30.00	€ 480.00	1.1
EqEx	Excavation Equipment	1	16.00	hr	€ 80.00	€ 1,280.00	1.2
FormW	Formwork	1	500.00	sq ft	€ 2.50	€ 1,250.00	
ReSt	Reinforcing Steel	1	800.00	lb	€ 0.75	€ 600.00	
ConcB	Concrete	1	30.00	cu yd	€ 120.00	€ 3,600.00	
LabCP	Concrete Pouring Labor	1	24.00	hr	€ 35.00	€ 840.00	2.1
EqCM	Concrete Mixer	1	8.00	hr	€ 45.00	€ 360.00	2.2
LabFI	Finishing Labor	1	24.00	hr	€ 30.00	€ 720.00	2.3
LabCu	Curing Labor	1	8.00	hr	€ 25.00	€ 200.00	3.1
EqOt	Other Equipment	1	10.00	hr	€ 15.00	€ 150.00	3.2
FuelD	Diesel for Equipment	1	40.00	gal	€ 3.50	€ 140.00	
MiscL	Lubricants and Maintenance	1	1.00	lump sum	€ 200.00	€ 200.00	
TransM	Transportation of Materials	1	1.00	lump sum	€ 300.00	€ 300.00	
OverH	Overhead Costs	1	1.00	percentage	10% of Total	€ 907.00	
Prof	Profit Margin	1	1.00	percentage	10% of Total	€ 907.00	

Code	Description	Unit	Unit cost
Exc_2	Delivery of the equipment to the site	hour	80
Exc_3	Delivery to the place of installation	hour	100
transport	Material profit	kg	10.00

חישובי עבודות בשיטת הערכת משאבים כוללים את זמני העבודה 5.31 איור.

כך, בחישובים ברמת ההערכות נכללים לא רק העלויות של דלק וחומרים (עלות רכישה, אלא גם זמן העבודה של טבלת העלויות מייצגת את חישוב העלות, (5.31 איור) מפעילים, ציוד ועובדים נוספים באתר הבנייה. בדוגמה המוצגת של התקנת בלוק יסוד, כולל שלבי עבודה מרכיבים, כמו הכנה, התקנת שלד ויציקת בטון, כמו גם חומרים וזמן עבודה נדרשים. במקביל, פעולות מסוימות, כמו עבודות הכנה, עשויות לא לכלול עלויות חומריות, אך להכיל עלויות זמן משמעותיות, המובאות בשעות אדם.

לתכנון רצף העבודות (לוח הזמנים של העבודות) (באתר הבנייה, מוסיפים ידנית לטבלת החישוב את המאפיין "הזמנת עבודה") "איור 5.31. (הוא מצוין בעמודה נוספת רק עבור פריטים שהיחידה שלהם נמדדת בזמן) (שעה, יום). (מאפיין זה משלים את קוד העבודה, התיאור, הכמות, יחידת המדידה) פרמטר "יחידה" (והעלויות. הרצף המספרי) פרמטר "הזמנת עבודה" (של העבודות מאפשר לקבוע את סדר ביצוע המשימות באתר הבנייה ולהשתמש בו בעת הכנת לוח הזמנים).

לוח זמנים לבנייה ואוטומציה שלו על בסיס נתוני חישוב

לוח הבנייה הוא ייצוג חזותי של תוכנית ביצוע העבודות והמשאבים שצריכים להתבצע במסגרת יישום הפרויקט. הוא נוצר על בסיס חישובי משאבים מפורטים) איור 5.31, (כאשר כל משימה-עבודה מפורטת, בנוסף לעלות המשאבים, גם בזמן ובסדר.

בניגוד לגישות ממוצעות, שבהן חישובי הזמן מבוססים על כמות שעות טיפוסית להתקנת חומרים או ציוד, בשיטת המשאבים התכנון מתבסס על נתונים ממשיים המוטמעים בחישוב. כל פריט בהצעת המחיר, הקשור להוצאות עבודה מתבסס על לוח שנה המחשב את תנאי השימוש האמיתיים במשאבים במהלך התקופה העבודה. התאמת שעות הייצור באמצעות מקדמים ברמת החישוב) איור 5.31 פרמטר "מקדמי הצעה", ("מאפשרת לקחת בחשבון הבדלים בפרודוקטיביות ותנאים עונתיים המשפיעים על זמני ביצוע העבודות.

כדי לקבוע את תאריכי התחלה וסיום של התהליך עבור לוח הבנייה על גבי תרשים גנטה, אנו לוקחים את ערכי מאפיין הזמן עבור כל פריט מתוך חישוב הבלוקים הבסיסיים ומכפילים אותם בכמות הבלוקים) במקרה זה כמות הבלוקים הבטוניים. (חישוב זה נותן את משך כל משימה. לאחר מכן אנו מציבים את משכי הזמן הללו על ציר הזמן, החל מתאריך התחלה של הפרויקט, כדי לבנות את הלוח, ובסופו של דבר מקבלים ייצוג חזותי המראה מתי כל משימה צריכה להתחיל "ולהסתיים. פרמטר "הזמנת עבודה" בתהליכים מאפשר לנו בנוסף להבין אם התהליך מתבצע במקביל ("הזמנת עבודה 1.1-1.2) לדוגמה 1.1-1.1 (או בסדר

תרשים גנטה הוא כלי גרפי לתכנון וניהול פרויקטים, המייצג משימות בצורה של פסים אופקיים על ציר הזמן. כל פס מציג את משך ביצוע המשימה, את תחילתה ואת סיומה.

לוח העבודות, או תרשים גנטה, מסייע למנהלי הפרויקטים ולעובדים להבין בצורה ברורה מתי ובאיזו סדרה יש לבצע את שלבי הבנייה השונים, תוך הבטחת שימוש יעיל במשאבים ושמירה על לוחות הזמנים.

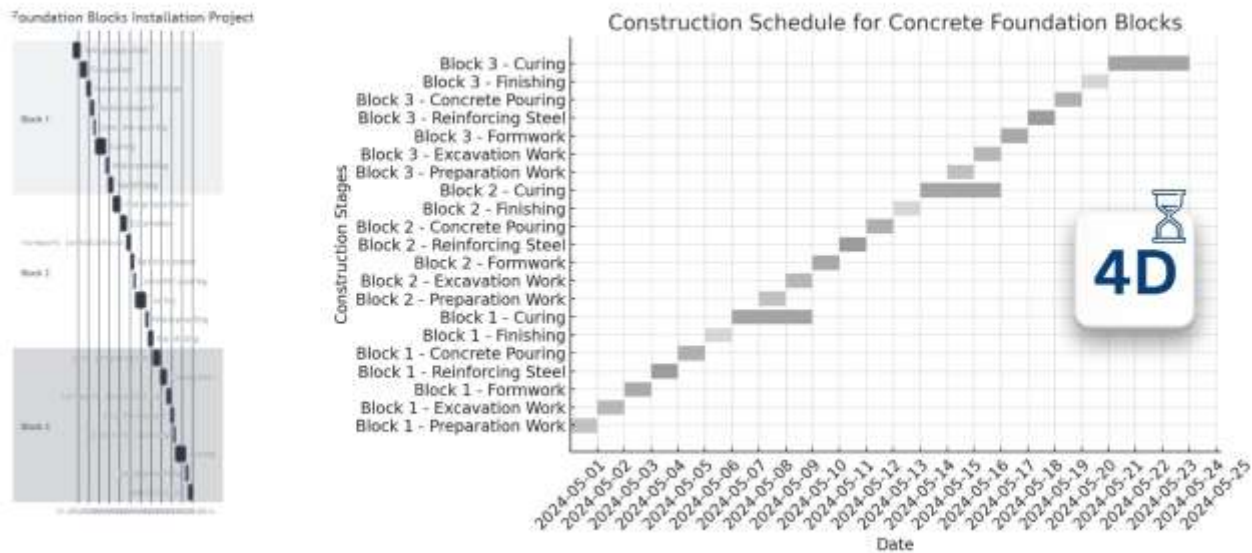
בניח תכנית עבודה להתקנת שלושה בלוקים בטוניים בסיסיים תוך שימוש בחישובים מתוך הטבלה למעלה. באמצעות לתכנן את התקנת 3 פריטי הבלוק הבסיסי, למשל, ל-1 LLM-טבלת העלויות) איור 5.31 (מהדוגמה למעלה, נבקש מ במאי 2024.

או פשוט להדביק צילום מסך XLSX אנו יכולים להעלות את טבלת החישוב בפורמט LLM-כדי לשלוח את החישוב ל ימצא בעצמו את הספרייה להמחשת LLM (איור 5.31) LLM ישירות בצ'אט של JPEG של טבלת החישוב בפורמט התמונה של הטבלה, וכאשר הוא יכפיל את מאפייני הזמן של העבודות מהטבלה בכמותן, הוא יסכם את כל הנתונים בלוח - הזמנים.

LLM: שלחו בקשה טקסטואלית ל

צור תרשים עבודה בצורת תרשים גנטה תוך שימוש בערכים מהטבלה הקשורים לזמן) איור 5.31
 בהוספה (להתקנת 3 בלוקי יסוד. העבודה על כל בלוק מתבצעת באופן סדרתי. התחלת JPEG-כ
 -ל העבודה נקבעה ל-01/05/2024

LLM: תשובת



מציג את שלבי הבנייה של שלושה בלוקי בטון, בהתאם לתנאים LLM איור 5.32 תרשים גנטה שנוצר אוטומטית על ידי מספר מההנחה.

התרשים שהתקבל (איור 5.32) (מהווה תרשים זמני, שבו כל פס אופקי תואם לשלב מסוים בביצוע העבודה על בלוק היסוד ומדגים את רצף הפעולות) פרמטר "סדר עבודה", ("כגון הכנה, עבודות עפר, התקנת קונסטרוקציה, חיזוק, יציאת בטון ועיבוד סופי, כלומר התהליכים שיש להם פרמטרים זמניים מלאים ורצף בחישובים).

תרשים דומה (איור 5.32) אינו לוקח בחשבון מגבלות הקשורות לימים עבודה, משמרות או תקני שעות עבודה, והוא מיועד אך ורק להמחשה קונספטואלית של התהליך. ניתן להוסיף לתוכנית מדויקת שתשקף את המקבילויות בעבודה על ידי הנחיות מתאימות או הנחיות נוספות בתוך הצ'אט.

באמצעות חישוב עלויות אחד) איור 5.31, (בזכות מאפייני הנפחים מהגיאומטריה התלת-ממדית, ניתן להעריך אוטומטית גם את עלות הפרויקט דרך אומדנים אוטומטיים וגם לחשב במקביל את המאפיינים הזמניים של הקבוצות בצורה של -טבלאות או תרשימים עבור אפשרויות שונות של הפרויקט) איור 5.33



במודל המידע המידע מרחיבים את ההתבוננות בהיבטים שונים של הפרויקט, מהתייעלות 6D, 7D, 8D ואור 5.34 מאפייני 6 אנרגטית ועד בטיחות.

בנוסף למסד הנתונים של הפרויקט (או מסגרת הנתונים) (עם מאפיינים גיאומטריים, 6D-6 נפחיים של רכיבים, מתווספת מידע) (מאפיינים-עמודות) (על קיימות סביבתית. זה כולל מידע הקשור להתייעלות אנרגטית, טביעת רגל פחמנית, אפשרות לחזור מחדש של חומרים ושימוש בטכנולוגיות ידעיות לסביבה. נתונים אלה מאפשרים להעריך את השפעת הפרויקט על הסביבה, לייעל את ההחלטות (ESG) התכנוניות ולהשיג את מטרות הקיימות.

תכונות משלימות תכונות הנדרשות לניהול תפעול הבניין. אלו נתונים על לוחות זמנים של תחזוקה, אורך 7D חיי רכיבים, תיעוד טכני והיסטוריית תיקונים. סט מידע כזה מאפשר אינטגרציה של המודל עם מערכות תפעול מאפשר תכנון יעיל של תחזוקה, החלפת ציוד ומספק תמיכה לאורך כל מחזור חיי (CAFM, AMS), הפרויקט.

שכבת תכונות נוספת כוללת מידע הקשור לבטיחות - הן בשלב הבנייה והן במהלך התפעול שלאחר מכן 8D. המודל מתווספות אמצעים להבטחת בטיחות העובדים, הוראות לפעולה במצבי חירום, דרישות למערכות פינוי והגנה מפני אש. אינטגרציה של נתונים אלו במודל הדיגיטלי מסייעת לקחת בחשבון סיכונים מראש ולפתח פתרונות ארכיטקטוניים, הנדסיים וארגוניים תוך התחשבות בדרישות של בטיחות ובריאות בעבודה.

(5.35 אזור) מייצגות תכונות נוספות בצורת עמודות עם ערכים מלאים עד 8D בצורה מסודרת בטבלה, השכבות מ-4 כגון שם, קטגוריה, סוג ומאפיינים תלת-ממדיים. הערכים בשכבות התכונה, D שנוספו לתכונות המלאות של מודל ה-3, מכילים נתונים טקסטואליים ודיגיטליים נוספים, כגון אחוז מחזור, טביעת רגל פחמנית, תקופת אחריות 6D, 7D 8D-8D. -'מחזור החלפה, תאריך התקנה, פרוטוקולי בטיחות וכו

3D				6D		7D		8D			
ID	Type Name	Width	Length	Recyclability	Carbon Footprint	Warranty Period	Replacement Cycle	Maintenance Schedule	Installation Date	Wellbeing Factors	Safety Protocols
W-NEW	Window	120 cm	-	90%	1622 kgCO ₂ e	8 years	20 years	Annual	-mon	XYZ Windows	ISO 45001
W-OLD1	Window	100 cm	140 cm	90%	1522 kgCO ₂ e	8 years	15 years	Biannual	08/22/2024	XYZ Windows	OSHA Standard
W-OLD2	Window	110 cm	160 cm	90%	1522 kgCO ₂ e	-	15 years	Biannual	08/24/2024	???	OSHA Standard
D-122	Door	90 cm	210 cm	100%	1322 kgCO ₂ e	15 years	25 years	Biennial	08/25/2024	Doors Ltd.	OSHA Standard

מוסיפים שכבות תכונה למודל המידע הדיגיטלי, שכבר מכיל תכונות גיאומטריות ותלת-ממדיות מתוך מודל 6D-8D 5.35 איור D-3.

-הבאות D-8D יכול לכלול את התכונות 3 (5.35 איור) W-NEW רכיב עם מזהה (4.41 איור) עבור החלון החדש שלנו -

CAD מידע גיאומטרי המתקבל ממערכות D תכונות 3

- "שם סוג - רכיב "חלון
- רוחב - 120 ס"מ
- כתכונה "BREP / MESH של הרכיב או את "גיאומטריית "Bounding Box" בנוסף ניתן להוסיף נקודות נפרדת

קיימות סביבתיות - D תכונות 6

- מדד "מיחזוריות" - 90%
- CO₂ טביעת רגל פחמנית - 1622 ק"ג

נתוני ניהול נכסים - D תכונות 7

- תקופת אחריות - 8 שנים
- מחזור החלפה - 20 שנים
- תחזוקה - נדרשת מדי שנה

הבטחת שימוש וביצוע בטוח של בניינים - D תכונות 8

- "XYZ Windows" חלון "הותקן" - על ידי חברת
- ISO 45001-תקן בטיחות - תואם ל"

נחוצים למומחים במחלקות שונות לצורך מיון (5.35 איור) כל הפרמטרים שנרשמו בבסיס הנתונים או בסט הנתונים חיפוש או חישובים. תיאור רב-ממדי כזה של אובייקטי הפרויקט על בסיס תכונות מאפשר לקבל תמונה מלאה על מחזור

חיים, דרישות תפעוליות ורבים אחרים, הנדרשים בתכנון, בנייה ותפעול הפרויקט.

וחישוב פליטות פחמן דו-חמצני בפרויקטים בבנייה CO₂ הערכת

בבנייה המודרנית מושם דגש מיוחד על הקיימות הסביבתית של D, לצד נושא הקיימות של פרויקטים בנייה בשלב 6 המתרחשות בשלבי CO₂ הפרויקטים, כאשר אחד ההיבטים המרכזיים הוא הערכה ומזעור של פליטות פחמן דו-חמצני. (מחזור חיי הפרויקט) למשל, במהלך הייצור וההתקנה)

הערכה וחישוב פליטות הפחמן של חומרי הבניין הוא תהליך שבו הפליטות הכוללות של הפחמן נקבעות על ידי הכפלת מאפיינים כמותיים של רכיבים או קבוצת רכיבים בשימוש בפרויקט, עם מקדם פליטה מתאים לקטגוריה זו.

סביבתיים (חברתיים) ESG התחשבות בפליטות הפחמן בהערכת פרויקטים בנייה, כחלק מהקריטריונים הרחבים יותר של ומנהליים, (מוסיפה רמה חדשה לניתוח המקיף. זה חשוב במיוחד עבור הלקוח-משקיע בעת קבלת תעודה מתאימה, כמו שיטת הערכת הסביבה של המכון לחקר הבניין (או BREEAM®, (מנהיגות באנרגיה ועיצוב סביבתי) LEED® (החברה הגרמנית לבנייה בת קיימא. (קבלת אחת מהתעודות הללו יכולה להעלות באופן משמעותי את DGNB® האטרקטיביות בשוק של הנכס, לפשט את תהליך ההפעלה ולהבטיח עמידה בדרישות השוכרים הממוקדים בקיימות איכות סביבתית גבוהה, תקן הבנייה הסביבתית HQE-בהתאם לדרישות הפרויקט, ניתן גם להשתמש ב. (ESG) מדד הקיימות הגלובלי) GRESB-1 (ממוקד בבריאות ובנוחות המשתמשים, WELL תקן הבניין) WELL, (הצרפתי (לנכסי נדל"ן).

(, סביבתי, חברתי ומנהלי (הוא סט רחב של עקרונות שניתן להשתמש בהם להערכת ניהול תאגידי ESG, השפעה חברתית וסביבתית של עסק הן בתוך החברה והן מחוצה לה.

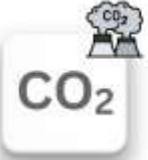
שפותח לראשונה בתחילת שנות ה-2000 על ידי קרנות השקעה כדי לספק למשקיעים מידע על קריטריונים ESG, רחבים של אופי סביבתי, חברתי ומנהלי, הפך למדד מרכזי להערכת חברות ופרויקטים, כולל בתחום הבנייה. על פי הופכת לחלק בלתי (ESG) מחקרים של חברות ייעוץ הגדולות ביותר, התחשבות בגורמים סביבתיים, חברתיים ומנהליים נפרד מתעשיית הבנייה.

לא ESG המסלול לנייטרליות פחמנית, "חברות המיישמות באופן פעיל עקרונות" (EY (2023 על פי נתוני רק מפחיתות סיכונים ארוכי טווח, אלא גם משפרות את היעילות של מודלי העסקים שלהן, דבר החשוב מצוין כי רמת "ESG-מודעות ל" PwC במיוחד בתנאים של טרנספורמציה גלובלית של השווקים. בדו"ח נעה בין 67% ל-97%, כאשר רוב הארגונים רואים במגמות ESG המודעות של חברות לגבי חשיבות גורמי אלו קריטיות לפיתוח בר קיימא בעתיד, וכי העסקים ברובם מדווחים על לחצים משמעותיים מצד בעלי עניין ESG. לשלב עקרונות

LEED, בפרויקטי בנייה לא רק תורם להשגת תעודות קיימות בינלאומיות, כמו ESG כך, שילוב עקרונות BREEAM, DGNB, מבטיח קיימות ארוכת טווח ותחרותיות של חברות בתעשייה.

אחד הגורמים המשמעותיים ביותר המשפיעים על טביעת הרגל הפחמנית הכוללת של פרויקט בנייה הם שלבי הייצור CO₂-והלוגיסטיקה של חומרי הבניין והאלמנטים. החומרים המשמשים באתר משפיעים לעיתים קרובות על פליטות הכוללות, במיוחד בשלבים המוקדמים של מחזור חיי הפרויקט - מהפקת חומרי הגלם ועד ההובלה לאתר הבנייה.

כדי לחשב את הפליטות לפי קטגוריות או סוגי אלמנטים בנייה, יש צורך בשימוש בקואפיציוני הפחמן המייצגים את כמות הנוצרת כתוצאה מייצור חומרים שונים. בין החומרים הללו ניתן למנות בטון, לבנים, פלדה ממוחזרת, אלומיניום CO₂-ה (מלאי של) UK ICE 2015 ועוד. ערכים אלו נלקחים בדרך כלל ממקורות מהימנים ומבסיסי נתונים בינלאומיים, כגון (סוכנות ההגנה על הסביבה של ארה"ב). (בטבלה הבאה) איור 5.36 (מוצגים US EPA 2006-פחמן ואנרגיה ו) קואפיציוני הפליטות הבסיסיים עבור מספר חומרים בנייה נפוצים. עבור כל אחד מהם מצוינים שני פרמטרים מרכזיים, (הספציפיות) בקילוגרמים לקילוגרם חומר (וקואפיציוני המרה מנפח למסה) בקילוגרמים למטר מעוקב (CO₂-פליטות ה QTO הנדרשים לשילוב החישובים במודל התכנון ולחיבור לקבוצת הנתונים).



Carbon Emitted in Production		UK ICE Database (2015) USEPA (2006)	UK ICE Database (2015) USEPA (2006)	Coefficient m3 to kg
Material	Abbreviated	Process Emissions (kg CO2e/ kg of product) (K1)	Process Emissions (kg CO2e/ kg of product) (K2)	Kg / m3 (K3)
Concrete	Concrete	0.12	0.12	2400
Concrete block	Concrete block	0.13**	0.14	2000
Brick	Brick	0.24	0.32	2000
Medium density fiberboard (MDF)	MDF	0.39*	0.32	700
Recycled steel (avg recy content)	Recycled steel	0.47	0.81	7850
Glass (not including primary mfg.)	Glass	0.59	0.6	2500
Cement (Portland, masonry)	Cement	0.95	0.97	1440
Aluminum (virgin)	Aluminum	12.79	16.6	2700

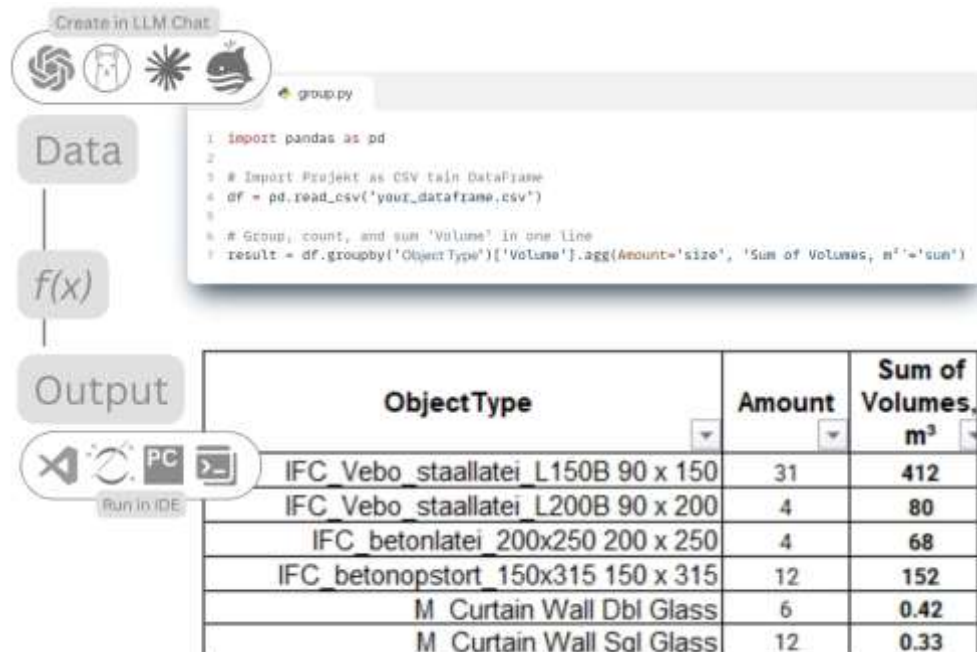
US EPA ו-UK ICE איור 5.36 כמות הפחמן המשתחררת במהלך ייצור חומרים בנייה שונים, בהתאם לבסיסי הנתונים.

יש לקבוע את נפח האטריבוטים, D-5 ודעבור הפרויקט, כמו במקרה של חישובי CO₂ 4 כדי לחשב את סך הפליטות של על ידי קבלת נפחי האטריבוטים, (QTO) של כל קבוצת אובייקטים. ניתן לעשות זאת באמצעות כלים לניתוח כמותי לאחר מכן, הנפחים שהתקבלו מוכפלים. Quantity take-off-במטרים מעוקבים, כפי שנדון בפירוט בפרק המוקדש לטכנולוגיות "של כל קבוצת חומרים CO₂ בקואפיציונים המתאימים לאטריבוט "פליטות

תוך כדי קיבוץ כל, CAD (BIM) בואו נחלץ אוטומטית את טבלת הנפחים לפי סוגי אלמנטים מתוך פרויקט LLM-הנתונים של הפרויקט, כפי שנעשה בפרקים קודמים. כדי לבצע משימה זו, נפנה ל

לפי פרמטר עמודת "שם CAD (BIM) מתוך פרויקט DataFrame-בבקשה, קבץ את טבלת ה אובייקט) "או "סוג ("והצג את כמות האלמנטים בכל קבוצה, כמו כן סכם את הפרמטר "נפח "עבור כל האלמנטים בסוג

LLM: תשובות ?

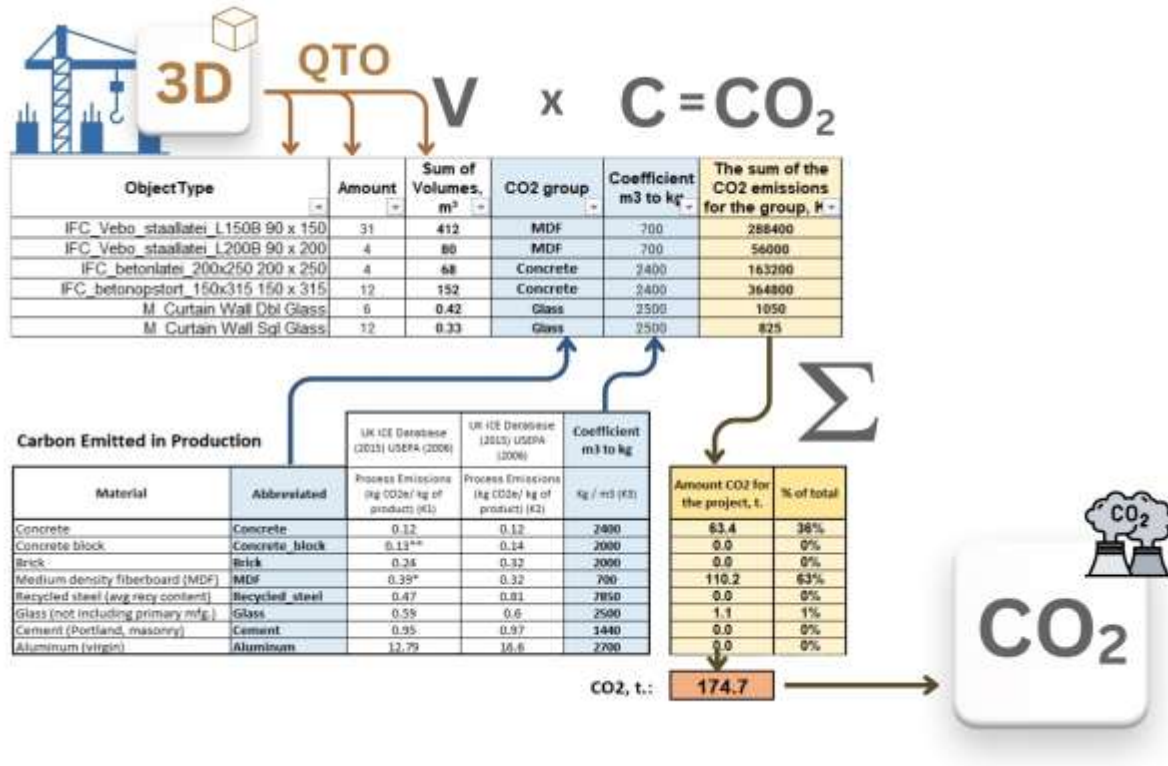


עם אטריבוט "נפח" מסוכם (ObjectType) קיבץ עבורנו את ישויות הפרויקט לפי סוג LLM-איור 5.37 הקוד שנוצר ב

עבור כל הפרויקט, מספיק להגדיר התאמה אוטומטית של הנתונים CO₂ כדי לאוטומט את חישוב הפליטות הכוללות של בטבלה או לקשר ידנית את סוגי האלמנטים (איור 5.37 (עם סוגי החומרים המתאימים) איור 5.36 (מתוך טבלת קואפיציוני הפליטות. את טבלת הקואפיציונים והנוסחאות המוכנות, כמו גם את הקוד לקבלת הנפחים מתוך פורמטים CAD (BIM) "CO₂-calculating-the-embodied-carbon. DataDrivenConstruction". לפי הבקשה GitHub-ניתן למצוא ב CO₂ ואוטומציה של הגדרת

מאפשרת לחשב אוטומטית את CAD של אלמנטים מתוך בסיס הנתונים QTO כך, אינטגרציית הנתונים לאחר קיבוץ פליטות הפחמן הדו-חמצני עבור אפשרויות תכנון שונות. זה מאפשר לנתח את השפעת החומרים השונים באופציות שונות לצורך קבלת תעודה מסוימת בעת מסירת CO₂-ולבחור רק את הפתרונות העומדים בדרישות הלקוח לגבי רמות פליטות ה-הבניין לשימוש.

באמצעות הכפלת מקדמים בנפחים של רכיבי פרויקט מקובצים - דוגמה טיפוסית CO₂ הערכת פליטות
עבור פרויקט (LEED למשל, הסמכת) ESG למשימה בתהליך קבלת דירוג

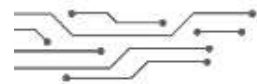


CO₂ מבטיחה דיוק ואוטומציה בהשגת הערכות נפחי פליטות CAD ממסדי נתונים QTO אינטגרציה קבוצות 5.38 אורן הסופיות.

באופן דומה, על ידי הגדרת נפחי קבוצות רכיבים, אנו יכולים לבצע חישובים לצורך בקרת ולוגיסטיקה חומרים, ניטור וניהול איכות, דימוי וניתוח צריכת אנרגיה ומגוון משימות נוספות להשגת סטטוס אטריבוטיבי חדש) פרמטר בטבלה (עבור קבוצות רכיבים נפרדות כמו גם עבור הפרויקט כולו.

כאשר מספר תהליכי חישוב כאלה בחברה מתחיל לגדול, עולה השאלה לגבי הצורך באוטומציה של חישובים כאלה והטמעת תוצאות החישובים בתהליכים ובמערכות ניהול הנתונים של החברה.

בשל המורכבות של פתרון כולל, חברות בינוניות וגדולות הפועלות בתעשיית הבנייה מעבירות אוטומציה כזו למיקור חברות המפתחות יוצרות עבור לקוחות גדולים מערכת מודולרית (PMIS או ERP) חוץ לחברות המפתחות מערכות כוללת אחת, לניהול שכבות מידע שונות רבות, כולל חישובי חומרים ומשאבים.



פרק 5.4. בבנייה ERP ו PMIS מערכות

בבנייה לדוגמת חישובים והערכות ERP מערכות

מודולריות מאחדות שכבות אטריבוטיות (מידע (שונות וזרמי נתונים למערכת כוללת אחת, המאפשרת ERP מערכות למנהלי פרויקטים לנהל באופן מסונכרן משאבים, כספים, לוגיסטיקה ואספקטים נוספים של הפרויקט במסגרת פלטפורמה לבנייה פועלת כ"מוח" של פרויקטי בנייה, מפשטת תהליכים חוזרים באמצעות אוטומציה, ומספקת ERP אחת. מערכת שקיפות ובקרה לאורך כל תהליך הבנייה.

לבנייה (תכנון משאבים ארגוניים (הן פתרונות תוכנה כוללים, המיועדים לניהול ERP מערכות לבנייה נמצאים מודולים לניהול ERP ואופטימיזציה של היבטים שונים בתהליך הבנייה. בליבת מערכות חישובי עלויות ולתכנון לוחות זמנים, מה שהופך אותן לכלי חשוב לתכנון משאבים יעיל.

מאפשרים למשתמשים להזין, לעבד ולנתח נתונים, תוך כיסוי מסודר של היבטים שונים של ERP מודולי מערכת הפרויקט, שיכולים לכלול ניהול עלויות חומר וכוח אדם, שימוש בציוד, ניהול לוגיסטיקה, משאבי אנוש, קשרים וסוגים נוספים של פעילויות בנייה.

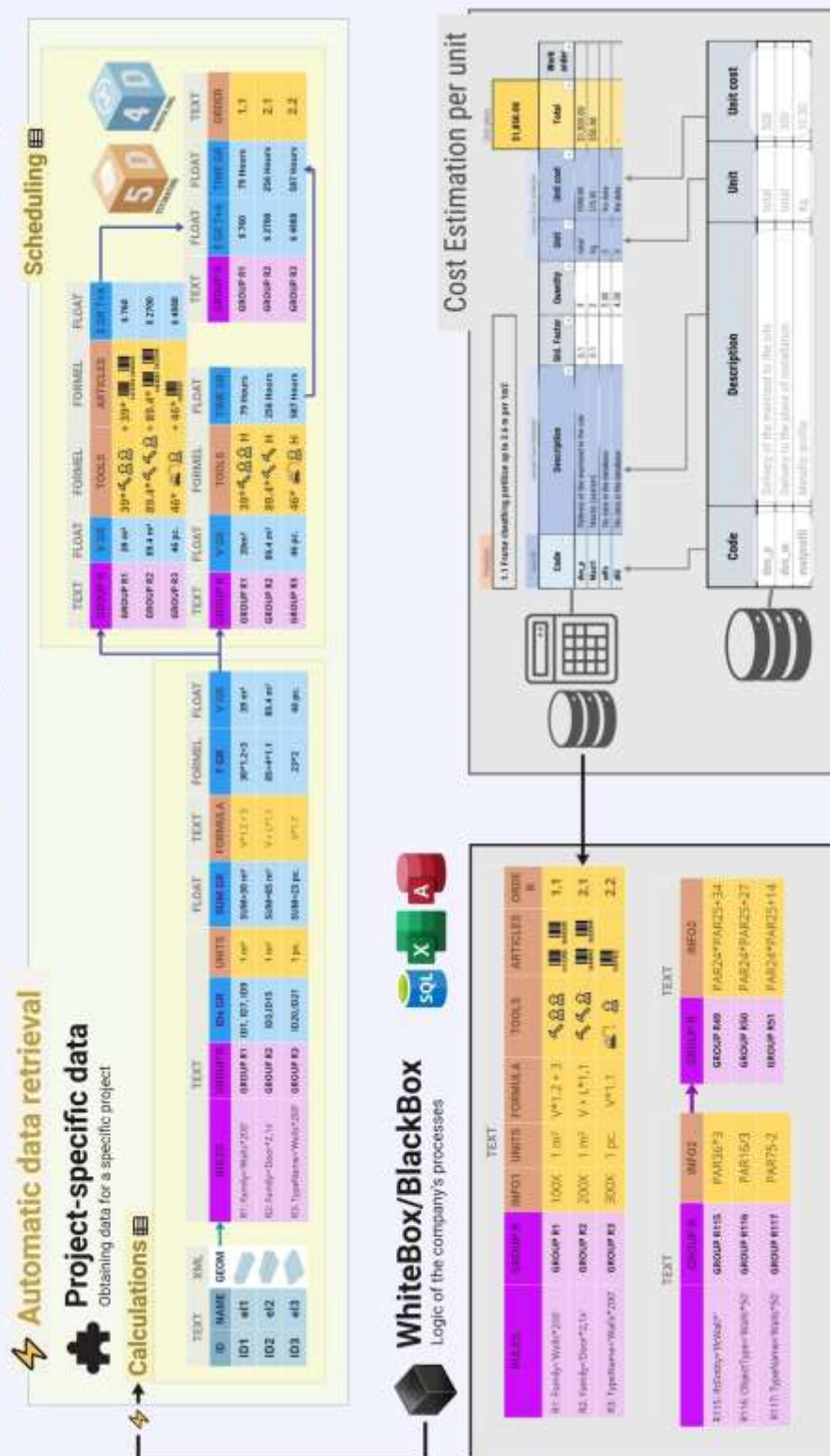
BlackBox/WhiteBox - אחד מהבלוקים הפונקציונליים של המערכת הוא מודול אוטומציה של לוגיקה עסקית, המהווה מרכז ניהול לתהליכים.

באמצעות זכויות גישה, לנהל באופן גמיש, ERP מאפשר למומחים המשתמשים במערכת BlackBox/WhiteBox המונחים, ERP היבטים שונים של העסק, שכבר הוגדרו מראש על ידי משתמשים או מנהלים אחרים. בהקשר של מערכות מציינים רמות שקיפות ובקרה על הלוגיקה הפנימית של המערכת WhiteBox ו-BlackBox

■ ("קופסה שחורה" - המשתמש אינטראקציה עם המערכת דרך ממשק, מבלי שיש לו גישה BlackBox ללוגיקה הפנימית של ביצוע התהליכים. המערכת מבצעת חישובים באופן עצמאי, על בסיס כללים שנקבעו מראש, המוסתרים מהמשתמש הסופי. הוא מזין נתונים ומקבל תוצאה, מבלי לדעת אילו אטריבוטים או מקדמים שימשו בפנים.

■ («קופסה לבנה» - הלוגיקה של התהליכים זמינה לצפייה, הגדרה ושינוי. משתמשים WhiteBox מתקדמים, מנהלים או אינטגרטורים יכולים לקבוע ידנית את האלגוריתמים לעיבוד נתונים, כללי חישוב ותסריטים לאינטראקציה בין ישויות הפרויקט.

Enterprise Resource Planning ERP



לבנייה, לקבלת אומדנים ולוחות זמנים בעבודה עם הזנת נתונים ידנית של מאפייני הכמות **ERP** ארכיטקטורת מערכת **5.41**. ריש

דוגמה לכך יכולה להיות מצב שבו משתמש מנוסה או מנהל קובע כלל: אילו מאפיינים באומדן צריכים להיות מוכפלים, זה בזה או מקובצים לפי קריטריון מסוים, ואיפה יש לרשום את התוצאה הסופית. בהמשך, אנשי מקצוע פחות מיומנים דרך ממשק המשתמש – ומקבלים אומדנים, לוחות זמנים או ERP-כמו מהנדסי אומדן, פשוט מעלים נתונים חדשים ל מפרטים מוכנים מבלי צורך לכתוב קוד או להבין בפרטים הטכניים של הלוגיקה

בסביבת מערכות LLM. בפרקים הקודמים נבחנו מודולי חישוב ולוגיקה בהקשר של אינטראקציה עם חישובים ושינויים כאלה מתבצעים בתוך מודולים, המוסתרים מאחורי ממשק של כפתורים וטפסים ERP

קובע כללים להתאמת מאפיינים BlackBox/WhiteBox במודול ERP מנהל מערכת (5.41. ריש) בדוגמה הבאה בזכות המודול המוגדר הזה) על ידי מנהל או מנהל מערכת (משתמש QTO של ישויות מהאומדן עם מאפיינים לקיבוץ מקבל אוטומטית, ERP) מהנדס אומדן או מהנדס, (שמוסיף ידנית מאפיין של כמות או נפח דרך ממשק המשתמש של אומדנים מוכנים ולוחות עבודה. כך תהליכי החישוב וההכנה של האומדנים, שנדונו בפרקים הקודמים בעזרת קוד, הופכים לקו ייצור חצי אוטומטי ERP-בתוך ה

דרך, למשל, העלאת פרויקט, (4.113. ריש) CAD (BIM) חיבור תהליך חצי-אוטומטי זה למאפייני הכמות מהמודלים המוגדר מראש לכך, הופך את זרם הנתונים למנגנון מסונכרן, המסוגל לעדכן באופן עצמאי ומיידי ERP למודול, CAD את העלות של קבוצות רכיבים בודדים או של כל הפרויקט בתגובה לכל שינוי בו בשלב התכנון, כאשר מודל -ERP-מועלה ל

יש להגדיר בצורה מסודרת את ERP-ל CAD (BIM) בין מערכות (5.42. ריש) כדי ליצור זרם נתונים אוטומטי על כך כבר דיברנו בפרק, CAD (BIM) התהליכים העיקריים ואת הדרישות לנתונים מתוך בסיסי הנתונים של מודלי -מחולק לשלבים דומים ERP-הקודם "דרישות והבטחת איכות הנתונים". תהליך זה ב

■ כללי. ERP. יצירת כללי בדיקה 1), (המשחקים תפקיד חשוב בהבטחת דיוק הנתונים המגיעים למערכת הבדיקה משמשים כמסננים, הבדקים את הישויות ואת המאפיינים שלהן, ומעבירים למערכת רק את האלמנטים שעברו את הדרישות. פרטים נוספים על בדיקה ואימות בפרק "יצירת דרישות ובדיקת איכות הנתונים".

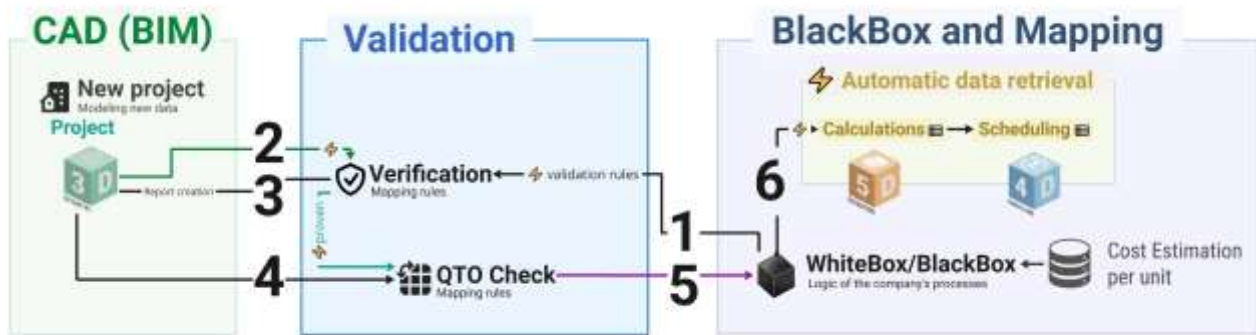
■ מתבצע תהליך אימות 2), (המאשר שכל האלמנטים-ישויות של הפרויקט עם ERP-לאחר מכן בתוך ה המאפיינים והערכים שלהם נוצרו כראוי ומוכנים לשלבים הבאים של העיבוד

■ אם מתעוררות בעיות עם נתונים מאפיינים חסרים, נוצר דוח 3), (והפרויקט יחד עם הוראות לתיקון נשלח לשיפור עד שיהיה מוכן ליטריציה הבאה

■ ליצירת טבלאות כמות 4) ERP לאחר שאושרו ובדקו את נתוני הפרויקט, הם משמשים במודול אחר של מייצרות את תכונות הכמות עבור (WhiteBox/BlackBox) אשר לפי כללים שנוצרו מראש, (QTO) קבוצות ישויות, חומרים ומשאבים

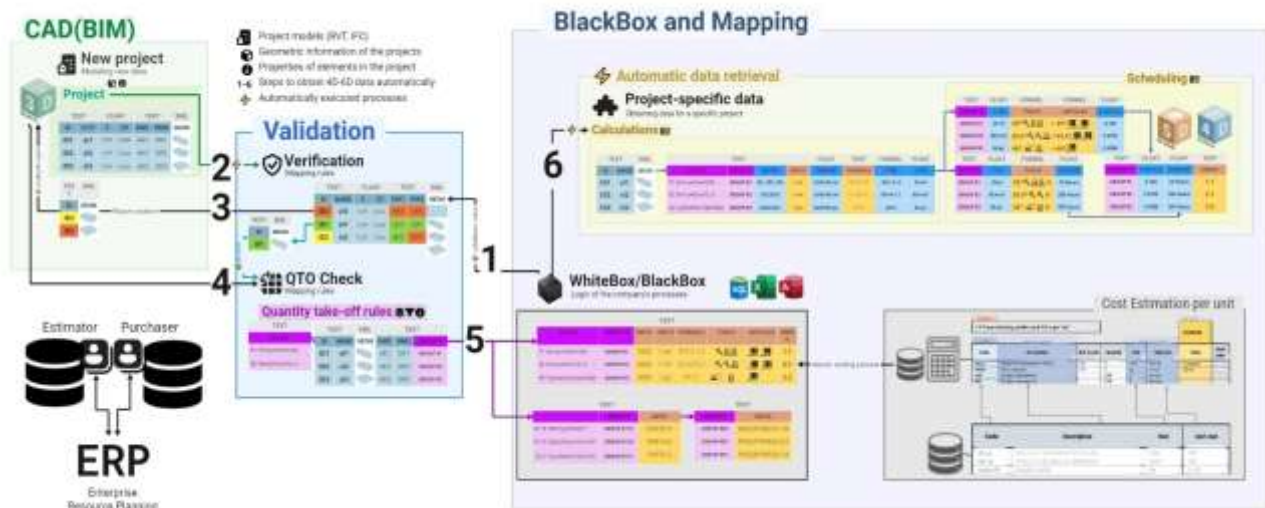
■ 5) משתלבים אוטומטית עם החישובים) למשל, עלויות וזמן QTO הנתונים הממוינים לפי כללי ההתאמה או

■ עם תכונות טבלאות QTO-המשתמש, מכפיל את תכונות הנפח מטבלת ה, ERP-בשלב האחרון, מערכת ה התהליכים) למשל, פריטי אומדן, (ומייצרת אוטומטית את תוצאות החישובים 6) (למשל, אומדני עלויות עבור כל קבוצת ישויות ועבור הפרויקט כולו CO₂) לוחות זמנים או פליטות

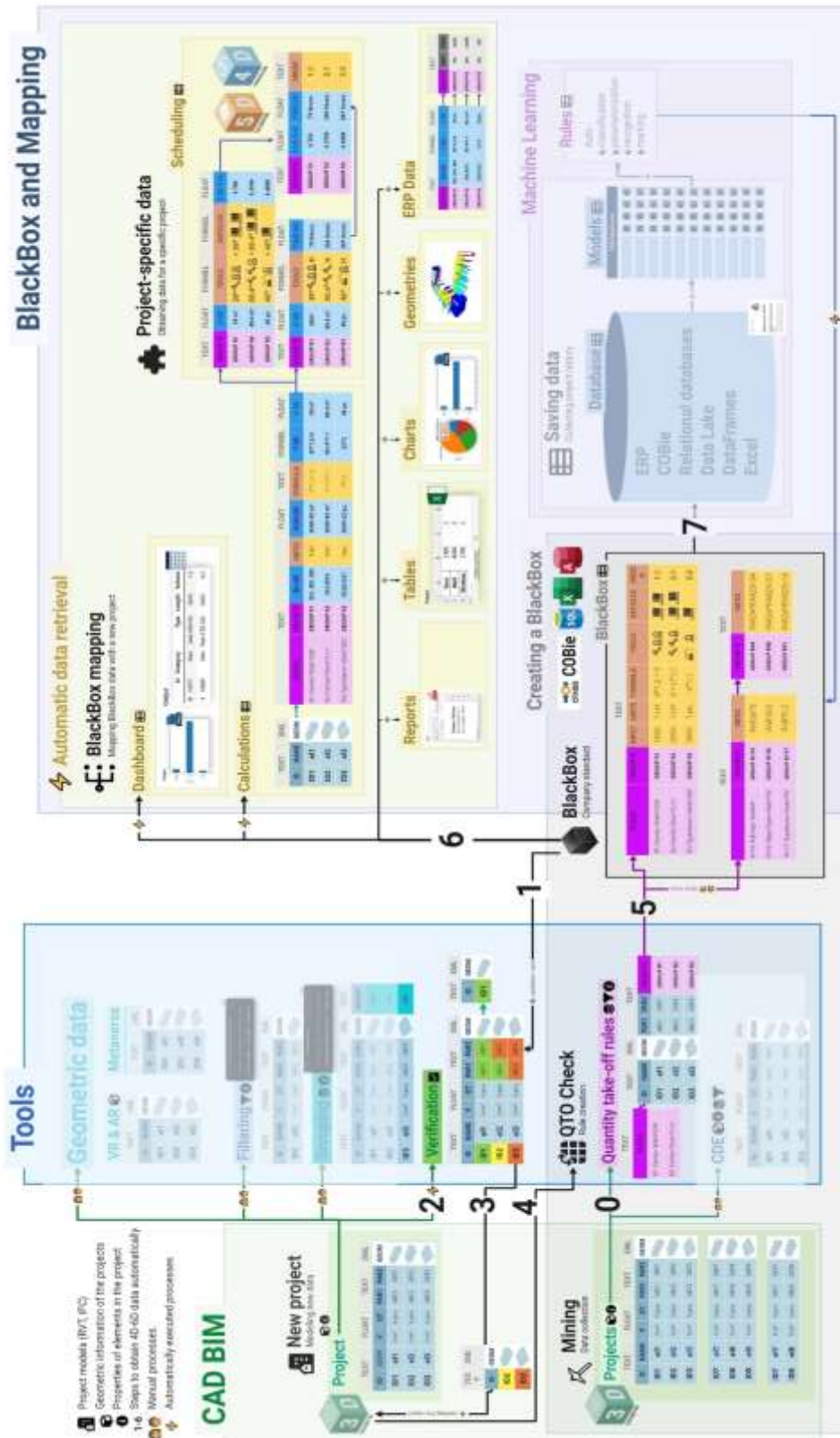


מהקמת כללי האימות (1) ועד חישוב אוטומטי של עלויות, CAD (BIM), לבנייה עם ERP ארכיטקטורת מערכת 5.42. ריש (5-6) ולוחות זמנים.

מודולרית, התהליכים משתלבים באמצעות תוכנה הכוללת ממשק משתמש. מאחורי הממשק נמצא החלק ERP במערכת הפנימי, שבו טבלאות מובנות מעבדות את הנתונים, מבצעות פעולות שונות שהוגדרו מראש על ידי מנהל או מנהל מערכת מקבל, (BlackBox/WhiteBox במודולים) כתוצאה מכך, המשתמש, בזכות הלוגיקה האוטומטית שהוגדרה מראש מסמכים מוכנים חצי-אוטומטית העונים על משימותיו.



מסייעת למנהלים ולמשתמשים לנוע בין טבלאות המומחים כדי לייצר נתונים חדשים ERP מערכת 5.43. ריש.



משולבת עם כלים אנליטיים ומאיצה את תהליך קבלת ההחלטות בחברה ERP מערכת 5.44. ריש

מהווים שרשרת של (5.43. שלבים 1-6 ריש) מההתחלה ועד החישוב הסופי, ERP באופן דומה, התהליכים במערכות

- צעדים מקושרים, אשר בסופו של דבר מבטיחים שקיפות, יעילות ודייקנות בתכנון

לבנייה מודרניות כוללות לא רק מודולים לחישוב עלויות וזמנים, אלא גם עשרות מודולים מוכנים מראש ERP מערכות בדרך כלל מכסים פונקציות של ניהול מסמכים, מעקב אחר התקדמות הפרויקט, ניהול חוזים, שרשראות אספקה מאפשרים ERP-ולוגיסטיקה, כמו גם אינטגרציה עם מערכות עסקיות ופלטפורמות אחרות. הכלים האנליטיים המובנים ב (KPI - key performance indicators) למשתמשים לאוטומט את יצירת לוחות מחוונים למעקב אחר מדדי ביצוע מרכזיים של הפרויקט. זה מבטיח ניהול מרכזי ועקבי של כל ההיבטים של פרויקט הבנייה, בניסיון לאחד כמות (indicators) גדולה של יישומים ומערכות על פלטפורמה אחת

תשתמש בשילוב עם למידת מכונה כדי לשפר את הדיוק ולייעל את תהליך חישוב תכונות ERP בעתיד, אנליטיקה יהפכו בעתיד, (ריש) 5.44 Big Data ב-ERP הפרויקט העתידיות. הנתונים והתכונות, שניתחו ואספו ממערכות לבסיס ליצירת מודלים חיזויים, שיכולים לחזות במדויק עיכובים פוטנציאליים, סיכונים או, למשל, שינויים אפשריים - בעלויות החומרים

- (מערכת מידע לניהול פרויקטים) PMIS-בתעשיית הבנייה לעיתים קרובות נעשה שימוש ב, ERP-כאלטרנטיבה ל מערכת לניהול פרויקטים, המיועדת לשליטה מפורטת על ביצוע המשימות ברמת פרויקט בנייה ספציפי

לאחר הבנייה ERP הקשר בין PMIS:

מתמקדת בניהול פרויקט ספציפי PMIS, אשר מכסה את כל שרשרת תהליכי העסקים של החברה, ERP-בניגוד ל ומספקת מעקב אחר לוחות זמנים, תקציב, משאבים ודוקומנטציה

(מערכת מידע לניהול פרויקטים) היא תוכנה לניהול פרויקטים בבנייה, המיועדת לתכנון, מעקב PMIS, ניתוח ודיווח על כל ההיבטים של הפרויקט

עשויה להיראות כפתרון כפול ביחס PMIS, מאפשרת לנהל דוקומנטים, לוחות זמנים, תקציבים ובמבט ראשון PMIS אך ההבדל המרכזי טמון ברמת הניהול, ERP-ל

■ מתמקדת בתהליכי העסקים של החברה בכללותה: ניהול עלויות, חוזים, רכש, כוח אדם ומשאבים ברמה ERP הארגונית.

■ מתמקדת בניהול פרויקטים בודדים, ומספקת תכנון מפורט, בקרת שינויים, ניהול דיווחים ותיאום בין PMIS המשתתפים

הופכת לשאלה של נוחות והעדפות PMIS כבר כוללות פונקציות מספקות, והטמעת ERP במקרים רבים, מערכות לא משום שהיא הכרחית, אלא משום שהיא נכפתה עליהם על ידי PMIS-החברה. קבלנים ורוכשים רבים משתמשים ב ספק או לקוח גדול, שמעוניין לאגד נתונים על פלטפורמה מסוימת

ניהול) PLM יש לציין כי במונחים בינלאומיים לניהול פרויקטים בבנייה קיימות גם מושגים פופולריים נוספים, כמו (ניהול הנדסה, רכש ובנייה -) שיטות התקשרות בתעשיית הבנייה EPC-M ו-EPC מחזור חיי המוצר (וכן

עשויה להתגלות כקישור מיותר PMIS עם מודולים לניהול פרויקטים, הטמעת ERP-אם בחברה כבר משתמשים ב עשויה להיות כלי נוח וקל יותר PMIS, המעתיק פונקציות. עם זאת, אם התהליכים אינם אוטומטיים והנתונים מפוזרים לתחזוקה

PMIS-ו-ERP-ספקולציות, רווח, סגירות וחוסר שקיפות ב

בבנייה ברוב המקרים מהוות פתרונות PMIS ו-ERP למרות הפשטות החיצונית של הממשקים והנהלים, מערכות סגורים ולא גמישים. מערכות אלו בדרך כלל מסופקות כקבוצת תוכנה מותאמת מראש מספק אחד, עם גישה מוגבלת לבסיסי נתונים פנימיים וללוגיקת התהליכים.

שכן בבסיסי הנתונים, CAD-(BIM) את הפיתוח והבקרה על מערכות אלו לוקחים על עצמם יותר ויותר ספקי מוצרים מאפיינים כמותיים ונפחיים של רכיבי הפרויקט. עם זאת, במקום לספק ERP: שלהם מצויה המידע הנדרש למערכות גישה לנתונים אלו בפורמט פתוח או קריא על ידי מכונה, ספקים מציעים רק תרחישי משתמש מוגבלים ולוגיקת עיבוד זה מפחית את גמישות המערכת ומקשה על התאמתה לתנאים. BlackBox סגורה - שהוגדרה מראש בתוך מודולי הספציפיים של הפרויקטים.

שקיפות מוגבלת של נתונים נותרת אחת הבעיות המרכזיות בתהליכים הדיגיטליים בבנייה. ארכיטקטורה סגורה של מסדי נתונים, חוסר גישה למערכות נתונים מלאות של רכיבי הבנייה, התמקדות במודולים וחוסר ממשקים פתוחים מגבירים משמעותית את הסיכונים של ביורוקרטיה BlackBox אוטומטיים מסוג תיעודית. מגבלות אלו יוצרות "צווארי בקבוק" בתהליך קבלת ההחלטות, מקשות על אימות המידע משתמשים בדרך כלל. ERP/PMIS ופותחות אפשרויות להסתתרות נתונים או ספקולציות בתוך מערכות חלקי - ללא אפשרות לתקשר עם API מקבלים גישה מוגבלת בלבד - בין אם מדובר בממשק מצומצם או CAD, מקורות הנתונים ישירות. זה במיוחד קריטי כאשר מדובר בפרמטרים שנוצרו אוטומטית מפרויקטי QTO. כמו כמיות, שטחים ומספרים, המשמשים לחישובי.

כתוצאה מכך, במקום לחפש יעילות דרך אוטומציה של תהליכים, שקיפות נתונים, צמצום עלויות עסקה ויצירת מודלים עסקיים חדשים, רבות מחברות הבנייה מתמקדות בניהול פרמטרים חיצוניים - מניפולציה של מקדמים, גורמים מתאימים זה יוצר קרקע לספקולציות, מעוות ERP/PMIS ושיטות חישוב המשפיעות על עלות הפרויקט בפלטפורמות סגורות את העלויות האמיתיות של הייצור ומפחית את האמון בין כל המשתתפים בתהליך הבנייה.

בבנייה, הרווח נוצר כהפרש בין ההכנסות מפרויקט שהושלם לבין העלויות המשתנות, הכוללות תכנון חומרים, משאבי עבודה והוצאות ישירות אחרות הקשורות ישירות למימוש הפרויקט. עם זאת, גורם מרכזי המשפיע על גובה העלויות הללו הוא לא רק הטכנולוגיה או הלוגיסטיקה, אלא גם מהירות ודיוק החישובים, כמו גם איכות ההחלטות הניהוליות בתוך החברה.

הבעיה מחמירה מכיוון שבמרבית חברות הבנייה תהליכי חישוב העלויות נותרו לא שקופים לא רק עבור הלקוחות, אלא גם עבור העובדים עצמם שאינם חלק ממחלקות האומדן או הכספים. סגירות זו תורמת להיווצרות קבוצת מומחים מועדפת בתוך החברה - נושאי "מומחיות פיננסית", המחזיקים בזכות הבלעדית לערוך את המאפיינים ואת המקדמים המתאימים. עובדים אלו, יחד עם ראשי החברות, יכולים למעשה לשלוט בלוגיקה הפיננסית של הפרויקט. ERP/PMIS במערכות

במצבים כאלה, המעריכים הופכים ל"להטוטנים פיננסיים", מאזנים בין מקסום הרווח של החברה לבין הצורך לשמור על מחיר תחרותי ללקוח. תוך כדי כך, הם נאלצים להימנע ממניפולציות גלויות וברוטליות כדי לא לפגוע במוניטין של החברה. בשלב זה בדיוק נקבעים המקדמים, המסתירים כמויות או עלויות מוגזמות של חומרים ועבודות.

כתוצאה מכך, התוכנית העיקרית לשיפור היעילות והרווחיות של חברות הפועלות בתעשיית הבנייה אינה אוטומציה והאצת תהליכי קבלת החלטות, אלא ספקולציה על מחירי חומרים ועבודות. העלאת עלויות העבודות והחומרים מתבצעת סגורות באמצעות העלאת אחוזים מעל המחירים בשוק או על ידי ERP/PMIS על ידי "חשבוניות אפורה" במערכות הגדלת כמויות העבודות בעזרת מקדמים, אשר נדונו בפרק "הכנת חישובים וחישוב עלויות עבודות על בסיס בסיס המשאבים".

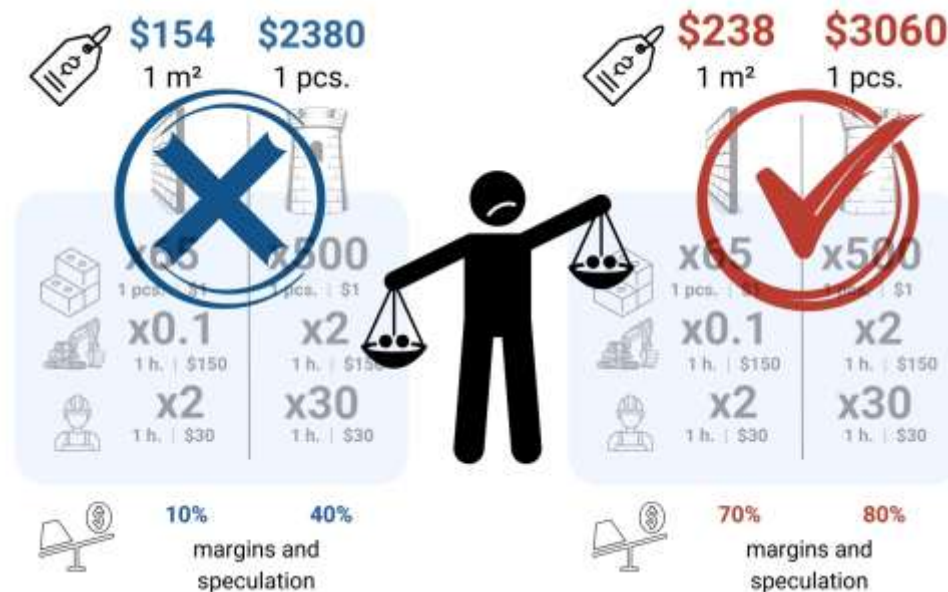
כתוצאה מכך, המזמין מקבל חישוב שאינו משקף את העלות או הכמויות האמיתיות של העבודות, אלא מהווה נגזרת ממספר מקדמים פנימיים סמויים. במקביל, קבלני המשנה, בניסיון להתאים את עצמם למחירים המופחתים שהציב הקבלן הראשי, לעיתים קרובות נאלצים לרכוש חומרים זולים ואיכותיים פחות, דבר שמחמיר את איכות הבנייה הסופית.

התהליך הספקולטיבי של חיפוש רווחים באוויר בסופו של דבר פוגע הן בלקוחות, שמקבלים נתונים לא מדויקים, והן בביצועיסטים, שנאלצים לחפש מודלים חדשים של ספקולציות.

כתוצאה מכך, ככל שהפרויקט גדול יותר, כך רמת הביורוקרטיה בניהול הנתונים והתהליכים גבוהה יותר. מאחורי כל שלב וכל מודול לעיתים קרובות מסתתרים מקדמים ותוספות לא שקופות, המוטמעות באלגוריתמים החישוביים ובנהלים הפנימיים. זה לא רק מקשה על הביקורת, אלא גם מעוות באופן משמעותי את התמונה הפיננסית של הפרויקט. בפרויקטים גדולים בתחום הבנייה, פרקטיקות כאלה לעיתים קרובות מביאות להגדלה מרובת (לפעמים עד עשרה פעמים) של העלות. - הסופית, כאשר הכמויות וההוצאות האמיתיות נשארות מחוץ לשליטת המזמין.

לפי דוח של מקינזי &חברה "מדמינים את העתיד הדיגיטלי של הבנייה" (2016), (פרויקטים גדולים בבנייה מסתיימים במוצע 20% מאוחר יותר מהתאריך המתוכנן ומעלים את התקציב עד 80%.

מחלקות האומדן והתקצוב הופכות להיות החוליה המוגנת ביותר בתוך החברה. הגישה אליהן מוגבלת באופן קפדני אפילו למומחים פנימיים, ובשל סגירות הלוגיקה והמבנים של בסיסי הנתונים, בלתי אפשרי להעריך באופן אובייקטיבי את היעילות של החלטות פרויקט ללא עיוותים. חוסר שקיפות מביא לכך שהחברות נאלצות לא ליעל תהליכים, אלא להיאבק על הישרדות באמצעות ניהול "יצירתי" של מספרים ומקדמים.



- מקדמים ספקולטיביים ברמת החישובים הם הרווח העיקרי של החברות ואמנות הלהטטים בין איכות העבודה לבין המוניטין

סגורות בבנייה. בתנאים של טרנספורמציה ERP/PMIS כל זה מעורר ספק לגבי הכדאיות של המשך השימוש במערכות דיגיטליות ודרישות הולכות וגדלות לשקיפות מצד המזמינים, סביר להניח שהיישום של פרויקטים בטווח הארוך לא יישאר תלוי בפתרונות קנייניים, המגבילים את הגמישות, מפריעים לאינטגרציה ומעכבים את התפתחות העסק

- וכמה שזה עשוי להיות משתלם לחברות הבנייה לעבוד עם סילו נתונים ונתונים לא שקופים בבסיסי נתונים סגורים העתיד של תעשיית הבנייה יהיה בלתי נמנע קשור למעבר לפלטפורמות פתוחות, מבנים נתונים קריאים על ידי מכונה, ושקופים, כמו גם אוטומציה המבוססת על עקרונות אמון. השינוי הזה ינוע "מלמעלה" - תחת הלחץ של לקוחות רגולטורים והחברה, שדורשים יותר ויותר שקיפות, קיימות ואחריות כלכלית

הסגורים: תעשיית הבנייה זקוקה לגישות חדשות ERP/PMIS-סוף עידן ה

מודולריות מסורבלות, המורכבות מעשרות מיליוני שורות קוד, מקשה מאוד על כל ERP/PMIS השימוש במערכות שינוי בהן. במקביל, המעבר לפלטפורמה חדשה בתנאים של מודולים שכבר מותאמים מראש לחברה, עשרות אלפי פריטים בבסיסי המשאבים) איור 5.13 (ואלפי חישובים מוכנים) איור 5.16 (הופך לתהליך יקר וממושך. ככל שיש יותר קוד ופתרונות ארכיטקטוניים מיושנים - כך עולה רמת חוסר היעילות הפנימית, וכל פרויקט חדש רק מחמיר את המצב. בחברות רבות, העברת נתונים ואינטגרציה של פתרונות חדשים הופכים לאפופיות שנמשכות שנים, מלוות בשיפורים מתמשכים וחיפוש אינסופי אחרי פשרות. כתוצאה מכך, לעיתים קרובות מתבצע חזרה לפלטפורמות ישנות --ומוכרות, למרות המגבלות שלהן

כפי שמודגש בדו"ח הגרמני "הספר השחור" "גרמניה" [108], [המוקדש לכשלים מערכתיים בניהול נתוני הבנייה, פיצול המידע וחוסר גישה מרכזית לניהולו - הם הסיבה המרכזית לירידת היעילות. ללא סטנדרטיזציה ואינטגרציה, הנתונים מאבדים את ערכם, והופכים לארכיון ולא לכלי ניהול

הסיבה העיקרית לאובדן איכות הנתונים היא תכנון ובקרה לא מספקים של פרויקטי הבנייה, מה שלעתים קרובות מביא לעלייה משמעותית בעלויות. בפרק "במרכז תשומת הלב: התפוצצות העלויות" של "הספר השחור" ניתוחים גורמים מרכזיים התורמים לתוצאות לא רצויות אלו. בין היתר - ניתוח לא מספק של הצרכים, חוסר הוכחות טכניות-כלכליות ותכנון לא מתואם, המוביל להוצאות נוספות שניתן היה להימנע מהן.

בשלה, החלפת מערכת מיושנת דומה להחלפת עמוד תומך בבניין שכבר נבנה. לא IT במערכת אקולוגית מספיק פשוט לפרק את הישנה ולהתקין את החדשה - חשוב לעשות זאת כך שהבניין יישאר יציב, התקרות לא יקרסו וכל התקשורת תמשיך לפעול. כאן טמונה הקושי: כל טעות עלולה להוביל לתוצאות חמורות עבור כל המערכת של החברה.

גדולים עבור תעשיית הבנייה ממשיכים להשתמש בכמות הקוד שנכתבה כטיעון בעד ERP עם זאת, מפתחי מוצרי "הפלטפורמה שלהם. בכנסים מיוחדים עדיין ניתן לשמוע ביטויים כמו: "כדי לשחזר מערכת כזו יידרשו 150 שנות אדם למרות שבמרבית הפונקציות של מערכות דומות מסתתרות מסדי נתונים ופונקציות פשוטות יחסית לעבודה עם טבלאות ארוזות בממשק משתמש קבוע ומיוחד. בפועל, כמות הקוד של "150 שנות אדם" הופכת יותר למעמסה מאשר ליתרון תחרותי. ככל שיש יותר קוד - כך עולה עלות התמיכה, קשה יותר להתאים לתנאים חדשים ועולה סף הכניסה למפתחים חדשים וללקוחות.

היום רבות ממערכות הבנייה המודולריות מזכירות מבנים מסורבלים ומיושנים, שבהם כל שינוי לא זהיר labyrinth, עלול להוביל לתקלות. כל מודול חדש רק מסבך את המערכת שכבר עמוסה, והופך אותה ל המובן רק לקבוצה מצומצמת של מומחים, מה שהופך אותה לקשה עוד יותר לתמיכה ולשדרוג.

המורכבות מוכרת גם על ידי המפתחים עצמם, אשר לעיתים קרובות עושים הפסקות לצורך ריפקטורינג - סקירה מחדש של הארכיטקטורה בהתחשב בהופעת טכנולוגיות חדשות. עם זאת, גם אם הריפקטורינג מתבצע באופן קבוע, המורכבות בלתי נמנעת הולכת וגדלה. האדריכלים של מערכות כאלה מתרגלים למורכבות ההולכת ומתרחקת, אך עבור משתמשים חדשים ומומחים זה הופך להיות מחסום בלתי עביר. כתוצאה מכך, כל המומחיות מתרכזת בידיהם של מספר מפתחים והמערכת מפסיקה להיות ניתנת להרחבה. בטווח הקצר, מומחים כאלה מועילים, אך בטווח הארוך - הם הופכים לחלק מהבעיה.

ארגונים ימשיכו לשלב "נתונים קטנים" עם המקבילים הגדולים שלהם, וטיפש הוא מי שמאמין שיישום אחד - לא משנה כמה יקר או אמין הוא - יכול להתמודד עם הכל. פיל סיימון, מנחה הפודקאסט שיחות על שיתוף פעולה

מתעורר שאלה הגיונית: האם באמת אנו זקוקים למערכות כה מסורבלות וסגורות לחישוב עלויות וזמנים של עבודות בצורה של טבלאות, כאשר בתחומים אחרים עם משימות דומות כבר מתמודדים עם כלים אנליטיים עם נתונים פתוחים ולוגיקה שקופה?

נכון לעכשיו, פלטפורמות מודולריות סגורות עדיין מבוקשות בתעשיית הבנייה, בראש ובראשונה בשל המאפיינים של וט "אפורות" או לא שקופות, ומאפשרות cxem ניהול החשבונאות. מערכות כאלה משמשות לעיתים קרובות לניהול להסתיר עלויות אמיתיות מהלקוח. עם זאת, ככל שהביקוש לדיגיטליזציה גובר, בראש ובראשונה מצד הלקוחות, וככל שלהן, יאבדו את ERP-שהתעשייה מתקדמת לעידן ה"אובריזציה", המתווכים, כלומר חברות הבנייה עם מערכות ה

חשיבותן בחישובי הזמן והעלות. זה ישנה באופן קבוע את פני תעשיית הבנייה. פרטים נוספים ניתן למצוא בחלק האחרון "של הספר ובפרק" בנייה 5.0: כיצד להרוויח כאשר כבר אי אפשר להסתיר.

אלפי פתרונות ישנים שהצטברו במהלך 30 השנים האחרונות עם אלפי שנות אדם שהושקעו בפיתוח יתחילו להיעלם במהירות. המעבר לניהול נתונים פתוח, שקוף וגמיש הוא בלתי נמנע. השאלה היא אילו חברות יצליחו להסתגל לשינויים הללו ואילו יישארו שבויות במודל הישן

הנתונים שלהם ממלאים כיום את הפרמטרים המאסיביים של ישויות, CAD (BIM) מצב דומה מתרחש גם בתחום הכלים (שפותח כבר בשנת 2002) (התבסס על קונספט של בסיס BIM הרעיון המקורי של ERP/PMIS פרויקט במערכות כיום דורשת סט שלם של תוכנות ופורמטים מיוחדים. מה שהיה BIM נתונים אינטגרטיבי אחד, אך בפועל, עבודה עם אמור לפשט את התכנון וניהול הבנייה הפך לשכבת פתרונות קנייניים נוספת, שמסבכת את האינטגרציה ומפחיתה את הגמישות של העסק.

צעדים נוספים: שימוש יעיל בנתוני פרויקטים

בחלק זה הראינו כיצד נתונים מובנים הופכים לבסיס לחישובים מדויקים של עלויות ולוחות זמנים של פרויקטים תכנון לוח זמנים וחשוב עלויות מפחיתה את העומס העבודה ומעלה באופן QTO בנייה. אוטומציה של תהליכי משמעותי את דיוק התוצאות

לסיכום חלק זה, כדאי להדגיש את הצעדים המעשיים העיקריים שיכולים לעזור ליישם את הגישות שנדונו במשימות היומיומיות שלכם. גישות אלו הן אוניברסליות - הן מועילות הן עבור טרנספורמציה דיגיטלית של החברה והן עבור העבודה היומיומית של אנשי מקצוע העוסקים בחישובים

■ אוטומטו חישובים שגרתיים

- ☐ נסו למצוא חישובים סטנדרטיים של עבודות שיכולים להיות רלוונטיים לעבודה שלכם
- ☐ -נתחו באילו שיטות מחושבות או נמדדות עבודות או תהליכים באתר הבנייה במדינתכם
- ☐ QTO חקרו את הפונקציות של חילוץ אוטומטי של מפרטים ונתוני - CAD אם אתם עובדים עם מערכות שלכם CAD (BIM) בתוכנת

☐ לכתיבת טיוטת קוד לאוטומציה של חישובים LLM-השתמשו ב

■ QTO פתחו כלים משלכם עבור

- ☐ צרו סקריפטים או גיליונות אקסל לאוטומציה של חישובי כמויות
- ☐ סטנדרטיזו קטגוריות וקבוצות של רכיבים לגישה עקבית להערכה
- ☐ תעדו את המתודולוגיה של החישובים כדי להבטיח את השחזור של התוצאות בפרויקטים חדשים

■ אינטגררו היבטים שונים של הפרויקט בעבודתכם

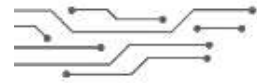
- ☐ אם אתם עובדים עם מערכות מודולריות, נסו להמחיש את התהליכים שלכם לא רק כסכמות או דיאגרמות, אלא גם ברמת הנתונים - במיוחד בצורה של טבלאות

☐ תוך, Python-עם חישובים - באמצעות קוד ב CAD שלטו באיחוד אוטומטי של נתונים שהופקו ממסדי שימוש בקיבוצים, סינון ואגרגציה

☐ כדי להציג מידע מורכב לעמיתים וללקוחות QTO צרו ויזואליזציות ברורות של קבוצות

צעדים אלו יסייעו לבנות מערכת חישובים יציבה, המבוססת על אוטומציה וסטנדרטיזציה של נתונים. גישה זו תעלה את הדיוק, תצמצם את השגרה בשאלות יומיומיות הקשורות לחישובים

עדיין קשה CAD ולסיבות לכך שמסדי נתונים CAD (BIM) הפרקים הבאים מוקדשים להיבטים טכניים של מוצרים בבנייה, בהתפתחות כלי BIM לשלב בתהליכי העבודה של חברות. אם אינכם מעוניינים כעת בהיסטוריה של יישום ובמאפיינים הטכניים של עבודה עם טכנולוגיות אלו, אתם יכולים לעבור מיד לחלק השביעי של הספר "קבלת CAD". החלטות מבוססות נתונים



מקסימום נוחות עם הגרסה המודפסת

לשימוש נוח יותר וגישה מהירה. Data-Driven Construction אתם מחזיקים בידיכם את הגרסה הדיגיטלית החינמית של לחומרים, אנו ממליצים לשקול את הגרסה המודפסת



תמיד בהישג

יד: ספר

בפורמט מודפס

יהפוך לכלי

עבודה אמין

המאפשר

למצוא

ולהשתמש

במהירות

בויזואליזציות

ובתרשימים

בכל מצב

עבודה

איכות גבוהה

של איורים: כל

התמונות

והגרפים

בגרסה

המודפסת

מוצגים

באיכות

מקסימלית

גישה מהירה

למידע: ניווט

נוח, אפשרות לעשות הערות, סימניות ולעבוד עם הספר בכל מקום.

ברכישת הגרסה המודפסת המלאה של הספר, אתם מקבלים כלי נוח לעבודה נוחה ויעילה עם מידע: אפשרות להשתמש בחומרים ויזואליים במשימות יומיומיות, למצוא במהירות את התרשימים הנדרשים ולעשות הערות. בנוסף, רכישתכם תומכת בהפצת ידע פתוח.

datadrivenconstruction.io/books: ניתן להזמין את הגרסה המודפסת של הספר בכתובת



VI חלק

ובימ: שיווק, מציאות ועתיד נתוני הפרויקטים בבנייה CAD

BIM ו-CAD החלק השישי של הספר מציע ניתוח ביקורתי של התפתחות טכנולוגיות BIM-והשפעתן על תהליכי ניהול הנתונים בבנייה. נבחנת ההתפתחות ההיסטורית של רעיון ה מהרעיון המקורי של מסד נתונים משולב ועד למבנים שיווקיים מודרניים המקדמים אותם ספקי תוכנה. נבחנת השפעת הפורמטים הקנייניים והמערכות הסגורות על היעילות בעבודה עם נתוני פרויקטים ועל הביצועים הכלליים של תעשיית הבנייה. ניתוח מעמיק של בעיות שונות והקשיים בשילובן עם תהליכי העבודה של חברות CAD התאימות בין מערכות USD, הבנייה. נבחנות מגמות עכשוויות לעבור לפורמטים פתוחים ומפשטים של נתונים, כמו והשפעתם הפוטנציאלית על התעשייה. מוצגים גישות חלופיות להפקת מידע ממערכות סגורות, כולל שיטות הנדסה הפוכה. נבחנות האפשרויות ליישום אינטליגנציה מלאכותית ולמידת מכונה לאוטומציה של תהליכי תכנון וניתוח נתונים בבנייה. מנוסחות תחזיות לפיתוח טכנולוגיות תכנון הממוקדות בצרכים האמיתיים של המשתמשים, ולא באינטרסים של ספקי התוכנה.

פרק 6.1 בתעשיית הבנייה BIM הופעת מושגי

לא היה בגרסה הראשונה של הספר. הנושאים של פורמטים, (BIM) CAD-בתחילה, החלק השישי הזה, המוקדש ל קנייניים, ליבות גיאומטריות ומערכות סגורות הם טכניים מדי, עמוסים בפרטים ובמבט ראשון עשויים להיראות חסרי תועלת עבור מי שרוצה להבין את העבודה עם נתונים. עם זאת, המשוב והבקשות להוסיף הבהרה לגרסה הראשונה של ליבות גיאומטריות, מגוון, CAD הספר הראו: ללא הבנה של כל הקשיים הקשורים למנגנונים הפנימיים של מערכות הפורמטים והסכמות הלא תואמות לאחסון אותם נתונים, אי אפשר באמת להבין מדוע הרעיונות המקדמים על ידי ספקים לעיתים קרובות מקשים על העבודה עם מידע ומונעים מעבר לתכנון פתוח ומבוסס פרמטרים. לכן חלק זה תפס - אינו בעדיפות עבורכם, אתם יכולים לעבור מיד לחלק הבא (BIM) CAD את מקומו הנפרד במבנה הספר. אם נושא "קבלת החלטות מבוססת נתונים, אנליטיקה, אוטומציה ולמידת מכונה: VII חלק"

CAD כמושגים שיווקיים של ספקי open BIM ו-BIM ההיסטוריה של הופעת

עם הופעת הנתונים הדיגיטליים בשנות ה-90, טכנולוגיות המחשב החלו להיכנס לא רק לתהליכי העבודה, אלא גם מודל מידע (BIM), מערכות תכנון ממוחשבות (ולאחר מכן CAD לתהליכי התכנון, מה שהוביל להופעת רעיונות כמו (בניין).

הפכו לשלב חשוב בהיסטוריה של BIM עם זאת, כמו כל חידוש, הם אינם נקודת הסיום של ההתפתחות. רעיונות כמו הם עשויים לפנות מקום לכלים ולגישות מתקדמות יותר, שיתאימו טוב יותר sooner or later תעשיית הבנייה, אך לאתגרים של העתיד.

שהופיע בשנת 2002, BIM ומסובך בקשיים של יישום עצמי, המושג CAD מושפע מהשפעת ספקי עשוי לא לשרוד עד יום הולדתו ה-30, כמו כוכב רוק שזרחה בוהק אך כבה במהירות. הסיבה פשוטה מצליחים להסתגל CAD הדרישות של אנשי המקצוע בתחום הנתונים משתנות מהר יותר ממה שספקי אליהן.

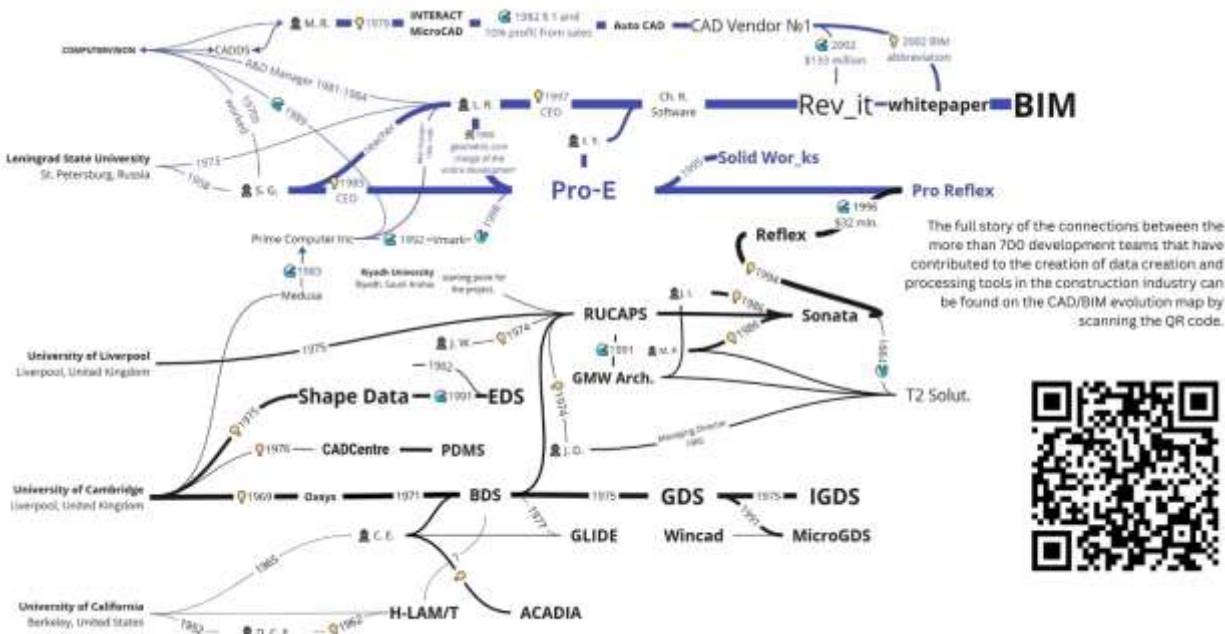
בהתמודדות עם חוסר בנתונים איכותיים, אנשי מקצוע מודרניים בתעשיית הבנייה דורשים תאימות בין פלטפורמות וגישה והסיכון בעיבוד CAD כדי לפשט את הניתוח והעיבוד שלהם. המורכבות של נתוני CAD לנתונים פתוחים מפרויקטי שלהם משפיעים לרעה על כל המשתתפים בתהליך הבנייה: מעצבים, מנהלי פרויקטים, קבלנים בשטח ולבסוף על הלקוח.

אשר דורשים ליבות, CAD במקום סט נתונים מלא לניהול, כיום הלקוח והמשקיע מקבלים קונטיינרים בפורמטים של כדי (BIM) CAD שמתעדכן מדי שנה ותוכנה מיוחדת של API גיאומטריות מורכבות, הבנה של סכמות נתונים, תיעוד לעבוד עם הנתונים. במקביל, חלק גדול מנתוני הפרויקט נשארים לא מנוצלים.

מביאה להנדסה יתרה של ניהול פרויקטים. חברות CAD כיום, בעולם התכנון והבנייה, המורכבות של גישה לנתוני נאלצות לתחזק קשרים הדוקים עם ספקי פתרונות, BIM או מפתחות פתרונות CAD בינוניות וגדולות, שעובדות עם נתוני יקרים SDK באמצעות שימוש בממירי CAD או לעקוף את המגבלות של ספקי, API כדי לגשת לנתונים דרך CAD להנדסה הפוכה, כדי להשיג נתונים פתוחים.

הגישה לשימוש בנתונים קנייניים הפכה מיושנת ואינה עונה על הדרישות של הסביבה הדיגיטלית המודרנית. העתיד יחלק את החברות לשני סוגים: אלו שמשמשות ביעילות בנתונים פתוחים ואלו שיפרשו מהשוק.

Whitepaper (הגדולים CAD) מודל מידע בניין (הופיע בתעשיית הבנייה עם פרסום של אחד מספקי BIM המושג (חשבוניות חומרים), (והחל את דרכו מהגישה הפרמטרית ליצירה BOM בשנת 2002, והושלם עם המושג ההנדסי BIM לתכנון הנדסה Pro-E ועיבוד נתוני פרויקט. הגישה הפרמטרית ליצירה ועיבוד נתוני פרויקט יושמה לראשונה במערכת המודרניים, כולל אלו שמשמשים בהם CAD מערכת זו הפכה לאב טיפוס עבור רבים מפתרונות. (MCAD) מכנית. -כיום בתעשיית הבנייה



ומושגים דומים BIM מפה היסטורית של הופעת המושג

עד תחילת שנות ה-2000 CAD שהקדישו את תשומת הלב שלהם לכלים של ספקי AEC עיתונאים ויועצים בתחום משנת 2002-Whitepaper BIM משנת 2002. למעשה, ה-Whitepaper BIM העבירו את תשומת הלב שלהם ל ומאמרים שפורסמו בשנים 2002, 2003, 2005, ו-2007 שיחקו תפקיד מרכזי בפופולריזציה של המושג 2004 BIM בתעשיית הבנייה.

— ליישום טכנולוגיות מידע בתעשיית הבנייה [CAD שם החברה ספק] מודל מידע בניינים — זו האסטרטגיה של Whitepaper BIM, 2002

בשנת 2002 עם עבודות CAD שפורסם על ידי ספק BIM-עד אמצע שנות ה-2000, חוקרים "החלו לקשר את מושג ה GLIDE, GBM, BPM, של צ'ארלס איסטמן, שהפכה לבסיס עבור מערכות כמו BDS מדעיות מוקדמות יותר, כמו

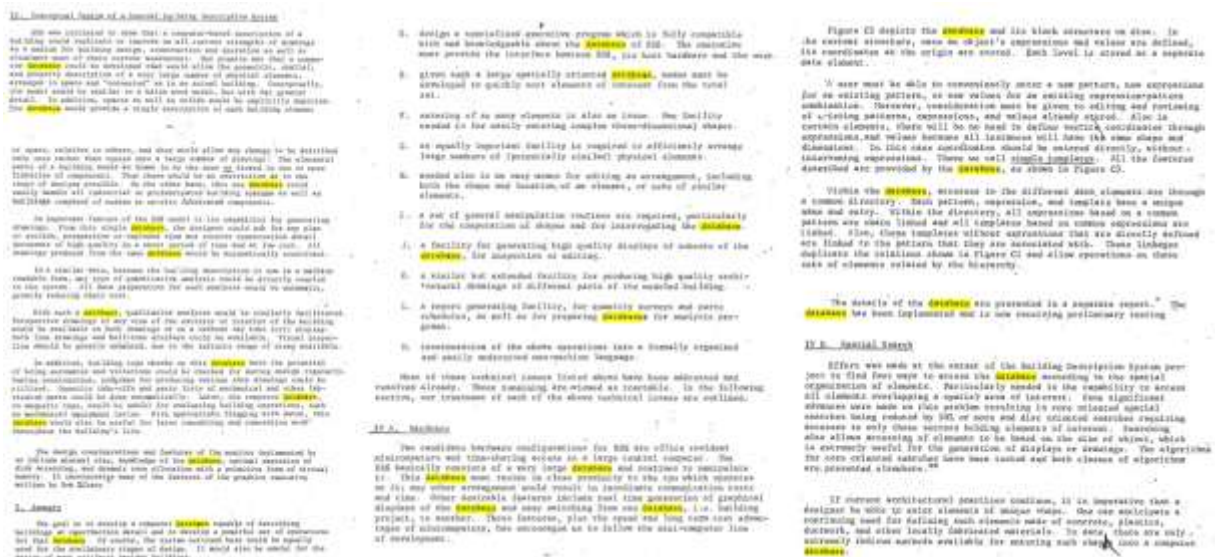
בעבודתו החדשנית "מערכת תיאור בניין" 1974, (צ'ארלס איסטמן הניח את היסודות התיאורטיים של RUCAPS. יותר מכל מונח אחר, למעט - (6.12 אור) מודל המידע המודרני. המונח "מאגר נתונים" מופיע בעבודתו 43 פעמים - "המילה" בניין

הרעיון המרכזי של איסטמן היה שכל המידע על הבניין - מהגיאומטריה ועד תכונות האלמנטים והקשרים, ביניהם - צריך להישמר במאגר נתונים מובנה אחד. מתוך מאגר זה ניתן לייצר באופן אוטומטי שרטוטים מפרטים, חישובים ולנתח עמידה בתקנים. איסטמן ביקר במפורש את השרטוטים כדרך תקשורת מיושנת ומיותרת, מצביע על כפילות המידע, בעיות בעדכון והצורך בעדכון ידני כאשר מתבצעות שינויים. במקום זאת, הוא הציע מודל דיגיטלי אחד במאגר הנתונים, שבו כל שינוי מתבצע פעם אחת ומשתקף אוטומטית בכל הייצוגים.

בעבודתו איסטמן לא שם את הוויזואליזציה במרכז. המידע - פרמטרים, קשרים, תכונות, אפשרויות, noteworthy, ניתוח ואוטומציה - היה במרכז המערכת שלו. השרטוטים, להבנתו, היו רק אחת מהצורות להציג את הנתונים מהמאגר. ולא המקור הראשוני של המידע התכנוני.

המוביל, הביטוי "מאגר נתונים" שימש באותה תדירות כמו CAD מספק BIM הראשונים על Whitepaper-במסמכי ה של צ'ארלס איסטמן - 23 פעמים [60] [בשבעה עמודים, והיה אחד מהמונחים הפופולריים ביותר במסמך לאחר BDS-ב [61] בניין, "מידע", "מודלינג" ו"עיצוב". "עם זאת, עד 2003 במסמכים דומים המונח "מאגר נתונים" מופיע רק פעמיים" ובסוף שנות ה-2000 הנושא של מאגרי נתונים כמעט נעלם מהדיון על נתוני תכנון. בסופו של דבר, המושג "מאגר נתונים" אינטגרטיבי אחד לניתוח ויזואלי וכמותי "לא מומש במלואו

של צ'ארלס איסטמן עם הדגש שלו על מאגרי נתונים BDS כך, תעשיית הבנייה עברה ממצב של מושג מתקדם כמו קודמו של פתרונות) Pro-E ורעיונותיו של סמואל גייסברג על עדכון אוטומטי של נתוני תכנון ממאגרים במוצר ההנדסה השיווקי המודרני, שבו ניהול הנתונים דרך מאגרי נתונים כמעט ולא-BIM הפופולריים בשימוש כיום בבנייה (ל CAD מוזכר, למרות שהמושג הזה היה הבסיס לפיתוחים התיאורטיים המקוריים



המתואר על ידי צ'ארלס איסטמן בשנת 1974, הביטוי "מאגר נתונים" מודגש בצהוב (שימש BDS, 43 במושג 6.12 אור - פעמים.

BIM. ומושגים דומים לו עד שנות ה-2000 פותחו כמאגר נתונים דיגיטלי של בניינים, ולא ככלי ויזואליזציה BDS ומושגים דומים בשנות ה-2002 הפך לכלי תכנון, שבו מאגר הנתונים ירד למקום שני. מה איבדנו במעבר מ: עד אמצע שנות ה-BIM-2010 ה-90 ל

- את העיצוב - BIM, ומושגים דומים הדגישו את האנליטיקה BDS: מאגרי נתונים פתוחים
- את התהליכים, שצריכים להיות - BIM, הדגישה את האנליטיקה של הנתונים BDS: גמישות בעבודה עם נתונים מבוססים על נתונים לא ברורים
- הפכו את בסיסי הנתונים BIM-CAD ב-20 נועדה להיות בסיס נתונים פתוח ומאוחד, בעוד שספקי BDS: שקיפות שלהם לסגורים לחלוטין ונלחמו במשך 20 שנה ללא הצלחה בכלים של הנדסה הפוכה, אשר פותחים פורמטים קנייניים.

במהלך 30 השנים האחרונות, המהנדסים לא קיבלו גישה ל"בסיס הנתונים המאוחד" ואחרי עשרים שנה של התלהבות תעשיית הבנייה מתחילה להבין את ההשלכות של התלהבות זו, BIM, שיווקית סביב כלי

במקום בסיסי נתונים משולבים - מערכות מודולריות סגורות: BIM מציאות

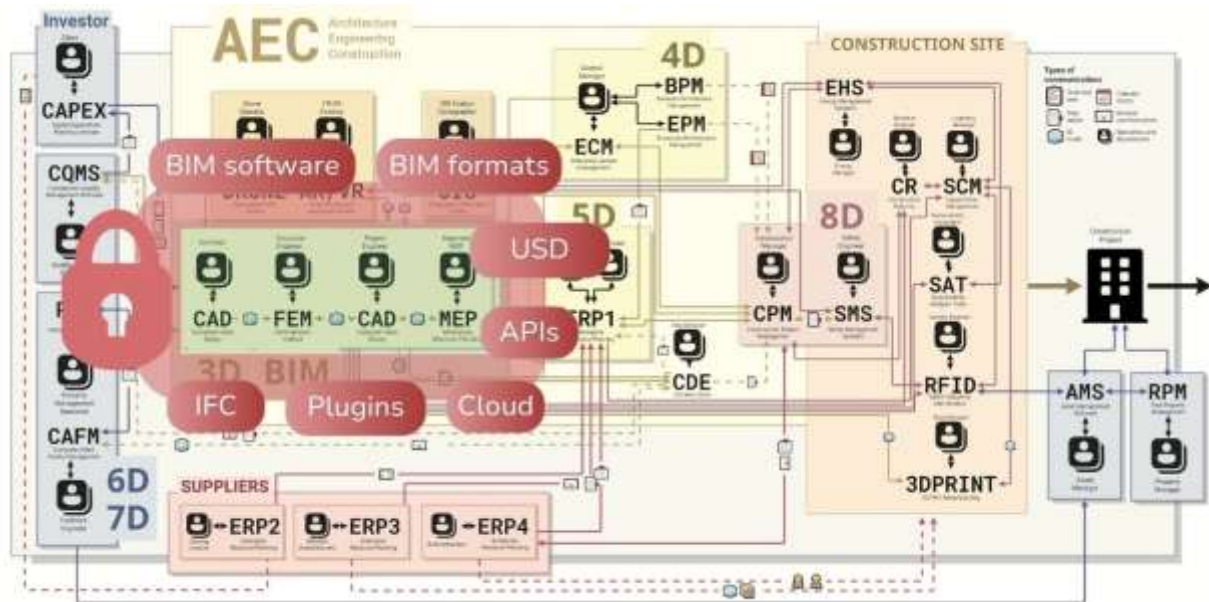
נאלצים (BIM) CAD במקום להתמקד בנתונים, במבנה שלהם ובאינטגרציה לתהליכים מאוחדים, משתמשי מערכות לעבוד עם אוסף מפורק של פתרונות קנייניים, שכל אחד מהם קובע את כללי המשחק שלו

- נותר מיתוס. למרות, BIM הראשונים של Whitepaper-בסיס הנתונים המאוחד, עליו דובר במסמכי ההצהרות הרועשות, הגישה לנתונים עדיין מוגבלת ומפוזרת בין מערכות סגורות
- הפכו לא רק לכלי, אלא לאקוסיסטמה סגורה. במקום החלפת מידע שקופה, המשתמשים נאלצים BIM מודלי קנייניים API-לשלם עבור מנויים ולהשתמש ב
- הנתונים שייכים לספקים, ולא למשתמשים. המידע על הפרויקטים סגור בפורמטים קנייניים או בשירותים בענן, ולא זמין בפורמטים פתוחים ועצמאיים

ולא CAD מהנדסי תכנון ומנהלי פרויקטים לעיתים קרובות אינם מקבלים גישה לא לבסיס הנתונים של מערכות לפורמט שבו מאוחסנים הנתונים של הפרויקט שלהם. זה מקשה על בדיקות מהירות של מידע או ניסוח דרישות ותוספים, מה API למבנה ואיכות הנתונים. גישה לנתונים כאלה דורשת סט של תוכנות מיוחדות, מחוברות דרך שמוביל לבירוקרטיה מופרזת בתהליכים בתעשיית הבנייה. בינתיים, נתונים אלה בשימוש בו זמנית על ידי עשרות -מערכות מידע ומאות מומחים

לאחסן אותם בצורה דיגיטלית ולמכור [CAD (BIM)] או צריכים להיות מסוגלים לנהל את כל הנתונים הללו יש עשרה אנשים [CAD בתוכנת] תוכנה לניהול מחזור החיים והתהליכים, כי על כל מהנדס [מתכנן] שמייצר משהו שעובדים עם הנתונים הללו

BIM, 2005, שיצר את הקונספט של CAD, מנכ"ל ספק -



ומנהלי נתונים באקוסיסטמת עסקי IT בסיסי נתונים נותרו אחד מהמערכות הסגורות האחרונות עבור מחלקות CAD- (BIM-) הבנייה.

הוא יותר אמצעי למיסוד בסיסי נתונים ולא כלי ניהול של ממש, עולה השאלה BIM כאשר ברור כי הלוגית: כיצד להחזיר את השליטה על הנתונים? התשובה היא שימוש במבני נתונים פתוחים, שבהם בעל המידע הוא המשתמש עצמו, ולא ספק התוכנה.

משתמשים ומפתחים של פתרונות בתעשיית הבנייה, בדומה לעמיתיהם בתעשיות אחרות, יתחילו בהכרח להתרחק מהטרמינולוגיה המעורפלת של ספקי התוכנה, ששלטה ב-30 השנים האחרונות, ויתמקדו בהיבטים המרכזיים של "הדיגיטציה" – נתונים ו"תהליכים".

עוד בסוף שנות ה-80, הכיוון המרכזי של פיתוח טכנולוגיות דיגיטליות בתחום הבנייה נתפס כשאלה של גישה לנתונים וניהול מידע פרויקט. עם הזמן, הדגשים השתנו. במקום לפתח גישות שקופות וגישות לעבודה עם נתונים, החלה קידום כניסיון להסיח את דעתם של אנשי המקצוע מהנושאים של ניהול בסיסי - Open BIM - ועקרון ה IFC פעיל של פורמט נתוני פרויקטים.

בתעשיית הבנייה IFC הופעת פורמט פתוח

מוצג כסטנדרט להבטחת תאימות בין מערכות IFC (Industry Foundation Classes) מה שנקרא פורמט פתוח הגדולים ביותר. על בסיס CAD שונות. הפיתוח שלו נעשה במסגרת ארגונים שנוצרו ונשלטו על ידי ספקי CAD (BIM) OPEN BIM. בשנת 2012 קונספט שיווקי בשם CAD פיתחו שתי חברות IFC פורמט

זהו סטנדרט פתוח להחלפת נתונים בתעשיית הבנייה – IFC (Industry Foundation Classes) שונות CAD (BIM) שנועד להבטיח תאימות בין מערכות

והחלפת מידע בין מערכות דרך פורמט פתוח להחלפת CAD עבודה עם מידע מבסיסי נתוני Open BIM-קונספט ה IFC - CAD נתוני

וחברות נוספות [CAD ספק 2]... [CAD 1 ספק]... היא קמפיין שיווקי, שהושק על ידי Open BIM תוכנית עם תקשורת מתואמת, AEC בכל תעשיית OPEN BIM-במטרה לעודד ולקדם את הקידום הגלובלי של קונספט ה OPEN BIM, 2012 תוכנית, CAD ומיתוג משותף, הזמין למשתתפי התוכנית. — מהאתר של ספק

בסוף שנות ה-80, ולאחר מכן נרשם על ידי STEP הותאם על ידי הטכניון של מינכן מפורמט ההנדסה המכאנית IFC בשנת (Industry Alliance for Interoperability) IAI גדול כדי ליצור את הברית CAD חברת תכנון גדולה וספק שונות והיה מבוסס על העקרונות שהונחו CAD פותח כדי להבטיח אינטרופראביליות בין מערכות IFC פורמט. 1994 שנוצר עוד בשנת 1979 על ידי קבוצת משתמשי IGES, אשר, בתורו, נוצר מפורמט, STEP בפורמט ההנדסה המכאנית (-). המכון הלאומי לתקנים וטכנולוגיה (ומשרד ההגנה של ארצות הברית NIST וספקים בתמיכת CAD

התלות הקרובה שלו בליבת הגיאומטריה, כמו גם אי-הבנות ביישום הפורמט על ידי IFC, עם זאת, המבנה המורכב של פתרונות תוכנה שונים, הביאו למגוון בעיות בעת השימוש המעשי בו. עם קשיים דומים - אובדן פירוט, הגבלת דיוק IGES, STEP והצורך להשתמש בפורמטים ביניים - נתקלו גם המומחים להנדסה מכאנית בעבודתם עם הפורמטים IFC. שממנו נוצר



6.14. CAD (BIM) מפה של קשרי צוותי המפתחים ומוצרי ריש.

פרסם (bS מאוחר יותר) IAI ויצר את הארגון IFC שרשם את פורמט CAD בשנת 2000, אותו ספק מסמך לבן "תכנון וייצור משולבים: יתרונות והצדקה". במסמך הודגש הצורך לשמור על פירוט מלא של IGES, הנתונים בעת החלפתם בין תוכניות בתוך אותה מערכת, מבלי להשתמש בפורמטים נייטרליים, כמו במקום זאת, הוצע להבטיח גישה ישירה של יישומים לבסיס הנתונים המרכזי של (IFC-שזהים ל) STEP. דבר שצריך היה למנוע אובדן דיוק במידע CAD,

ומבוסס עליו מפתח (פרטים בחלק השלישי, 3.118 ראה) פרמטרי BOM רוכש מוצר CAD בשנת 2002, אותו ספק כתוצאה מכך, בהחלפת נתונים של פרויקטים בנייה כיום נעשה שימוש רק בפורמטים סגורים. BIM את הקונספט של עצמו בשנת 2000, שהכניס פורמט זה לתחום CAD-על מגבלותיו כתב ספק ה, (STEP) IFC או בפורמט CAD של הבנייה.

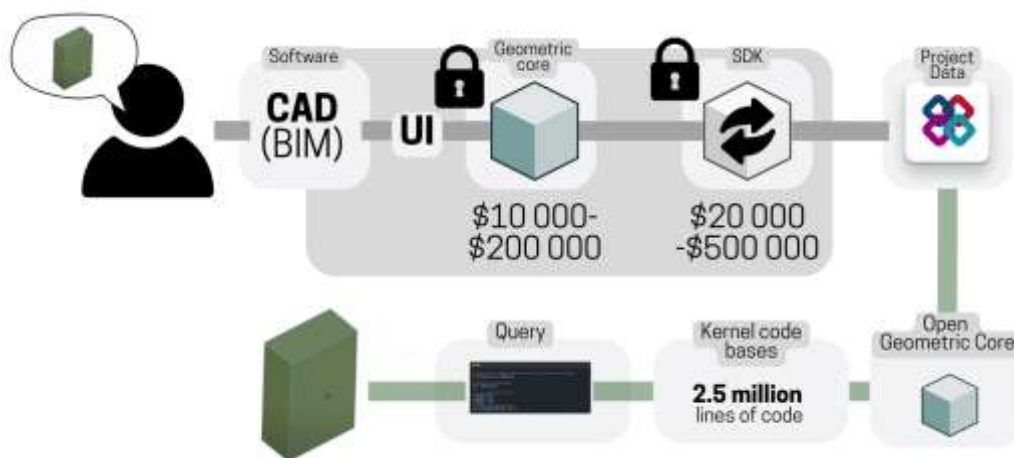
ההיסטוריה המפורטת של האינטראקציה בין יותר מ-700 צוותי מפתחים, שהשתתפו ביצירת כלים ליצירה ועיבוד נתוני [116] "CAD (BIM) בנייה, מוצגת במפה "אבולוציית

מורכב מתיאור גיאומטרי של רכיבי הפרויקט ותיאור של מטא-מידע. לצורך הצגת הגיאומטריה IFC הפורמט הפתוח הפך BREP עם זאת, הייצוג הפרמטרי: CSG ו-Swept Solids נעשה שימוש בדרכים שונות, כגון IFC בפורמט CAD (BIM) שכן פורמט זה נתמך בעת ייצוא מתוכנות IFC לסטנדרט המוביל להעברת גיאומטריה של רכיבים בפורמט CAD. לתוכנות IFC ומאפשר פוטנציאל לעריכה של רכיבים בעת ייבוא חוזר של

כתלות בליבת הגיאומטריה IFC בעיית פורמט

קשה להמחיש או לקבל תכונות גיאומטריות (BREP) מוגדרת פרמטרית IFC-ברוב המקרים, כאשר הגיאומטריה ב מכיוון שעבור עבודה עם גיאומטריה והמחשה שלה, IFC כגון נפח או שטח של ישויות הפרויקט, כאשר יש רק קובץ -אשר אינה קיימת מלכתחילה, (6.15 ראה) במקרה זה נדרש ליבת גיאומטריה

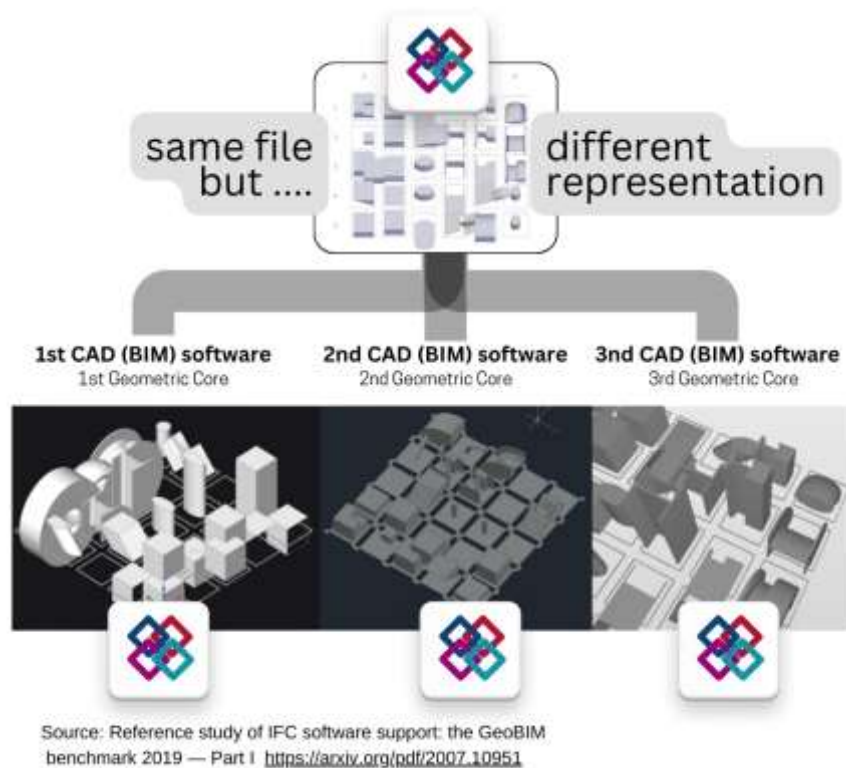
ליבת גיאומטריה היא רכיב תוכנה המספק אלגוריתמים בסיסיים ליצירה, עריכה וניתוח של אובייקטים ויישומים הנדסיים אחרים. היא אחראית על בניית גיאומטריה דו- (BIM), CAD (CAD) גיאומטריים ב ממדית ותלת-ממדית, כמו גם על פעולות עליה, כגון: פעולות בוליאניות, החלקה, חיתוכים, טרנספורמציות והמחשה.



אשר לעיתים, SDK-כיום מתבצעת דרך ליבות גיאומטריות פרופרטאריות ו CAD יצירת גיאומטריה דרך תוכנות 6.15 ראה קרובות אינן שייכות לספקי CAD.

יש ליבת גיאומטריה משלה או שנרכשה IFC ולכל תוכנה העובדת עם פורמטים פרמטריים או פורמט CAD לכל תוכנת לא עשויות להתעורר בעיות, ובתוכנות עם ליבות גיאומטריות שונות IFC-BREP ואם עם רכיבים פרימיטיביים בפורמט רכיבים אלה עשויים להופיע דומים, אז מעבר לבעיות עם מנועי הליבות הגיאומטריות השונות, קיימים רכיבים רבים -שיש להם תכונות ייחודיות להופעה נכונה. בעיה זו נבחנה בפירוט במחקר בינלאומי "מחקר התייחסות לתמיכת תוכנה ב IFC", שפורסם בשנת 2019, [117].

אותם סטים נתונים סטנדרטיים נותנים תוצאות סותרות, כאשר מתגלות מעט תבניות משותפות, ובעיות חמורות ככל הנראה בשל אותה מורכבות גבוהה של מודל הנתונים הסטנדרטי. חלקית, [IFC] נמצאות בתמיכה בסטנדרט הסטנדרטים עצמם אשמים בכך, מכיוון שהם לעיתים קרובות משאירים כמה פרטים לא מוגדרים, עם רמת חופש גבוהה ופרשנויות שונות אפשריות. הם מאפשרים מורכבות גבוהה בארגון ואחסון של אובייקטים, דבר שאינו תורם IFC-להבנה אוניברסלית יעילה, יישומים ייחודיים ומידול נתונים מתואם [117]. -מחקר התייחסות לתמיכת תוכנה ב 2021.



ליבות גיאומטריות שונות מספקות ייצוגים שונים לאותה גיאומטריה, המתוארת בפרמטרים (על בסיס חומרים [117] 6.16 איור).

כתוצאה מכך, מי IFC הבנה נכונה של "מצבים מוגדרים" זמינה לחברים בתשלום בארגונים מיוחדים, העוסקים בפיתוח גדולים, או להגיע להבנה CAD ינסה לשתף פעולה עם ספקי IFC שמעוניין לגשת לידע חשוב על תכונות מסוימות של איכותית של התכונות באמצעות מחקר עצמאי.

ושואל את עמיתך הספקים: "למה כך מועברת המידע IFC אתה נתקל בשאלה על ייבוא וייצוא נתונים דרך פורמט במפרט הפתוח לא נאמר דבר על כך. "התשובה מספקים "היות IFC על העברת פרמטרים של חדרים בקובץ [118] CAD 2021 יודעים "האירופיים": נכון, לא נאמר, אבל זה מותר. — מתוך ריאיון עם מפתח

CAD, מתאר גיאומטריה דרך פרימיטיבים פרמטריים, אך אינו מכיל ליבה מובנית — תפקידה מתבצע על ידי תוכנת IFC, אשר דרך הליבה הגיאומטרית מקמפלת את הגיאומטריה. הליבה הגיאומטרית מבצעת חישובים מתמטיים ומגדירה מכיל פאות לא תקינות, תוכנות שונות עם ליבות IFC רק מספק נתונים לפרשנות שלה. אם IFC חיתוכים, בעוד גיאומטריות שונות יכולות להתעלם מהן או להנפיק שגיאות, בהתאם לליבה.

יש לענות על השאלה המרכזית, שעליה קשה למצוא תשובה חד IFC בסופו של דבר, כדי לעבוד עם פורמט משמעותי - איזה כלי, עם איזו ליבה גיאומטרית יש להשתמש כדי להשיג את איכות הנתונים, שהייתה IFC? ממנה התקבל, CAD, לפרויקט במקור בתוכנת

אינן מאפשרות שימוש ישיר בנתוני הפרויקט לאוטומציה של תהליכים IFC בעיות איכות הנתונים ומורכבות פורמט

סגורים CAD לניתוח ועיבוד נתונים, מה שלעתים קרובות מביא את המפתחים לצורך בלתי נמנע להשתמש בפתרונות [65] בשנת 1994 IFC עם "גישה איכותית" לנתונים [63], על כך כתב גם הספק עצמו, שרשם את

בליבת הגיאומטריה יכולות להתממש רק על ידי צוותי מפתחים IFC-כל התכונות של הצגת וגרפיקה של פרמטרים ב IFC, גדולים, שיש להם ניסיון בעבודה עם ליבות גיאומטריות. לכן, הפרקטיקה הנוכחית של תכונות ומורכבות פורמט, ויש לה הרבה במשותף עם האסטרטגיה של ספקי תוכנה גדולים "לאמץ, להרחיב CAD מועילה בראש ובראשונה לספקי [94] להשמיד", כאשר המורכבות הגוברת של הסטנדרט למעשה יוצרת מחסומים לשחקנים קטנים בשוק

האסטרטגיה של ספקים גדולים באסטרטגיה כזו עשויה לכלול התאמת סטנדרטים פתוחים, הוספת הרחבות ופונקציות משלהם, כדי ליצור תלות של משתמשים במוצרים שלהם לצורך דחיקת מתחרים

שונות, בפועל ממלא תפקיד של CAD (BIM) שנועד להיות גשר אוניברסלי בין מערכות IFC, פורמט STEP, שונות, בדומה לפורמט CAD אינדיקטור לבעיות תאימות בין ליבות גיאומטריות של פלטפורמות. ממנו הוא נוצר במקור

גדולים, אשר יכולים CAD אפשרי רק עבור ספקי IFC בסופו של דבר, היום יישום מלא ואיכותי של האונטולוגיה להשקיע משאבים משמעותיים בתמיכה בכל הישויות ומיפוי שלהן עם הליבה הגיאומטרית הפנימית שלהם, אשר אינה כסטנדרט. ספקים גדולים גם יכולים לתאם ביניהם את הפרטים הטכניים של המאפיינים, אשר עשויים IFC קיימת עבור IFC להיות לא נגישים אפילו למשתתף הפעיל ביותר בארגונים העוסקים בפיתוח הפורמט

עבור צוותים קטנים ועצמאיים ופרויקטים עם קוד פתוח, השואפים לתמוך בפיתוח פורמטים אינטרופוראביליים, היעדר ליבה גיאומטרית הופך לבעיה רצינית. בלעדיה, כמעט בלתי אפשרי לקחת בחשבון את כל המגוון של הדקויות והניואנסים הקשורים בהחלפת נתונים בין פלטפורמות

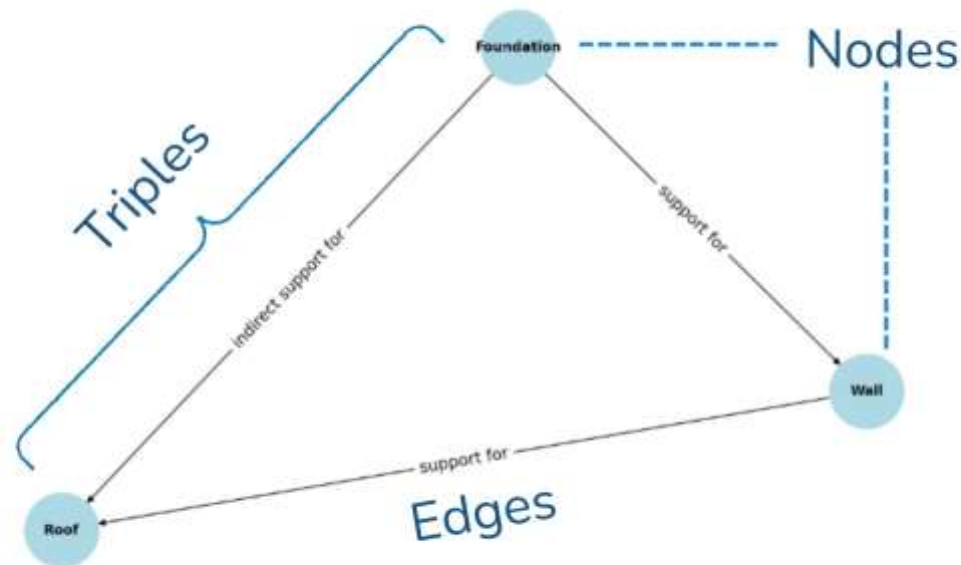
התעוררו דיונים בתעשיית הבנייה על תפקיד open BIM, הפרמטרי ומושג ה IFC עם ההתפתחות של פורמט האונטולוגיה והסמנטיקה בניהול נתונים ותהליכים

הופעת נושא הסמנטיקה והאונטולוגיה בבנייה

הסמנטיקה, IFC בזכות רעיונות יצירת האינטרנט הסמנטי בסוף שנות ה-90 ומאמצי הארגונים העוסקים בפיתוח פורמט ואונטולוגיות הפכו לאחד המרכיבים המרכזיים של הסטנדרטיזציה, אשר נדונה בתעשיית הבנייה עד אמצע שנות ה-2020.

טכנולוגיות סמנטיות הן האחידות, הסטנדרטיזציה והשינוי של כמויות גדולות של נתונים הטרוגניים, כמו גם יישום חיפוש מורכב

המוצגת בצורה, OWL (Web Ontology Language) לשמירת נתונים סמנטיים נעשה שימוש בשפת האונטולוגיות, שייכת למודלים גרפיים של נתונים OWL (Resource Description Framework). טריפלטים-RDF של גרפי -"על סוגיהם דיברנו בפירוט בפרק "מודלי נתונים: יחס בנתונים וקשרים בין רכיבים



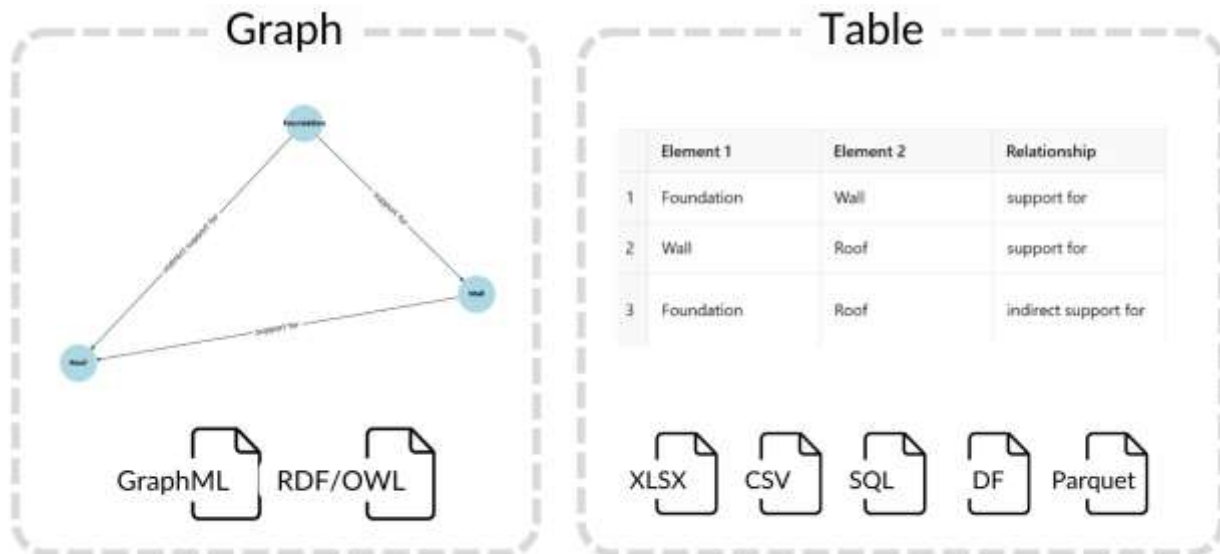
הממחיש את הקשרים בין רכיבי הבניין (Triples) וטריפלים (Edges) קשרים (Nodes), צמתים (Nodes) מודל נתונים.

תיאורטית, הסקת לוגיקה של ריזונרים (תוכנה להסקה לוגית אוטומטית) מאפשרת להפיק הצהרות חדשות על בסיס האונטולוגיות. לדוגמה, אם באונטולוגיה הבנייתית נרשם כי "יסוד הוא תמיכה לקיר", וכי "קיר הוא תמיכה לגג", הריזונר יוכל באופן אוטומטי להסיק כי "יסוד הוא תמיכה לגג".

מנגנון כזה מועיל לאופטימיזציה של ניתוח נתונים, מכיוון שהוא מאפשר להימנע מהגדרה מפורשת של כל התלויות. עם זאת, הוא אינו יוצר ידע חדש, אלא רק מגלה ומסדר את העובדות שכבר ידועות.

הסמנטיקה אינה יוצרת משמעות או ידע חדש מעצמה ואינה עולה על טכנולוגיות אחרות לאחסון ועיבוד נתונים בהיבט זה. הצגת נתונים מבסיסי נתונים רלציוניים בצורה של טריפלים אינה עושה אותם יותר משמעותיים. החלפת טבלאות במבנים גרפיים עשויה להיות מועילה לאחידות מודלי הנתונים, חיפוש נוח ועריכה בטוחה, אך היא אינה עושה את הנתונים "חכמים" יותר - המחשב אינו מתחיל להבין טוב יותר את תוכנם.

(SQL) קשרים לוגיים בנתונים ניתן לארגן גם ללא טכנולוגיות סמנטיות מורכבות. מסדי נתונים רלציוניים מסורתיים מאפשרים לבנות תלות דומה. לדוגמה, במסד נתונים עמודתי ניתן להוסיף שדה, XLSX או CSV כמו גם פורמטים גרפים, RDF, OWL, תמיכה לגג "ולחבר אוטומטית את הגג עם היסוד בעת יצירת קיר. גישה זו מתבצעת ללא שימוש ב" או ריזונרים, וממשיכה להיות פתרון פשוט ויעיל לאחסון וניתוח נתונים.



השוואת מודלים גרפיים וטבלאיים של נתונים המייצגים את אותם קשרים לוגיים.

לעקוב אחרי הקונספט של האינטרנט הסמנטי IFC ההחלטה של מספר חברות בנייה גדולות וארגון העוסק בפיתוח פורמט שנראה מבטיח בסוף שנות ה-90, השפיעה באופן משמעותי על פיתוח הסטנדרטים בתעשיית הבנייה

עם זאת, הפרדוקס הוא שהקונספט של האינטרנט הסמנטי, שנועד במקור לאינטרנט, לא זכה להפצה רחבה אפילו בסביבתו האינטרנט הסמנטי המלא לא הופיע כפי שתוכנן, והקמתו כבר נראית לא, RDF ו-OWL המקורית. למרות הפיתוח של סבירה.

מדוע טכנולוגיות סמנטיות אינן עומדות בציפיות בתעשיית הבנייה

תעשיות אחרות נתקלות במגבלות טכנולוגיות בשימוש בסמנטיקה. בתעשיית המשחקים, ניסי לתאר אובייקטים במשחק, ואינטראקציות שלהם באמצעות אונטולוגיות לא היו יעילים בשל הדינמיות הגבוהה של השינויים. בסופו של דבר יחד עם פתרונות אלגוריתמיים, הפכו להיות מועדפים. מצב, XML ו-JSON פורמטים פשוטים יותר של נתונים, כמו דומה התרחש גם בתחום הנדל"ן: בשל הבדלים אזוריים במונחים ושינויים תכופים בשוק, השימוש באונטולוגיות היה התמודדו טוב יותר עם משימות החלפת נתונים RETS מורכב מדי, בעוד שמסדי נתונים פשוטים וסטנדרטים כמו

קשיים טכניים, כמו מורכבות הסימון, עלות התחזוקה הגבוהה ומיעוט המוטיבציה של המפתחים, עיכבו את קישורם של יישומי האינטרנט הסמנטי גם בתחומים אחרים בכלכלה. לא הפך לסטנדרט המוני, ואונטולוגיות היו מורכבות מדי ולא כלכליות (RDF (Resource Description Framework

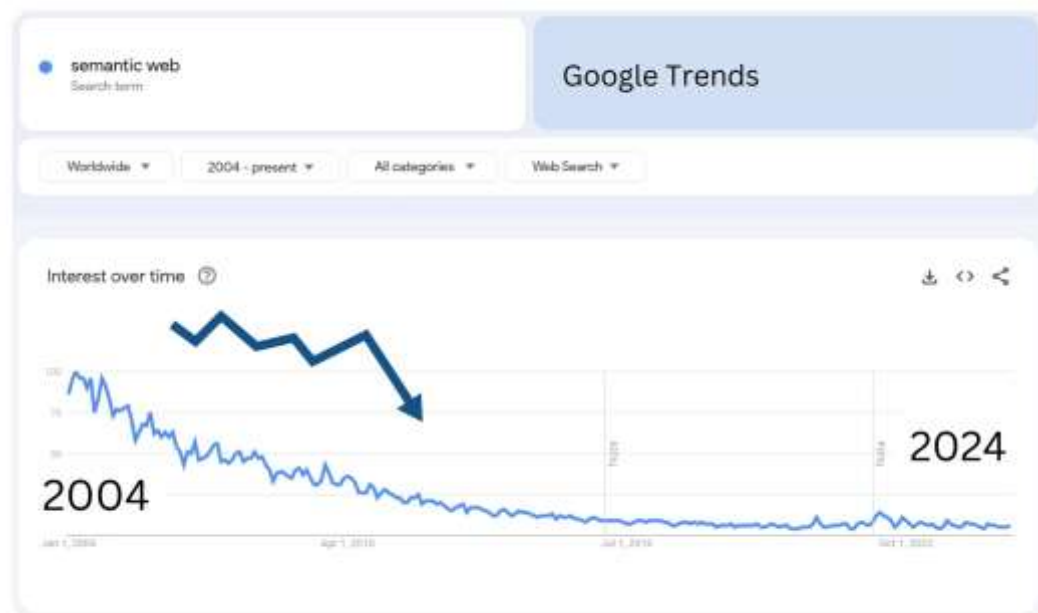
כתוצאה מכך, הרעיון השאפתני של יצירת אינטרנט סמנטי גלובלי לא התממש. אף על פי שחלק מהאלמנטים של מצאו שימוש בפתרונות ארגוניים, המטרה המקורית של יצירת מבנה נתונים SPARQL-הטכנולוגיה, כמו אונטולוגיות וכולל לא הושגה.

הקונספט של אינטרנט שבו מחשבים יכולים להבין את משמעות התוכן התגלה כמורכב טכנית ולא משתלם כלכלית. לכן לצרכים RDF ו-OWL המקורית שתרמו ברעיון זה צמצמו עם הזמן את השימוש בו לכלים מועילים ספציפיים, והשאירו את

בשני העשורים האחרונים מאפשר להעריך כי Google Trends ארגוניים מיוחדים, ולא לאינטרנט בכללותו. ניתוח -ייתכן שאין יותר פוטנציאל לפיתוח האינטרנט הסמנטי

אין צורך להכפיל ישויות ללא צורך. אם קיימות מספר הסברים לוגיים שאינם סותרים לתופעה כלשהי, המבחרים אותה באותה מידה, יש להעדיף, בתנאים שווים, את הפשוט ביותר מביניהם. להב של אוקאם

בבניית נתונים, אם אפשר לעבד SPARQL-כאן עולה שאלה הגיונית: מדוע בכלל להשתמש בטרפלטים, ריזונרים ו SQL באפליקציות ארגוניות? (SQL, Pandas, Apache®) נתונים באמצעות שאלות מובנות פופולריות לעומת זאת, דורש מבנים גרפיים מורכבים ותוכנה מיוחדת ואינו מושך, SPARQL. הסטנדרט לעבודה עם מסדי נתונים את תשומת הלב של מפתחים לפי מגמות בגוגל



עניין בשאלות "האינטרנט הסמנטי" לפי סטטיסטיקות גוגל 6.19. ריש.

מסדי נתונים גרפיים ועצי סיווג עשויים להיות מועילים במקרים מסוימים, אך השימוש בהם לא תמיד מוצדק עבור רוב המשימות היומיומיות. בסופו של דבר, יצירת גרפי ידע ושימוש בטכנולוגיות של האינטרנט הסמנטי יש משמעות רק כאשר יש צורך לאחד נתונים ממקורות שונים או לבצע מסקנות לוגיות מורכבות

המעבר מטבלאות למודלים גרפיים של נתונים מאפשר לשפר את החיפוש ולאחד את זרימת המידע, אך אינו עושה את הנתונים ליותר משמעותיים עבור מכונות. השאלה אינה אם יש להשתמש בטכנולוגיות סמנטיות, אלא היכן הן באמת מביאות תועלת. לפני שמטמיעים אונטולוגיה, סמנטיקה ומסדי נתונים גרפיים בחברה שלכם, בדקו אילו חברות כבר משתמשות בהצלחה בטכנולוגיות אלו, ואיפה הן לא עמדו בציפיות.

למרות הציפיות השאפתניות, טכנולוגיות סמנטיות לא הפכו לפתרון אוניברסלי למבנה נתונים בתעשיית הבניין. בפועל, טכנולוגיות אלו לא הביאו לפתרון אוניברסלי, אלא רק הוסיפו מורכבויות חדשות, והמאמצים הללו משקפים את השאיפות

הלא ממומשות של רעיון האינטרנט הסמנטי, שבו הציפיות עלו בהרבה על המציאות.

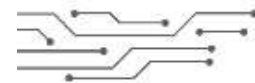


מורכבות ועד נתונים ממושטות לניתוח CAD ו-BIM גיאומטריה ומידע בתהליכי בנייה: ממערכות 6.110. ריש

IoT, הכישלונות של האינטרנט הסמנטי הוקצו על ידי הופעת טכנולוגיות חדשות (נתונים גדולים IT-אם בתחום ה-AR/VR), למידת מכונה, אז לתעשיית הבניין אין סיבות כאלה,

מעבר לבעיות בשימוש בקונספטים להעברת קשרים בין נתונים בין רכיבי הפרויקט, קיימת בעיה בסיסית – זמינות הנתונים עצמם. בתעשיית הבניין עדיין שולטות מערכות סגורות, המקשות על עבודה עם נתונים, חילופי מידע ושיפור היעילות של התהליכים.

בדיוק האופי הסגור של הנתונים הופך לאחד מהמכשולים המרכזיים המונעים את התפתחות הפתרונות הדיגיטליים בבניין כל תוכנה משתמשת (BIM) CAD שבה פורמטים פתוחים ומאוחדים הפכו לסטנדרט, במגזר, IT-בניגוד לתעשיית ה-בפורמט ייחודי משלה, מה שיוצר אקוסיסטמות סגורות ומגביל את המשתמשים באופן מלאכותי.



פרק 6.2. פורמטים סגורים של פרויקטים ובעיות אינטרופראביליות

CAD (BIM) נתונים סגורים וירידת פרודוקטיביות: פינה ללא מוצא בתעשיית

הביא לכך שכל תוכנה יש לה פורמט נתונים ייחודי, שהוא או סגור ואינו נגיש CAD האופי הפרופרטארי של מערכות - או נגיש בצורה חצי-מובנית דרך תהליך המרה מורכב למדי, RVT, PLN, DWG, NDW, NWD, SKP, IFC, STEP ו- ifcXML, ifcJSON, BIMJSON, ifcSQL, CSV, JSON, XML (CPIXML).

פורמטים שונים של נתונים, שבהם ניתן לאחסן את אותם נתונים על אותם פרויקטים, לא רק שמבדילים במבנה, אלא, כוללים גם גרסאות שונות של תיוג פנימי, שעל המפתחים לקחת בחשבון כדי להבטיח תאימות של יישומים. לדוגמה של שנת 2026, אך אותו פרויקט לא ייפתח לעולם בכל הגרסאות CAD של שנת 2025 ייפתח בתוכנת CAD פורמט שיכלו להיות לפני שנת CAD 2025 של תוכנת.

מבלי לספק גישה ישירה למסדי נתונים, ספקי תוכנה בתעשיית הבנייה לעיתים קרובות יוצרים פורמט ייחודי וכלים אליו, שעל המומחה (מהנדס מתכנן או מנהל נתונים) להשתמש בהם כדי לגשת, לייבא ולייצא נתונים.

מעלים באופן קבוע את המחירים על השימוש (ERP/PMIS כגון) ופתרונות קשורים CAD (BIM) כתוצאה מכך, ספקי במוצרים, והמשתמשים הרגילים נאלצים לשלם "עמלה" בכל שלב של העברת נתונים בפורמטים: על חיבור, ייבוא, ייצוא ועבודה עם נתונים שהמשתמשים יצרו בעצמם.

פופולריים בשנת 2025 תגיע לדולר אחד לכל עסקה CAD- (BIM-) עלות הגישה לנתונים באחסון בענן ממוצרי לבנייה עבור חברות בינוניות מגיעים לסכומים של חמישה ושישה ספרות בשנה ERP ומגויים על מוצרים.

מהות תוכנת הבנייה המודרנית היא שלא האוטומציה או העלאת היעילות, אלא היכולת של המהנדסים להבין תוכנה ספציפית מאוד משפיעה על איכות ועלות עיבוד הנתונים של פרויקט הבנייה, כמו גם על הרווחיות והקיום ארוך הטווח של חברות המיישמות פרויקטים בנייה.

שמשמשות בעשרות מערכות ובמאות תהליכים, וככתוצאה מכך, חוסר CAD, חוסר גישה למסדי נתונים של מערכות תקשורת איכותית בין אנשי מקצוע שונים, הביא את תעשיית הבנייה למעמד של אחד הסקטורים הבלתי יעילים ביותר בכלכלה מבחינת פרודוקטיביות.

טכנולוגיות בנייה (ERP), הופעת מערכות חדשות CAD- (BIM-) במהלך 20 השנים האחרונות של שימוש בתכנון חדשות וחומרים, פרודוקטיביות של כל תעשיית הבנייה ירדה ב-20%, בעוד שהפרודוקטיביות הכוללת של כל הסקטורים - בכלכלה, שאין להם בעיות גידול למסדי נתונים ודומיהם, עלתה ב-70% (96% בתעשייה המעובדת).



בשל הבידוד והמורכבות של נתוני הפרויקטים, שעליהם תלויים עשרות מחלקות ומאות תהליכים בתעשיית הבנייה, מהירות קבלת ההחלטות נמוכה בכמה פעמים מאשר בתעשיות אחרות.

חברת הבנייה הגדולה CAD עם זאת, ישנם גם דוגמאות בודדות לגישות חלופיות ליצירת אינטרופראביליות בין פתרונות שהחל כבר בשנת 2018, מדגימה כיצד ניתן לחרוג מהלוגיקה הקלאסית של SCOPE, ביותר באירופה עם פרויקט או להסתמך על ליבות גיאומטריות קניניות, המפתחים של IFC במקום לנסות להכניע את CAD (BIM). מערכות שונות, להמירם לפורמטים CAD של הנדסה הפוכה כדי לחלץ נתונים מתוכניות SDK וב-API-משתמשים ב SCOPE ולהשתמש בהם במאות תהליכי עסקים, OCCT על בסיס ליבת גיאומטריה פתוחה CPIXML או OBJ נייטרליים כמו של חברות בנייה ועיצוב. עם זאת, למרות ההתקדמות של הרעיון, פרויקטים כאלה מתמודדים עם מגבלות ומורכבות של ליבות גיאומטריות חינומיות והם עדיין נשארים חלק מאקוסיסטמות סגורות של חברה אחת, המייצרות את הלוגיקה של פתרונות חד-ספקיים.

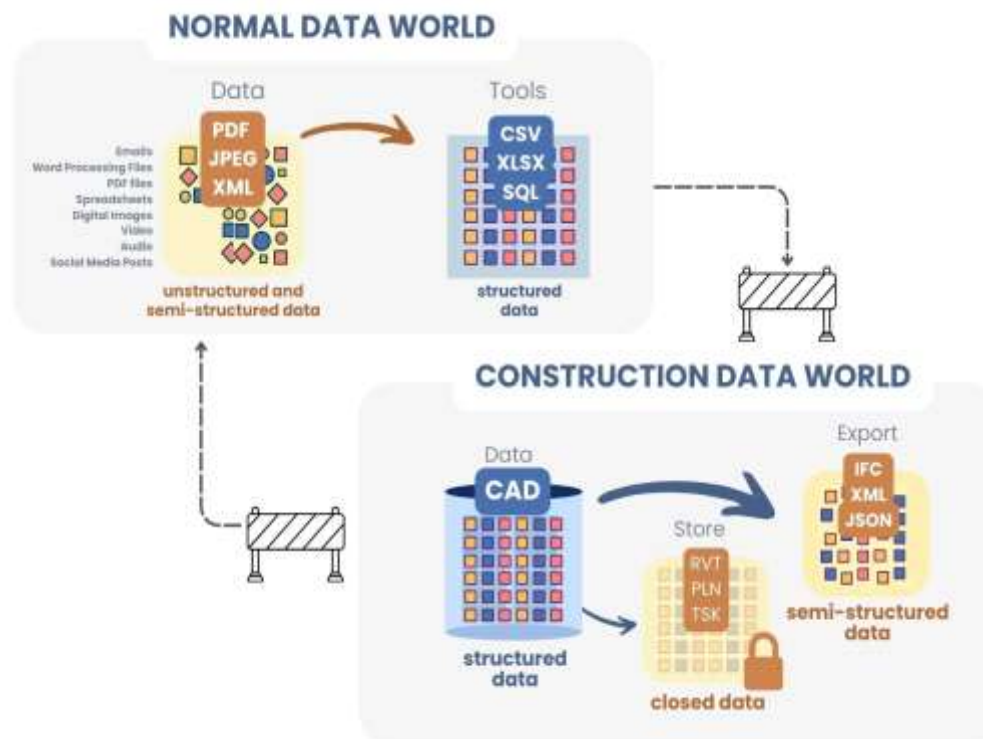
בשל המגבלות של מערכות סגורות וההבדלים בפורמטי הנתונים, כמו גם חוסר הכלים היעילים לאיחוד שלהם, חברות מתמודדות עם הצטברות של כמויות משמעותיות של נתונים בדרגות שונות של מבנה CAD הנאלצות לעבוד עם פורמטי וסגירות. נתונים אלה אינם מנוצלים כראוי ונעלמים בארכיונים, שם עם הזמן הם נשארים נשכחים ולא בשימוש.

נתונים שהושגו בעמל רב בשלב התכנון, בשל המורכבות והסגירות שלהם, הופכים להיות בלתי נגישים לשימוש נוסף.

כתוצאה מכך, במשך 30 השנים האחרונות, מפתחים בתעשיית הבנייה נאלצים להתמודד שוב ושוב עם אותה בעיה: כל פתוחות וסגורות קיימות. ניסיונות CAD פורמט סגור חדש או פתרון קנייני יוצר צורך באינטגרציה עם מערכות, שונים רק מסבכים את האקוסיסטמה של הנתונים CAD ו-BIM מתמשכים להבטיח אינטרופראביליות בין פתרונות. במקום לקדם את הפשטתה וסטנדרטיזציה שלה.

CAD מיתוס האינטרופראביליות בין מערכות

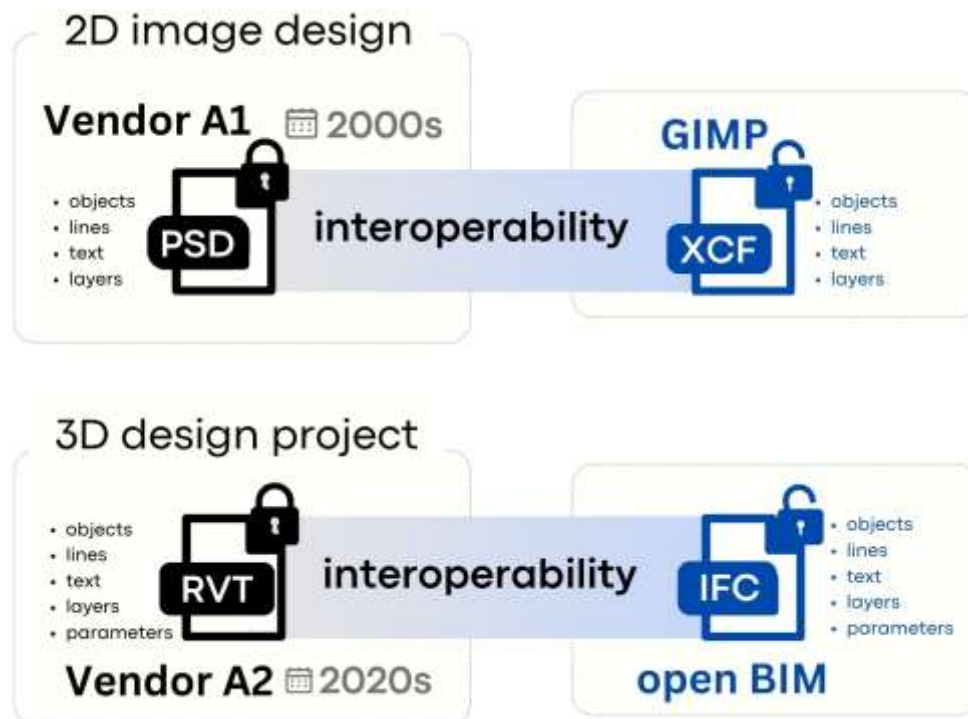
DWG היה פריצת הפורמט הקנייני CAD אם באמצע שנות ה-90 הכיוון המרכזי של פיתוח האינטרופראביליות בסביבת ובפתיחת הפורמט הפופולרי ביותר של שרטוטים עבור כל תעשיית הבנייה Open DWG שהסתיימה בניצחון של ברית - אז באמצע שנות ה-2020 הדגש השתנה. בתעשיית הבנייה מתפתחת מגמה חדשה: צוותי מפתחים רבים מתמקדים - (פתוח BIM) ופתרונות פתוחים IFC פורמט, (סגור BIM) סגורות CAD ביצירת מה שמכונה "גשרים" בין מערכות המספקות חיבור טכני בין פלטפורמות, OCCT וליבת גיאומטריה IFC רוב היוזמות הללו מבוססות על השימוש בפורמט שונות. גישה זו נחשבת לכיוון מבוטא, שיכול לשפר באופן משמעותי את חילופי הנתונים ולהגביר את התאמת הכלים התוכנתיים.



CAD בעוד שענפים אחרים עובדים עם נתונים פתוחים, תעשיית הבנייה נאלצת לעבוד עם פורמטים סגורים או חלשים-מובנים של (BIM).

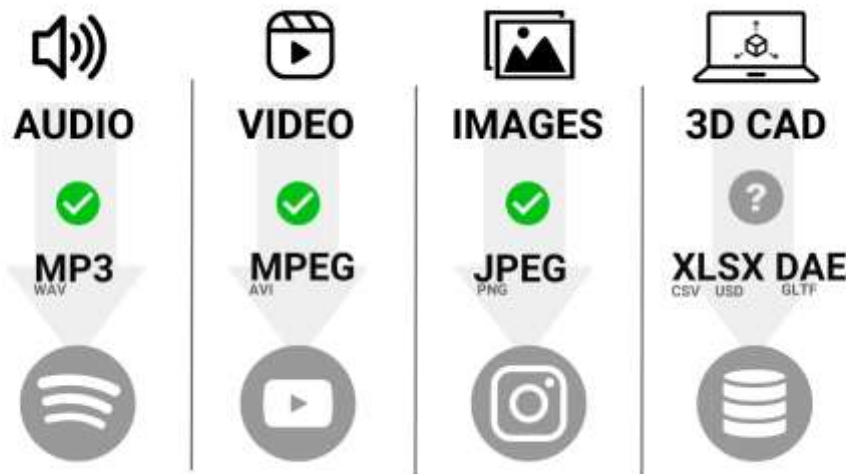
גישה דומה יש לה מקבילות היסטוריות. בשנות ה-2000, מפתחים, בניסיון להתגבר על הדומיננטיות של הספק הגדול ניסו ליצור אינטגרציה חלקה בין הפתרון הקנייני שלו לבין (Dבעולם ה-2) ביותר של תוכנות עריכת גרפיקה, אז, כמו היום בבנייה, מדובר בניסיון לחבר בין מערכות סגורות ופתוחות. GIMP האלטרנטיבה החינמית בקוד פתוח - תוך שמירה על פרמטרים מורכבים, שכבות ולוגיקה פנימית של פעולת התוכנה

עם זאת, המשתמשים למעשה חיפשו פתרונות פשוטים - נתונים שטוחים ופתוחים ללא מורכבות מיותרת של שכבות המשתמשים שאפו לפורמטים פשוטים ופתוחים של נתונים. (CAD-אנלוגיים לגרעין הגיאומטרי ב) ופרמטרים בתוכנות כיום הם בשימוש ברשתות. GIF ו-PNG, JPEG-חופשיים מלוגיקה מיותרת. בתחום הגרפיקה, פורמטים כאלה הפכו ל חברתיות, באתרי אינטרנט וביישומים - הם מעובדים ומפורשים בקלות, ללא תלות בפלטפורמה או ביצרן התוכנה



GIMP ויזואליזציה 6.23: החלפת פורמטים של נתונים בבנייה דומה למסלול שנעשה בניסיון לחבר בין מוצר קנייני פופולרי לבין בקוד פתוח בשנות ה-2000.

עבור יישומים XCF או PSD כתוצאה מכך, בתעשיית התמונה כיום כמעט אף אחד לא משתמש בפורמטים סגורים כמו רשתות חברתיות כמו פייסבוק ואינסטגרם או כתוכן באתרי אינטרנט. במקום זאת, ברוב המשימות נעשה שימוש המבטיחים פשטות שימוש ותאימות רחבה. פורמטים פתוחים, GIF ו-PNG, JPEG בפורמטים שטוחים ופתוחים כמו הפכו לסטנדרט בהחלפת תמונות בזכות הרבגוניות שלהם והתמיכה הרחבה, מה שמקל על השימוש, JPEG ו-PNG כמו בהם בפלטפורמות שונות. מעבר דומה מתרחש גם בפורמטים אחרים להחלפה, כמו וידאו ואודיו, שבהם פורמטים בולטים בזכות היעילות של דחיסת הנתונים והתאימות הרחבה. המעבר הזה MP3 ו-MPEG אוניברסליים כמו לסטנדרטיזציה הפך את ההחלפה והשחזור של תוכן ומידע לפשוטים יותר, והפך אותם לנגישים לכל המשתמשים בפלטפורמות שונות.



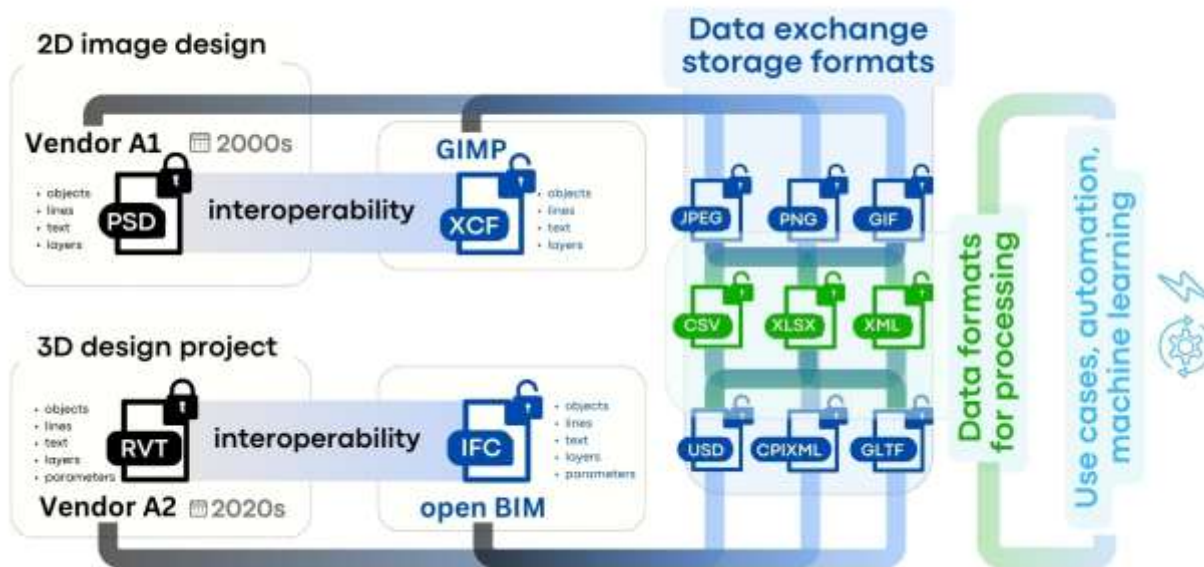
ויזואליזציה 6.24: פורמטים פשוטים ללא פונקציות עריכה מורכבות הפכו לפופולריים להחלפת נתונים ושימוש בהם.

USD, OBJ, glTF, DAE, DXF, SQL ו-XLSX (BIM). פורמטים פשוטים ופתוחים כמו גרעינים גיאומטריים, BREP, את כל המידע הנדרש, כולל גיאומטריה ומטא-נתונים, מבלי צורך לתפעל מבנה מורכב של המוצעים, NWC, SVF, SVF2, CPIXML ו-CP2 או מסווגים פנימיים של ספקים ספציפיים. פורמטים קנייניים, כמו על ידי ספקי תוכנה מובילים, מבצעים גם הם פונקציות דומות, אך נשארים סגורים, בניגוד לסטנדרטים הפתוחים.

noteworthy (and it is worth reminding once again, as already noted in the previous chapter) that this idea - the rejection of intermediate neutral and parametric formats such as IGES, STEP, and IFC - was supported back in 2000 by the main CAD vendor, which created the BIM Whitepaper and registered the IFC format in 1994. In the 2000 Whitepaper "Integrated Design and Production," the CAD vendor emphasizes the importance of native access to the CAD data repository within the software environment, without the need to use intermediate translators and parametric formats to maintain the completeness and accuracy of information.

The construction industry still needs to agree on either the tools for accessing CAD databases, or their forced reverse engineering, or the adoption of a common simplified data format for use outside CAD platforms (BIM). For example, many large companies in Central Europe and German-speaking regions operating in the construction sector use the CPIXML format in their ERP systems. This proprietary format, which is a type of XML, combines CAD project data (BIM), including geometric and metadata, into a single organized simplified structure. Large construction companies are also creating new proprietary formats and systems, as seen in the SCOPE project mentioned in the previous chapter.

The closed logic of parametric CAD formats or complex parametric IFC (STEP) files are often redundant in most business processes. Users are looking for simplified and flat formats such as USD, CPIXML, XML&OBJ, DXF, glTF, SQLite, DAE&XLSX, which contain all the necessary information about the elements but are not burdened by the excessive logic of BREP geometry construction, dependence on geometric kernels, and internal classifications of specific CAD and BIM products.-



For most use cases, users choose the simplest formats that are not dependent on vendor software

The emergence of flat image formats such as JPEG, PNG, and GIF, free from the excessive logic of internal vendor engines, has contributed to the development of thousands of compatible solutions for processing and using graphics. This has led to the emergence of a variety of applications, from retouching and filtering tools to social networks such as Instagram, Snapchat, and Canva, where these simplified data can be used without being tied to a specific software developer.

Standardization and simplification of CAD design formats will stimulate the emergence of many new convenient and independent tools for working with construction projects.

The rejection of the complex logic of vendor applications tied to closed geometric kernels and the transition to universal open formats based on libraries of simplified elements creates the prerequisites for more flexible, transparent, and efficient data management. This also opens access to information for all participants in the

construction process - from designers to clients and operational services.

ינסו שוב לשנות את הדגש בדיון על אינטרופראביליות וגישה CAD עם זאת, יש סבירות גבוהה שבשנים הקרובות ספקי מדובר כבר על "קונספטים חדשים" - כמו נתונים גרנולריים, גרפים אינטלקטואליים, מודלים CAD למסדי נתונים של פדרטיביים, "דיגיטליים תאומים במאגרי ענן - וכן על יצירת בריתות וסטנדרטים תעשייתיים, הממשיכים את הדרך של למרות המינוח המושך, יוזמות כאלה עשויות שוב להפוך לכלים לשימור משתמשים במסגרת BIM ו-open BIM. החל (USD (Universal Scene Description אקוסיסטמות קנייניות. דוגמה אחת היא הקידום הפעיל של פורמט (BIM)-CAD משנת 2023 כ"סטנדרט חדש" לאינטראקציה בין פלטפורמות ב

ונתונים גרנולריים USD-המעבר ל

בשנת 2023 מסמנת תפנית חשובה בתעשיית הבנייה. אנו עדים להתחלה של מציאות חדשה AOUSD הופעת הברית בעבודה עם נתוני בנייה דרך כמה שינויים משמעותיים. השינוי הראשון והמשמעותי נוגע, CAD המיוצרת על ידי ספקי אנשי מקצוע המעורבים בשלבים המוקדמים של תכנון קונספטואלי מבינים יותר ויותר כי יצירת CAD לתפיסת נתוני היא רק נקודת התחלה. הנתונים הנוצרים במהלך התכנון הופכים עם הזמן לבסיס לניתוח, הפעלה CAD פרויקט בסביבת וניהול של אובייקטים. זה אומר שהם צריכים להיות זמינים ומתאימים לשימוש במערכות החורגות מגבולות הכלים CAD המסורתיים של

BIM המוביל בתעשייה, שיצר את הקונספט CAD במקביל לכך מתרחשת מהפכה בגישות של המפתחים המובילים. ספק מבצע תפנית בלתי צפויה באסטרטגיה שלו. החל משנת 2023 החברה מתרחקת מאחסון נתונים IFC, ואת פורמט בקבצים נפרדים, וממקדת את מאמציה בעבודה עם נתונים גרנולריים (נורמליזציה ומבנה) (עם מעבר לגישה מונחת נתונים

ספקים עוקבים אחרי מגמות היסטוריות של תעשיות אחרות: לרוב המשתמשים לא נדרשים פורמטים עם לוגיקת (GIMP-דומים ל) IFC או קבצים פרמטריים מורכבים של (PSD-דומים ל) CAD סגורים של אינסטגרם של) CAFM-שכבות. (הם זקוקים לתמונות פשוטות של אובייקטים שניתן להשתמש בהם ב PDF, ובמסמכי Excel) פייסבוק (ובאלפי תהליכים אחרים, המלאים בטבלאות ERP, (הבנייה

המגמות הנוכחיות בתעשיית הבנייה יוצרות פוטנציאל לנסיגה הדרגתית מפורמטים פרמטריים ומורכבים לטובת פורמטים עם מידע מטה גם בתוך פורמטים היברידיים, וגם) USD, GLTF, DAE, OBJ יותר אוניברסליים ועצמאיים כמו בפורמטים מובנים או חלשים. (מנהיגים היסטוריים, כולל חברות תכנון הגדולות ביותר, שבעבר קידמו באופן פעיל את מדגישים את הפשטות והאוניברסליות שלו. ההטמעה, USD, באמצע שנות ה-90, היום מקדמים בגלוי את הפורמט IFC ו-Blender, Unreal Engine ואינטגרציה פעילה בכלים כמו GLTF-במוצרים, התאמה ל USD ההמונית של מעידה על הפוטנציאל להתחלת פרדיגמה חדשה בעבודה עם נתונים. לצד הפופולריות של פתרונות, Omniverse, האירופיים הפופולריים, עשוי לחזק את מעמד ERP-המשמש ב, CPIXML - שטוח אירופי USD מקומיים, כמו פורמט מה, USD-כבר מתאימים את האסטרטגיה שלהם ל IFC במרכז אירופה. ארגונים העוסקים בפיתוח פורמט USD-ה-שמאשר את הבלתי נמנע של השינוי

Technical Specifications 			Comparison / Notes
File Structure	Monolithic file	Uses ECS and linked data	IFC stores all data in one file; USD uses Entity-Component-System and linked data for modularity and flexibility
Data Structure	Complex semantics, parametric geometry	Flat format, geometry in MESH, data in JSON	IFC is complex and parametric; USD is simpler and uses flat data
Geometry	Parametric, dependent on BREP	Flat, MESH (triangular meshes)	IFC uses parametrics; USD uses meshes for simplified processing.
Properties	Complex structure of semantic descriptions	Properties in JSON, easy access	Properties in USD are easier to use thanks to JSON
Export/import	Complex implementation, dependent on third-party SDKs	Easy integration, wide support	USD integrates more easily and is supported in many products
Format Complexity	High, requires deep understanding	Low, optimized for convenience	The time required to understand the structure of the file and the information stored in it.
Performance	Can be slow when processing large models	High performance in visualization and processing	USD is optimized for speed and efficiency. Simulations, machine learning, AI, smart cities will be held in the Nvidia Omniverse
Integration with 3D Engines	Limited	High, designed for graphics engines	USD excels with native support for real-time visualization platforms
Support outside CAD Software	BlenderBIM, ifcOpenShell	Unreal Engine, Unity, Blender, Omniverse	USD is widely supported in graphics tools
Cloud Technology Support	Limited	Well-suited for cloud services and online collaboration	USD is optimized for cloud solutions
Ease of Integration into Web Applications	Difficult to integrate due to size and complexity	Easy to integrate, supports modern web technologies	USD is preferable for web applications
Change Management	Versions through separate files	Versioning built into the format core	IFC handles changes via separate files, while USD embeds versioning directly into its structure
Collaboration Support	Supports data exchange between project participants	Designed for collaborative work on complex scenes	USD provides efficient collaboration through layers and variations
Learnability	Steep learning curve due to complexity	Easier to master thanks to a clear structure	USD is easier to learn and implement

USD ו-IFC השוואת מפרטי טכנולוגיה של פורמטים 6.26. ריש

עשוי להפוך לסטנדרט זה פקטור, מבטיח להתגבר על מגבלות רבות הקשורות בעיקר USD, על הרקע הזה הקיימים ותלותם בפרשנות של ליבות גיאומטריות (BIM-) CAD- למורכבות של פורמטי

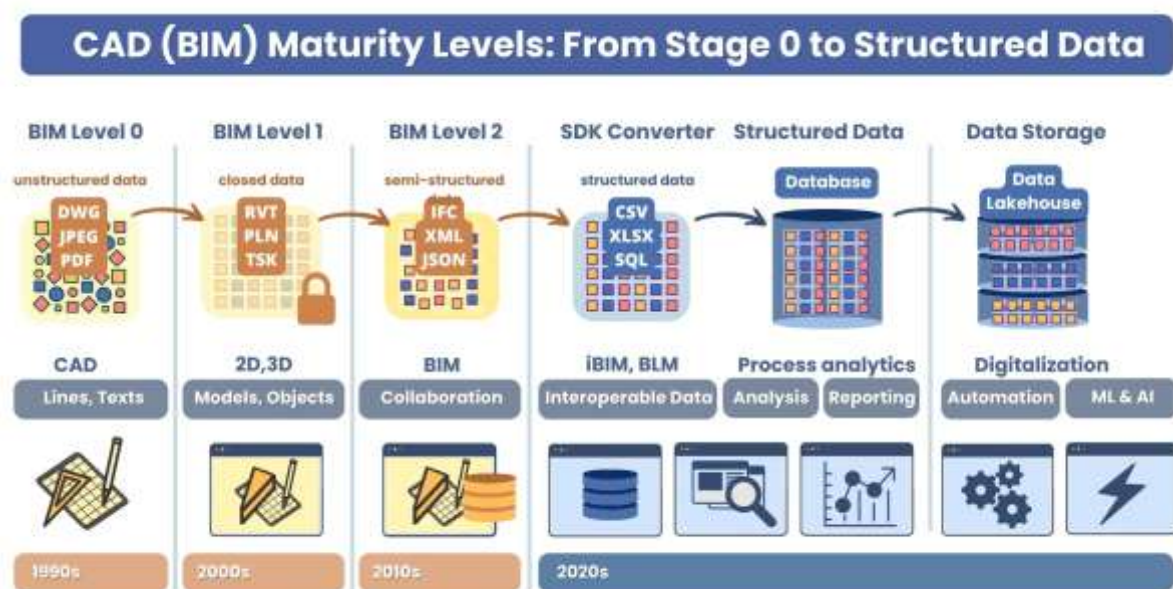
עם מידע USD, glTF, DAE, OBJ פורמטי נתונים פשוטים כמו, IFC-פרמטריים ומורכבים ו CAD במקום פורמטים יזכו למקום בתעשיית הבנייה בזכות הפשטות והגמישות שלהם CSV, XLSX, JSON, XML-מטה של רכיבים ב

המיושן ל IFC-השינויים הנוכחיים בתעשיית הבנייה נראים במבט ראשון כמו פריצת דרך טכנולוגית הקשורה למעבר מ כתב בעצמו על, IFC שפיתח את CAD המודרני יותר. עם זאת, יש לקחת בחשבון שכבר בשנת 2000, אותו ספק USD USD - הבעיות שלו ועל הצורך בגישה למסדי נתונים, ועכשיו הוא מקדם באופן פעיל את המעבר לסטנדרט החדש

ו"רעיונות" חדשים לניהול נתונים מפורטים, דרך אפליקציות USD מאחורי חזית "הנתונים הפתוחים" של עשויה להסתתר כוונה של ספקים למונופול על ניהול נתוני פרויקטים, CAD ענן שמתחילים לקדם ספקי שבהם המשתמשים מוצאים את עצמם במצב שבו הבחירה בפורמט קשורה יותר לאינטרסים תאגידיים מאשר לצרכים אמיתיים.

ניתוח של עובדות מרכזיות מראה כי המטרה העיקרית של שינויים אלה היא לאו דווקא נוחות המשתמשים, אלא בראש ובראשונה שמירה על שליטה על האקוסיסטמות וזרמי הנתונים לטובת הספקים, שלא הצליחו במשך 40 שנה לספק CAD גישה למסדי נתונים של.

ייתכן שעכשיו הגיע הזמן לחברות להפסיק לחכות לרעיונות חדשים מספקי התוכנה ולהתמקד בפיתוח עצמאי בכיוון מונחה נתונים. שחרור מהבעיות של גישה לנתונים, דרך כלים של הנדסה הפוכה, יאפשר לתעשייה לעבור באופן עצמאי לכלים מודרניים, חינוכיים ונוחים לעבודה וניתוח נתונים, מבלי להיות כפופים לרעיונות חדשים שמוכתבים.



מנתונים לא מובנים לנתונים מובנים ולאחסון: CAD (BIM) רמת בשלות 6.27. ריש

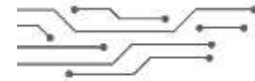
גישה למסדי נתונים, נתונים פתוחים ופורמטים יהפכו ללא ספק לסטנדרט בתעשיית הבנייה, ללא קשר לניסיונות הספקים לעכב את התהליך הזה - זו רק שאלה של זמן. קצב המעבר הזה עשוי לגדול משמעותית אם יותר ויותר אנשי זמינים להנדסה הפוכה, המאפשרים גישה SDKs-מקצוע יכירו פורמטים פתוחים, כלים לעבודה עם מסדי נתונים ו-CAD. ישירה לנתונים של מערכות.

העתיד הוא בנתונים פתוחים, מאוחדים ונגישים לניתוח. כדי להימנע מהתלות בפתרונות של ספקים ולא למצוא את עצמם כבולים לאקוסיסטמות סגורות, חברות בנייה ועיצוב יאלצו בשלב כלשהו להמר על פתיחות ועצמאות, בבחירת פורמטים ופתרונות המבטיחים שליטה מלאה על הנתונים.

הנתונים שנוצרים היום בתעשיית הבנייה יהפכו למשאב מרכזי לקבלת החלטות עסקיות בעתיד. הם ישמשו כדלק

אסטרטגי, המניע את ההתפתחות והיעילות של חברות הבנייה. העתיד של תעשיית הבנייה טמון ביכולת לעבוד עם נתונים ולא בבחירת פורמטים או מודלים של נתונים.

CAD לבין פורמטים פרופירטוריים של USD, glTF, DAE, OBJ כדי להבין את ההבדל בין פורמטים פתוחים כמו פרמטרי, חשוב לשקול אחד מהאלמנטים המורכבים והחשובים ביותר בנתונים בהדמיה ובחישובי פרויקטים הגיאומטריה ותהליכי היווצרותה. כדי להבין כיצד נתוני גיאומטריה הופכים לבסיס לניתוח ולחישובים בבנייה, יש צורך לחקור לעומק את המנגנונים של יצירת גיאומטריה, המרתה ואחסונה.



פרק 6.3. גיאומטריה בבנייה: מקווים לקוביות

מתי קווים הופכים לכסף או למה גיאומטריה חשובה לבונים

גיאומטריה בבנייה היא לא רק הדמיה, אלא גם הבסיס לחישובים כמותיים מדויקים. במודל הפרויקט, הגיאומטריה עם מאפיינים נפחיים חשובים, כמו אורך, שטח ונפח (3.116 איור) משלימה את רשימות הפרמטרים של האלמנטים ערכים אלו של פרמטרים נפחיים מחושבים אוטומטית בעזרת ליבות גיאומטריות ומהווים את נקודת המוצא עבור אומדנים, לוחות זמנים ודגמי משאבים. כפי שכבר דנו בחלק החמישי של הספר ובפרק "חישובי עלות ואומדנים של 1-ERP מהווים את הבסיס עבור מערכות CAD פרויקטים בנייה, "פרמטרים נפחיים של קבוצות אובייקטים ממודלי מודרניות. הגיאומטריה משחקת תפקיד יסודי לא רק בשלב התכנון, אלא גם בניהול ביצוע הפרויקט, בבקרת PMIS לוחות זמנים, בתקצוב ובתפעול. כמו לפני אלפי שנים בבניית הפירמידות המצריות, הדיוק של הפרויקט תלוי במדדי משפיע ישירות על התוצאה: מהתקציב CAD אורך כמו אמה וקוביט, כך גם היום הדיוק בפרשנות הגיאומטריה בתוכנות - והזמנים ועד לבחירת קבלנים ולוגיסטיקת אספקה

בתנאי תחרות גבוהה ותקציב מוגבל, הדיוק של חישובי הנפח, התלוי ישירות בגיאומטריה, הופך לגורם CAD מודרניות תלויות ישירות במאפיינים נפחיים מדויקים המתקבלים ממודלי ERP הישרדות. מערכות לכן, תיאור גיאומטרי מדויק של אלמנטים הוא לא רק הדמיה, אלא כלי מפתח בניהול עלויות. BIM-1 וזמנים בבנייה.

- היסטורית, גיאומטריה הייתה השפה העיקרית של אינטראקציה הנדסית. מהקווים על פפירוס ועד מודלים דיגיטליים שרטוטים וייצוגים גיאומטריים שימשו כאמצעי להעברת מידע בין מעצבים, מנהלי עבודה ומערכי עלויות. לפני הופעת המחשבים, חישובים בוצעו ידנית, בעזרת סרגלים ומחוגות. היום, משימה זו אוטומטית הודות למידול נפחי: ליבות ממירות קווים ונקודות לגופים תלת-ממדיים, מהם נשלפות אוטומטית כל המאפיינים CAD גיאומטריות של תוכנות הנדרשים.

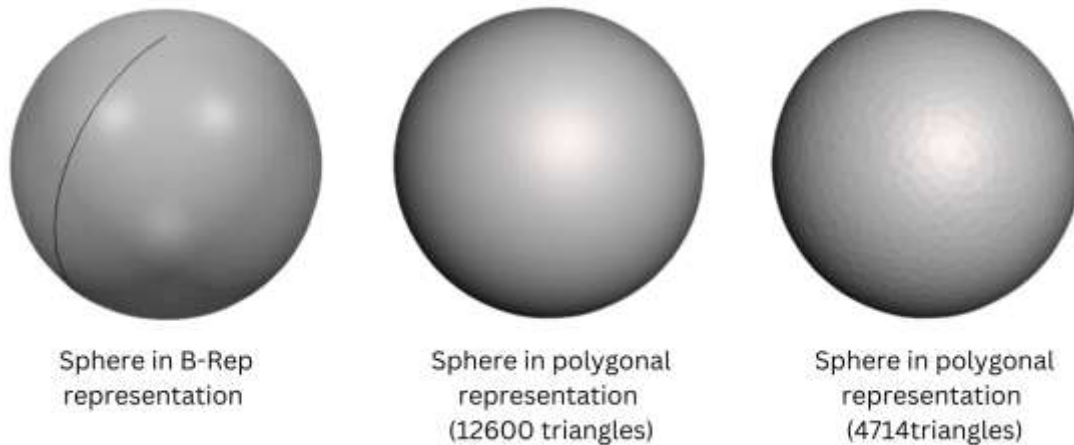
CAD יצירת אלמנטים גיאומטריים לחישובים מתבצעת דרך ממשק המשתמש של תוכנות CAD, בעבודה בתוכנות להמרת נקודות וקווים לגופים נפחיים, נעשה שימוש בליבה גיאומטרית, המבצעת את המשימה המרכזית - המרת (BIM). גיאומטריה למודלים נפחיים, מהם לאחר האפרוקסימציה מחושבים אוטומטית המאפיינים הנפחיים של האלמנט

מקווים לנפחים: כיצד שטח ונפח הופכים לנתונים

בפרקטיקה ההנדסית, נפחים ושטחים מחושבים על בסיס משטחים גיאומטריים, המתוארים אנליטית או דרך מודלים (ייצוג גבולות של אלמנטים) BREP במסגרת (רציונליים לא אחידים-B ספליינים) NURBS פרמטריים, כמו

זו שיטה מתמטית לתיאור קווים ומשטחים, בעוד ש — (רציונליים לא אחידים B-Splines) NURBS היא מבנה לתיאור גיאומטריה תלת-ממדית מלאה של אובייקט, כולל גבולותיו, שניתן להגדירם BREP. NURBS באמצעות

הם דורשים משאבים חישוביים חזקים ואלגוריתמים מורכבים. עם זאת, חישובים, NURBS ו-BREP למרות הדיוק של ישירים על תיאורים מתמטיים מדויקים כאלה הם לעיתים קרובות מורכבים מבחינה חישובית, ולכן בפועל כמעט תמיד נעשה שימוש בטסלציה – המרה של משטחים לרשתות של משולשים, מה שמפשט את החישובים הבאים. טסלציה היא שיטה זו משמשת להמחשה, חישובי נפחים, CAD/CAE פיצול של משטח מורכב למשולשים או פוליגונים. בסביבת וניתוח התנגשויות. דוגמה מהטבע – דבורי דבש, שבהן הצורה MESH חיפוש התנגשויות, ייצוא בפורמטים כמו - המורכבת מפוצלת לרשת סדירה) איור 6.31



ובייצוג פוליגוני עם מספר שונה של משולשים BREP בתיאור פרמטרי esfera איור 6.31 אותה.

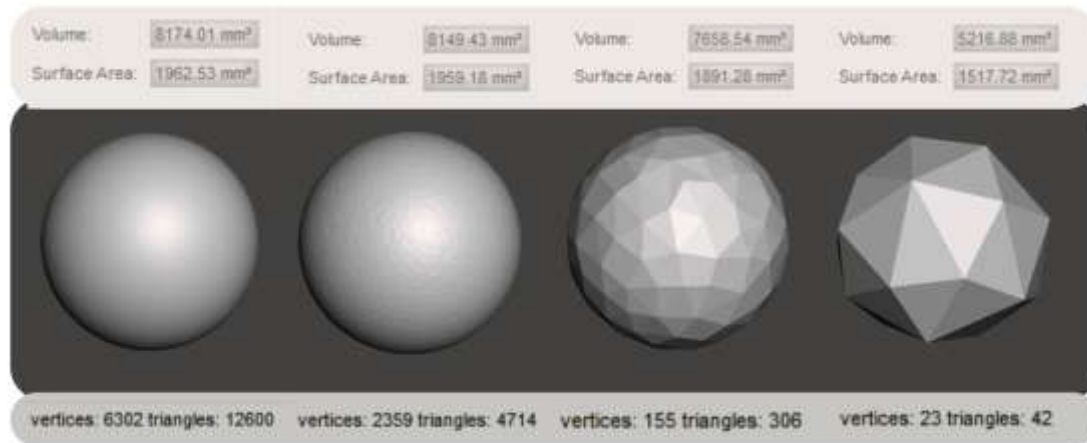
אינה מודל גיאומטרי בסיסי. שיטה זו נוצרה ככלי נוח לייצוג מעגלים וספליינים, CAD-המיושמת ב, BREP (NURBS) רציונליים ולמזעור אחסון נתונים על גיאומטריה. עם זאת, יש לה מגבלות – לדוגמה, חוסר יכולת לתאר בדיוק סינוסואידה, שהיא הבסיס לקווים ולמשטחים ספירליים, וכן הצורך בשימוש בליבות גיאומטריות מורכבות.

רשתות משולשות וטסלציה של דמויות פרמטריות, לעומת זאת, מתאפיינות בפשטות, בשימוש יעיל בזיכרון וביכולת, לעבד כמויות גדולות של נתונים) איור 6.32. (יתרונות אלו מאפשרים להימנע מליבות גיאומטריות מורכבות ויקרות, עם עשרות מיליוני שורות קוד, בחישובי צורות גיאומטריות

ברוב המקרים הבנייתיים אין זה משנה כיצד מוגדרות התכונות התלת-ממדיות – באמצעות מודלים הגיאומטריה נשארת. (USD, glTF, DAE, OBJ) או באמצעות פוליגונים (BREP, IFC) פרמטריים תמיד מדובר בתיאור משוער של הצורה, MESH או NURBS צורת קירוב: בין אם באמצעות

נשארת במידה מסוימת רק דרך קירוב עם תיאור משוער של BREP (NURBS)-גיאומטריה המוגדרת בפוליגונים או ב צורה רציפה. בדומה לכך, אינטגרלים של פרנל אינם בעלי ביטוי אנליטי מדויק, דיסקרטיזציה של גיאומטריה באמצעות משולש MESH תמיד היא קירוב, כמו גם NURBS פוליגונים או

נחוצה בעיקר שם שחשוב גודל מינימלי של נתונים ויש אפשרות להשתמש בליבות BREP גיאומטריה פרמטרית בפורמט אשר משתמשים, CAD גיאומטריות יקרות ודורשות משאבים לעיבוד והצגה שלה. לרוב זה מאפיין מפתחים של תוכנות בתהליך BREP במוצרים שלהם. עם זאת, אפילו בתוך תוכנות אלו, מודלים של MCAD בליבות גיאומטריות של ספקי (JPEG-מפושטות ל PSD בדומה לכך, קבצי) הטסלציה להמחשה ולחישובים לעיתים קרובות מומרות למשולשים



איור 6.32 ההבדל בתכונות תלת-ממדיות בין דמויות עם מספר שונה של פוליגונים.

הפרמטרי, יש להם יתרונות ומגבלות, אך המטרה שלהם היא אחת - לתאר את BREP-המשטח הפוליגוני, כמו גם ה-גיאומטריה בהתחשב בצרכי המשתמש. בסופו של דבר, הדיוק של המודל הגיאומטרי תלוי לא רק בשיטת הייצוג שלו אלא גם בדרישות של המשימה הספציפית.

ברוב המשימות הבנייתיות, הצורך בגיאומטריה פרמטרית ובגרעינים גיאומטריים מורכבים עשוי להיות מיותר.

בכל משימה ספציפית של אוטומציה של חישובים, יש לבחון האם לא מוגזמת החשיבות של הגיאומטריה הפרמטרית על שמעוניינים לקדם ולמכור את המוצרים התוכנתיים שלהם, CAD, ידי המפתחים של.

ופוליגונים: שימוש בטסלציה לגיאומטריה MESH, USD-המעבר ל

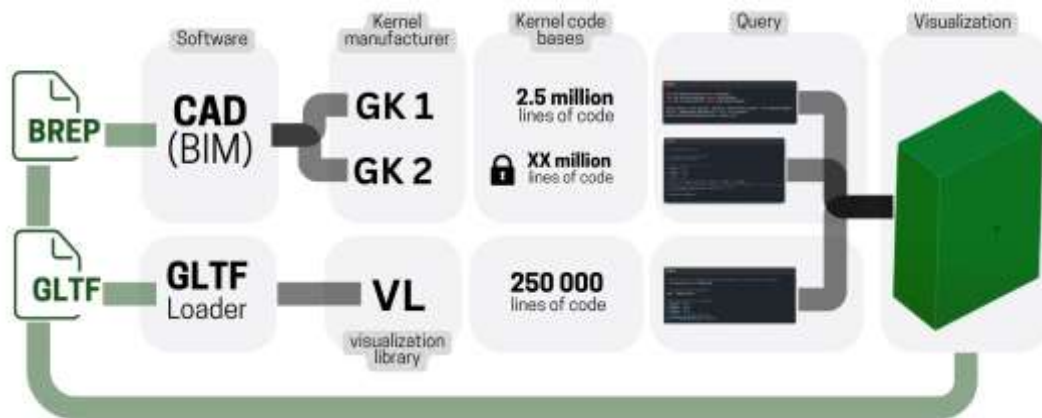
בתעשיית הבנייה, בעבודה רציפה, בפיתוח מערכות, בבסיסי נתונים או באוטומציה של תהליכים העובדים עם מידע ומהגרעינים הגיאומטריים CAD פרויקט וגיאומטריה של רכיבים, חשוב לשאוף לעצמאות מהעורכים הספציפיים של.

בבסיס פורמט ההחלפה, שייעשה בו שימוש הן במחלקות החישוב והן באתר הבנייה - לא צריכה להיות ספציפית. המידע הגיאומטרי צריך להיות מוצג בפורמט ישירות דרך טסלציה, ללא CAD (BIM) תוכנת CAD. תלות בגרעין גיאומטרי או בארכיטקטורת.

יכולה להיחשב כמקור ביניים, אך לא כבסיס לפורמט אוניברסלי. רוב התיאורים CAD-הגיאומטריה הפרמטרית מכל מקרה מומרות למשטח פוליגוני לצורך עיבוד מאוחר יותר. אם התוצאה (NURBS ו-BREP כולל) הפרמטריים היא אותה תוצאה) טסלציה ופוליגונים, (וההליך הוא פשוט יותר, הבחירה ברורה. זה דומה לבחירה בין אונטולוגיה גרפית --rarely justified לטבלאות מובנות) על כך דיברנו בחלק הרביעי: (מורכבות מיותרת

משתמשים במבנה אוניברסלי של רשתות OBJ, STL, glTF, SVF, CPIXML, USD ו-DAE: פורמטים פתוחים כמו משולשות, מה שנותן להם יתרונות משמעותיים. פורמטים אלה מציעים תאימות מצוינת - קל לקרוא ולוויזואליזציה

בעזרת ספריות פתוחות זמינות ללא צורך בגרעינים גיאומטריים מורכבים, המכילים מיליוני שורות קוד. פורמטים ועד IKEA™-גיאומטריים אוניברסליים אלה משמשים בתחומים שונים - מכלים יחסית פשוטים לתכנון מטבחים ב יתרון חשוב הוא קיומם של מספר רב של VR. מערכות מורכבות של ויזואליזציה של אובייקטים בקולנוע וביישומי .ספריות פתוחות וחינמיות לעבודה עם פורמטים אלה, זמינות לרוב הפלטפורמות ושפות התכנות



אותה ייצוג גיאומטרי מושג באמצעות שימוש בפורמטים פרמטריים וגרעינים גיאומטריים, או באמצעות פורמטים 6.33 התמונה - משולשים וספריות ויזואליזציה פתוחות

פתוח IFC זרים או CAD מתמודדים עם בעיות בפרשנות פורמטים פרמטריים של CAD כמו המשתמשים עצמם, ספקי של הנדסה SDK-ללא יוצא מן הכלל, משתמשים ב CAD, בשל ההבדלים בגרעינים הגיאומטריים. בפועל, כל ספקי למטרות אינטרופראביליות USD או IFC הפוכה להעברת נתונים בין מערכות, ואף אחד מהם לא סומך על פורמטים כמו

שאותם הם עצמם לא משתמשים, יהיה CAD, במקום להשתמש בקונספטים שמקדמים בריתות של ספקי להתמקד בהבנת היתרונות של כל גישה בהקשר ספציפי CAD מועיל יותר למפתחים ולמשתמשי פתרונות ולבחור את סוג הגיאומטריה בהתאם למקרה השימוש. הבחירה בין ייצוגים גיאומטריים שונים היא פשרה בין דיוק, יעילות חישובית וצרכים מעשיים של המשימה הספציפית

המורכבות הקשורה בשימוש בליבות גיאומטריות, שמספקים ספקים גדולים לתעשיית הבנייה בעת עיבוד נתוני פרויקטים, יכול להפוך לסוג של "תיבת MESH המבוסס על גיאומטריית USD, לעיתים קרובות מתבררת כמיותרת. פורמט ומבני IFC פנדורה "עבור התעשייה, פותח בפני מפתחים הזדמנויות חדשות לארגון חילופי נתונים - מחוץ למסגרות CAD. פרמטריים, האופייניים לספקי BREP

ועוד, מתברר שישנם פורמטים פתוחים ופשוטים יותר USD, DAE, glTF, OBJ כאשר מתעמקים במבנה של המאפשרים לארגן בצורה יעילה את העברת השימוש במידע גיאומטרי מבלי להסתמך על פרמטריקה מורכבת וליבות, גיאומטריות סגורות. גישה כזו לא רק מפחיתה את הסף הטכני עבור מפתחים, אלא גם תורמת לפיתוח פתרונות גמישים ניתנים להרחבה ופתוחים באמת לבנייה דיגיטלית

LOD, LOI, LOMD – סיווג ייחודי של פירוט ב-CAD (BIM)

בנוסף לפורמטים הגיאומטריים, בעולם שבו תעשיות שונות משתמשות ברמות שונות של פירוט ועומק נתונים, מציעות מערכות סיווג ייחודיות, המארגנות את הגישה למילוי המידע במודלים של בניינים CAD (BIM) מתודולוגיות.

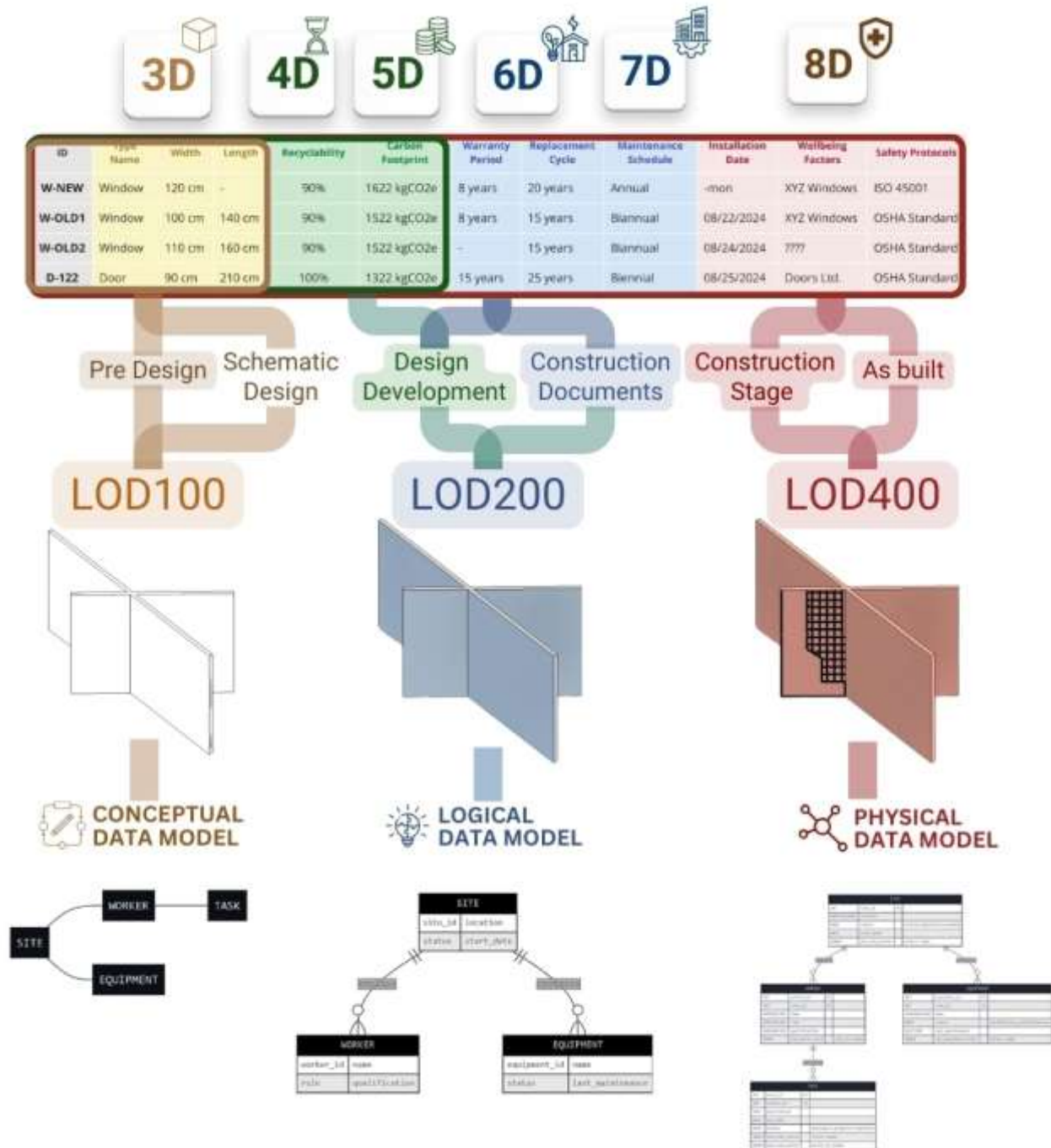
דוגמה אחת לגישות חדשות לסטנדרטיזציה היא הכנסת רמות פיתוח מודל, המשקפות את דרגת המוכנות והאמינות של LOD (Level Of Detail) הוצג CAD (BIM) החלק הגרפי והמידע. כדי להבחין בין רמות המילוי המידע בעבודה עם נתוני רמת הפיתוח של הנתונים - LOI (Level Of Information) - רמת הפירוט של החלק הגרפי של המודל, ו- (Detail) - LOG (Level of Accuracy) - לוגיקה האלמנטים המוצגים ו- LOA (Level of Accuracy) בנוסף, לצורך גישה כוללת הוצג המושג לקביעת דיוק הייצוג הגרפי (of Geometry).

מייצגת LOD 100. מסומנות במספרים מ-100 עד 500, המשקפים את דרגת הפיתוח של המודל (LOD) רמות הפירוט LOD. כוללת מידות וצורות מדויקות יותר, אך עם פירוט מותנה LOD 200. מודל קונספטואלי עם צורות וגדלים כלליים מכילה מידע מפורט הנדרש LOD 400. היא מודל מפורט עם מידות, צורות ומיקומים מדויקים של אלמנטים 300 משקפת את המצב האמיתי של האובייקט לאחר הבנייה ומשמשת לתפעול LOD 500. לייצור ולהתקנה של אלמנטים במידע בשלבים שונים של מחזור החיים, כולל CAD (BIM) ולתחזוקה. רמות אלו מתארות את מבנה המילוי של מודל 3D, 4D, 5D ועוד.

LOD100 לעיתים קרובות מתברר כמיותר ודי להשתמש בגיאומטריה (LOD400) בפרויקטים אמיתיים, פירוט גבוה או אפילו בשרטוטים שטוחים, בעוד שיתר הנתונים ניתן להשיג או בדרך חישובית או מאלמנטים מקושרים, שעשויים לא להיות בעלי גיאומטריה ברורה. לדוגמה, חללים ואלמנטים של חדרים (קטגוריות של אלמנטים "חדרים") ("עשויים לא להיות בעלי גיאומטריה ויזואלית, אך עדיין להכיל כמויות משמעותיות של מידע ומסדי נתונים, שסביבם נבנים רבים מתהליכי העסקים).

לכן, לפני תחילת התכנון, חשוב להגדיר בצורה ברורה את רמת הפירוט הנדרשת. עבור מקרים של שימוש המשימה. LOD100 וגיאומטריה מינימלית DWG לעיתים קרובות מספיקים אפילו שרטוטי 4D-7D, המרכזית בתהליך העבודה עם הדרישות היא למצוא את האיזון בין העושר לבין השימושים המעשית של המודל.

כבסיס נתונים) כפי שהם באמת, (תיאור העושר של המודל באמצעות ראשי CAD (BIM) למעשה, אם נבחן את נתוני תיבות חדשים מהווה לא יותר מאשר מודל נתונים בשלבים עבור מערכות מידע, החל מרמה קונספטואלית ועד לרמה משמעותה הוספת LOI-וב LOD-פיזית) איור 6.34, (שנדונה בפירוט בחלקים השלישי והרביעי של הספר. כל עלייה ב D-מידע הנדרש למשימות חדשות: חישובים, ניהול בנייה, תפעול, ומאופיינת בהעשרת המודל בשכבות מידע נוספות 3 בצורת פרמטרים שונים, עליהם דיברנו בחלק החמישי של הספר (8D).



איור 6.34 תהליך ההעשרה המידע של הפרויקט דומה למידול נתונים מרמה קונספטואלית לרמה פיזית של נתונים.

גיאומטריה היא רק חלק מסוים מנתוני התכנון, הצורך בה לא תמיד מוצדק בפרויקטים בנייה, והשאלה המרכזית בעבודה היא לא כל כך איך המודלים מוויזואליזציה, אלא יותר איך הנתונים מהמודלים הללו יכולים לשמש מעבר CAD עם נתוני (BIM) CAD לתוכנות.

בסביבות אמצע שנות ה-2000, תעשיית הבנייה נתקלת בבעיה חסרת תקדים, הקשורה לעלייה המהירה בכמות הנתונים העלייה החדה בכמות הנתונים תפסה את (BIM) CAD במערכות ניהול ועיבוד נתונים, במיוחד אלו שהגיעו ממחלקות

המנהלים בחברות לא מוכנים, והם לא היו מוכנים לדרישות הגוברות לאיכות הנתונים ולניהולם.

CAD (BIM) - AIA, BEP, IDS, LOD, COBie - תקנים חדשים ל

ולתחרות המוגבלת בשוק עיבוד הנתונים, כמו גם באמצעות CAD תוך כדי ניצול חוסר הגישה הפתוחה לבסיסי נתונים - CAD ארגונים העוסקים בפיתוח גישות לעבודה עם נתוני BIM, קמפיינים שיווקיים הקשורים לראשי תיבות החדשים. החלו ליצור סטנדרטים חדשים ורעיונות, אשר דה-יורה אמורים להיות מכוונים לשיפור הפרקטיקה של ניהול הנתונים.

CAD (BIM), למרות שכמעט כל היוזמות, הנתמכות במישרין או בעקיפין על ידי ספקים ומפתחים של היו מכוונות לאופטימיזציה של תהליכי העבודה, הן הביאו להופעתם של סטנדרטים רבים, המיוצגים על ידי בעלי עניין שונים, מה שהוביל את תעשיית הבנייה למצב של עמימות ובלבול בתהליכי עיבוד הנתונים.

שהופיעו בשנים האחרונות בתעשיית LOD, LOI, LOA, LOG, נציין כמה מהסטנדרטים החדשים לנתונים, מלבד הבנייה:

■ בפרויקט, תוך הגדרת שיטות CAD (BIM) - מתארת כיצד לשלב ולהשתמש ב - (BIM תוכנית ביצוע) BEP ותהליכי עיבוד נתונים.

■ (דרישות מידע של המזמין) - נערך על ידי המזמין לפני פרסום המכרז ומכיל דרישות EIR/AIA מסמך בפרויקט המתאים BEP - מהקבלן בנוגע להכנה והגשה של מידע. הוא משמש כבסיס ל

■ לאחר מסירת והשלמת הפרויקט, מודל הנתונים נקרא BIM) מודל מידע של נכס - חלק מתהליך AIM. היא ניהול, תחזוקה ותפעול של הנכס המיושם AIM המטרה של AIM. מודל מידע של נכס "או"

■ (מפרט אספקת מידע) - מגדיר את הדרישות ואת סוגי הנתונים ובאיזה פורמט הם נדרשים בשלבים IDS שונים של פרויקט בנייה.

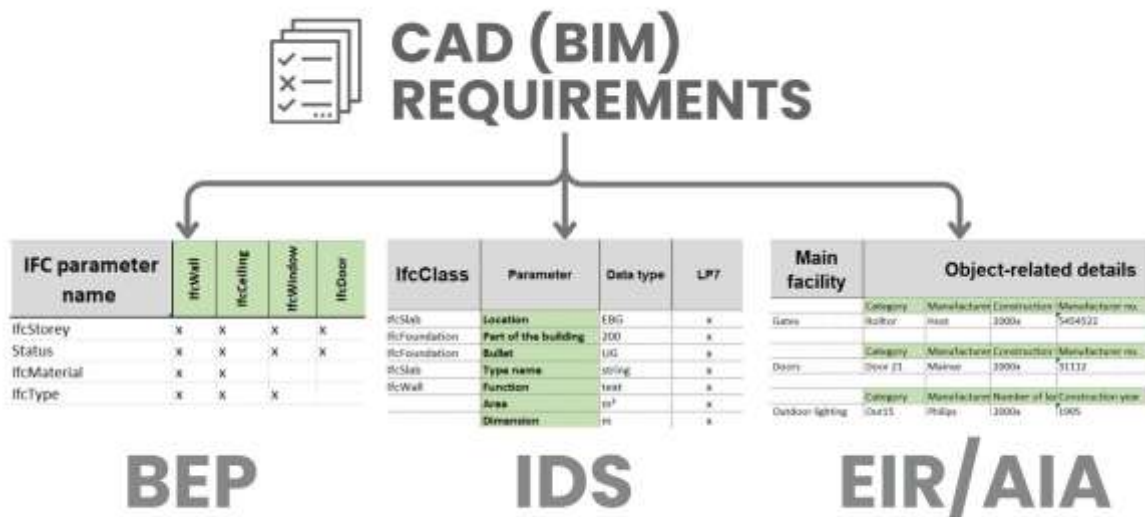
■ הוא מגדיר עד כמה מפורט ומלא המידע. BIM שבה המידע מוצג במודל, LOD רמת הפירוט - iLOD במודל, החל מייצוגים גיאומטריים בסיסיים ועד מפרטים מפורטים ונתונים.

■ הוא מגדיר את דרגת המידול של כל CAD (BIM) של רכיבים בודדים במודל LOD רמת הפירוט - eLOD רכיב ואת המידע הקשור אליו, כגון מידות, חומרים, תכונות שימושיות ואטריבוטים רלוונטיים אחרים.

■ מתארים את הכלים והתשתיות - CAD (BIM) שירותי פלטפורמה (ומוצרים אחרים מספקים גדולים APS הנדרשות ליצירת מודלים נתונים מקושרים ופתוחים.

AIM, AIA, EIR, BEP, LOD, LOI, LOA, LOG, כמו - CAD (BIM) אף על פי שהמטרה המוצהרת של יישום תקני היא לשפר את איכות ניהול הנתונים ולהרחיב את אפשרויות האוטומציה, בשטח השימוש בהם - iLOD, eLOD, IDS כסוג של בסיס נתונים, מתברר CAD (BIM) לעיתים קרובות מביא ליתר מורכבות ולפיצול תהליכים. אם נביט על מודל כי רבים מהתקנים הללו משכפלים גישות שהיו קיימות ויעילות כבר זמן רב בתחומים אחרים בכלכלה בעבודה עם מערכות מידע. במקום לפשט ולאחד, יוזמות כאלה לעיתים קרובות יוצרות עומס טכנולוגי נוסף ומקשות על יישום פתרונות באמת פתוחים וגמישים.

מעניין, רבות מהקונספטים החדשים הללו למעשה מחליפים את תהליך המידול ובדיקת הנתונים, שנדונו בפירוט בחלקים הראשונים של הספר וכבר זמן רב בשימוש בתחומים אחרים בכלכלה. בתחום הבנייה, תהליך הסטנדרטיזציה לעיתים קרובות מתקדם בכיוון ההפוך – נוצרים פורמטים חדשים לתיאור נתונים, תקנים חדשים וקונספטים חדשים לבדיקתם שלא תמיד מובילים לאיחוד אמיתי וליישום מעשי. כתוצאה מכך, במקום לפשט ולאטומט את העיבוד, התעשייה נתקלת ברמות נוספות של רגולציה ובירוקרטיה (איור 6.31, מה שלא תמיד תורם לשיפור היעילות).



איור 6.31 דרישות לנתונים ולתוכן המידע מצטמצמות לתיאור של אטריבוטים והערכים הגבוליים שלהם, המתוארים באמצעות טבלאות.

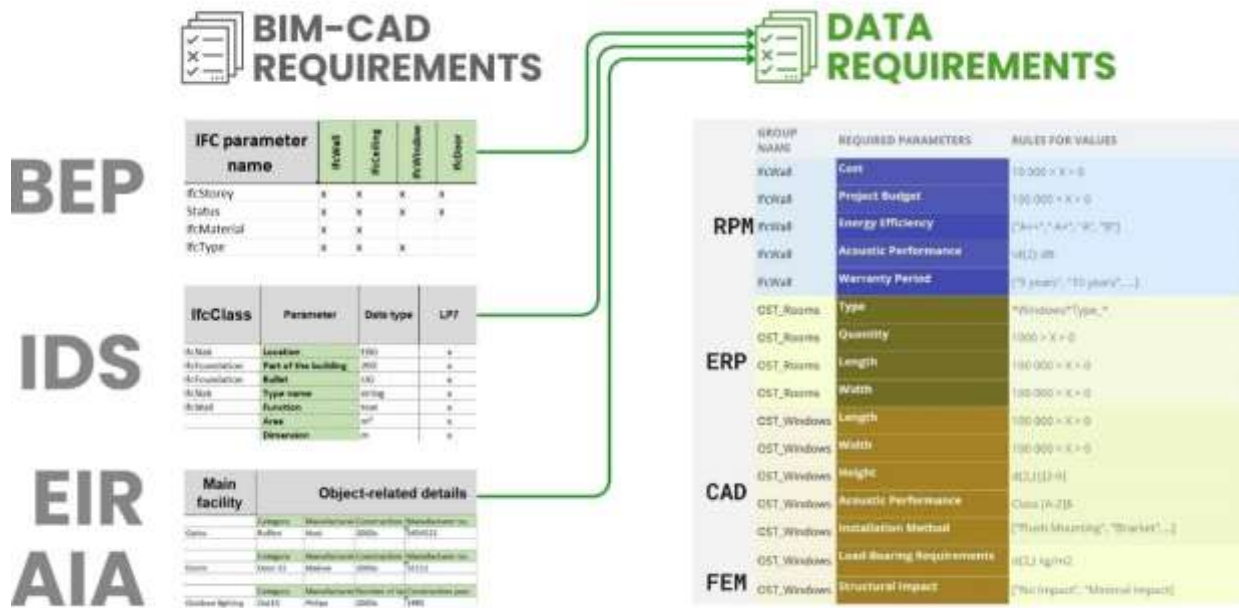
לעיתים קרובות CAD (BIM) במקום לפשט את עיבוד הנתונים, הקונספטים החדשים הקשורים לנתוני יוצרים מורכבויות נוספות ומחלוקות כבר בשלב הפרשנות וההגדרות הבסיסיות.

(שצץ בשנת 2020) (המאפשר לתאר את הדרישות IDS אחד מהדוגמאות האחרונות לקונספטים חדשים הוא פורמט מתארות מידע על אטריבוטים והערכים IDS דרישות open BIM. להרכב האטריבוטים של המודל המידע בקונספט אשר לאחר מכן מתורגמת לסימון בפורמט חצי-מובנה (MySQL או Excel) הגבוליים שלהם בצורה של טבלה מסודרת IDS. לאקראי XML-ששונה מ XML.

כי העבודה עם open BIM ו-BIM בניגוד לדעה המקדמת על ידי ספקים ונתמכת על ידי קונספטים של פורמטים של נתונים, CAD ו-BIM נתונים בבנייה היא ייחודית בשל השימוש בכלים מיוחדים, כמו ושיטות ניהול נתונים בתחום זה אינם שונים מהפורמטים והקונספטים של עיבוד נתונים בתחומים אחרים.

באמצעות טבלה אחידה של דרישות עם עמודות CAD (BIM) ניתן לפשט את כמות הדרישות לפרויקטים ולפורמטים של של תכונות, המתוארת בפירוט בפרק "המרת דרישות לצורת מבנה", מבלי הצורך להמיר דרישות מובנות מראש (מתואר מלכתחילה באמצעות טבלה IDS) לפורמטים שאינם טבלאיים.

הכוללת עמודות עבור מזהי ישויות, תכונות וערכי גבול, שנדונו בפירוט בפרקים הקודמים, (6.32 איור) הגישה המפורשת שיטה זו מספקת IDS-XML מאפשרת להימנע מהמרת דרישות לפורמט, (7.310 איור, 4.416 איור, 4.49 איור) מנגנון ישיר, פחות מסורבל ויותר שקוף לבקרת איכות הנתונים. היא מתבססת על כלים בשימוש נרחב: החל מביטויים סטנדרטיים – בדיוק כאלה שמשתמשים ETL וצינורות Pandas ועד מסגרות נתונים, ספריית (RegEx) רגולריים ---בהם אנשי מקצוע בתחומים אחרים של הכלכלה בעבודתם עם נתונים



דרישות הנתונים בתעשיות אחרות ממושטות לתיאור מובנה של תכונות וערכי גבול שלהן 6.32 איור.

עם הזמן, בתעשיית הבנייה, בשל סגירות הנתונים, צצים יותר ויותר גישות ושיטות חדשות לבקרת וניהול הנתונים המגוונים הללו, אף על פי שהנתונים בפרויקטים של בנייה, בעצם, אינם שונים מנתונים בתחומים אחרים. בעוד שתחומים אחרים מסתדרים בהצלחה עם גישות סטנדרטיות לעיבוד נתונים, הבנייה ממשיכה לפתח פורמטים ייחודיים חדשים. דרישות ומושגים לבדיקתן

השיטות והכלים המשמשים לאיסוף, הכנה וניתוח נתונים בבנייה, אינם צריכים להיות שונים באופן עקרוני מאלו המיועדים לאנשי מקצוע בתחומים אחרים של הכלכלה.

בתעשייה התפתחה אקוסיסטמה של מונחים מיוחדים, אשר דורשת חשיבה ביקורתית והערכה מחדש

תוך הוספת קטגוריזציה בנייתית, מבלי להתחשב במגבלות של פורמט IFC, מוצג תחת שם חדש STEP פורמט עצמו STEP.

הפרמטרי מיושם בתהליכי העברת נתונים, למרות חוסר קיום של ליבת גיאומטריה מאוחדת IFC פורמט הנדרשת להמחשה ולחישובים.

מבלי לדון בפרטים של בסיסי נתונים, "BIM" מקודמת תחת המונח CAD הגישה לבסיסי נתונים של מערכות אלו ובגישה אליהם.

לעיתים קרובות מבלי ליישם אותם בפועל, תוך, IFC ו-USD ספקים מקדמים אינטרופראביליות דרך פורמטים

שימוש בהנדסה הפוכה יקרה, נגד אשר הם עצמם נאבדו.

משמשים בכל מקום LOD, LOI, LOA, LOG, BEP, EIR, AIA, AIM, IDS, iLOD, eLOD המונחים לתיאור אותם פרמטרים של ישויות, מבלי לקשרם לכלים של מודל וולידציה, המיושמים כבר זמן רב בתחומים אחרים.

תעשיית הבנייה מראה שכל מה שצוין, אף על פי שזה נשמע מוזר, אפשרי – במיוחד אם המטרה העיקרית היא מונטיזציה של כל שלב בעיבוד הנתונים דרך מכירת שירותים מיוחדים ותוכנה. מבחינת עסקים אין בכך דבר פסול. עם זאת, השאלה מוסיפים ערך ומפשטים את התהליכים המקצועיים, נותרת CAD (BIM)-האם אכן ראשי התיבות והגישות הקשורות ל פתוחה.

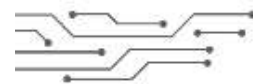
בתעשיית הבנייה מערכת כזו פועלת, מכיוון שהאינדוסטריה עצמה מפיקה את הרווח הספקולטיבי העיקרי בדיוק במבוכים של מערכות ואקראי. חברות שמעוניינות בתהליכים שקופים ובנתונים פתוחים נפגשות, לעיתים רחוקות. המצב המורכב הזה עשוי להימשך זמן בלתי מוגבל - עד שהלקוחות, הלקוחות המשקיעים, הבנקים ונציגי ההון הפרטי לא יתחילו לדרוש גישות ברורות ומבוססות יותר לניהול מידע.

התעשייה צברה כמות מופרזת של אקראי, אך כולם מתארים במידה שונה את אותם תהליכים ודרישות לנתונים. התועלת האמיתית שלהם בפשטות תהליכי עבודה נשארת שאלה פתוחה.

בזמן שהקונספטים ואקראי השיווק מגיעים והולכים, התהליכים עצמם של בדיקת דרישות לנתונים יישארו תמיד חלק בלתי נפרד מתהליכי העסקים. במקום ליצור כל הזמן פורמטים ותקנות מיוחדות חדשות, תעשיית הבנייה צריכה להסתכל על כלים שכבר הוכיחו את יעילותם בתחומים אחרים, כמו פיננסים, תעשייה וטכנולוגיית מידע.

שפע המונחים, האקראי והפורמטים יוצר אשליה של עיבוד מעמיק של תהליכי הבנייה הדיגיטלית. עם זאת, מאחורי הקונספטים השיווקיים והטרמינולוגיה המורכבת לעיתים קרובות מסתתרת אמת פשוטה אך לא נוחה: הנתונים נשארים קשים להשגה, מתועדים בצורה גרועה ומחוברים באופן נוקשה לפתרונות תוכנה ספציפיים.

לא ככלים קסומים CAD (BIM) כדי לצאת ממעגל האקראי והפורמטים למען הפורמטים, יש להסתכל על מערכות לניהול מידע, אלא כפי שהן באמת - בסיסי נתונים מיוחדים. ורק דרך הפריזמה הזו אפשר להבין היכן מסתתרים השיווק ומתחילה העבודה האמיתית עם המידע.



פרק 6.4. CAD לעבודה עם LLM-פרמטריזציה של תכנון ושימוש ב

הדרך לניתוח ופורמטים פתוחים: CAD (BIM) אשליית הייחודיות של נתוני

המודרניות שינו באופן משמעותי את הגישה לתכנון ולניהול מידע בנייה. אם בעבר הכלים (BIM) CAD פלטפורמות הללו שימשו בעיקר ליצירת שרטוטים ודגמים תלת-ממדיים, כיום הם מבצעים פונקציות של מחסני נתונים פרויקטיים מלאים. במסגרת הקונספט של מקור אמת יחיד, המודל הפרמטרי הולך והופך יותר ויותר למקור המידע העיקרי ולעיתים קרובות למקור המידע היחיד על הפרויקט, ומבטיח את שלמותו ואת עדכניותו לאורך כל מחזור חיי האובייקט.

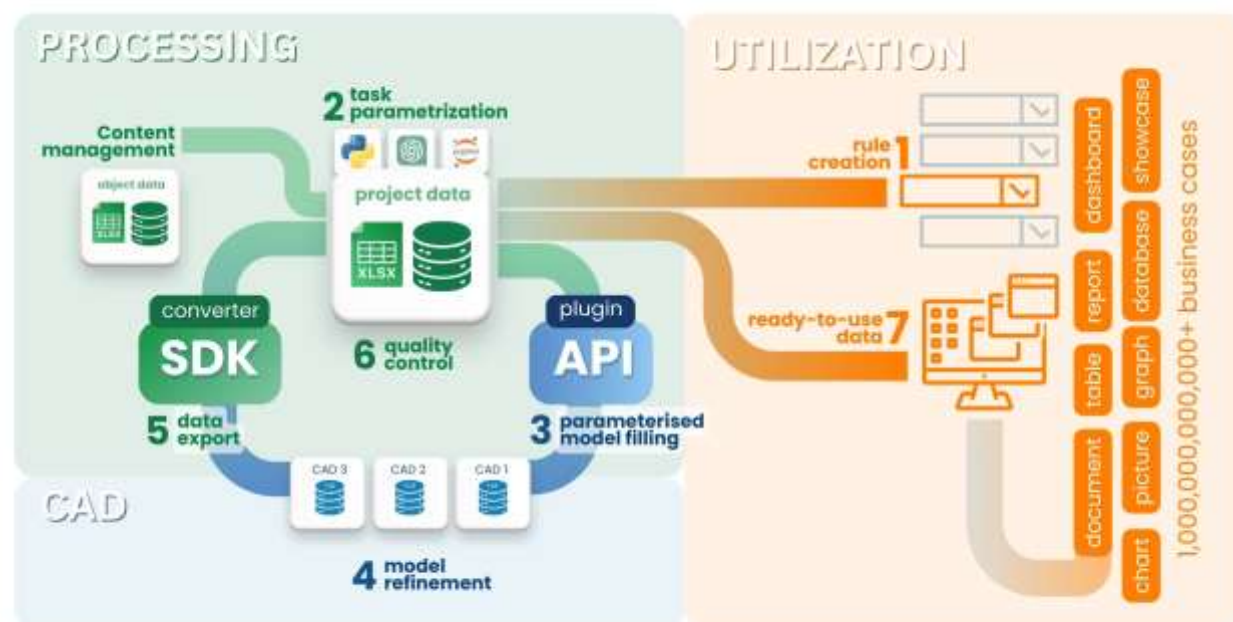
לבין מערכות ניהול נתוני בנייה אחרות הוא הצורך (BIM) CAD ההבדל המרכזי בין פלטפורמות (כדי לגשת למידע) מקור האמת היחיד. (בסיסי נתונים אלו אינם API-בשימוש בכלים מיוחדים וב אוניברסליים במובן המסורתי: במקום מבנה פתוח ואינטגרציה גמישה, הם מציעים סביבה סגורה, מחוברת באופן נוקשה לפלטפורמה ולפורמט ספציפיים.

עולה שאלה חשובה יותר, החורגת מהיישום הטכני: מה בעצם הם בסיסי CAD, למרות כל הקושי בעבודה עם נתוני כדי לענות על שאלה זו, יש לצאת מעבר לאקראי ולמונחים המוכרים, המוכתבים על ידי מפתחי (BIM) CAD הנתונים התוכנה. במקום זאת, יש להתמקד במהות העבודה עם מידע פרויקטואלי: נתונים ותהליכים של עיבודם.

אלא מהגדרת הדרישות לפרויקט, BIM או CAD תהליך העסקים בבנייה מתחיל לא מעבודה בכלים ומידול הנתונים. תחילה נקבעים פרמטרי המשימה: רשימת ישויות, מאפיינים התחלתיים וערכים גבוליים שיש לקחת בחשבון בעת פתרון המשימה הספציפית. רק לאחר מכן, על בסיס הפרמטרים שנקבעו, נוצרות (BIM) CAD המודלים והאלמנטים במערכות.

משחזר לחלוטין את תהליך מידול הנתונים, שנדון (BIM) CAD התהליך שמקדים את יצירת המידע בבסיסי הנתונים - "בפירוט בחלק הרביעי של הספר ובפרק "מידול נתונים: מודל קונספטואלי, לוגי ופיזי

בדיוק כמו בתהליך מידול הנתונים, אנו מייצרים דרישות לנתונים, שאותן נרצה לעבד בבסיס הנתונים, עבור בסיסי מנהלי הפרויקטים מייצרים דרישות לתכנון בצורה של מספר עמודות בטבלה או רשימות של זוגות CAD הנתונים נוצרות) או ליתר דיוק, מדויקות (API מפתח-ערך. "ורק על בסיס הפרמטרים הראשוניים הללו, באמצעות" לאחר מכן הם נבדקים שוב על מנת לוודא שהם עומדים בדרישות, (BIM) CAD האובייקטים בבסיסי הנתונים המקוריות. תהליך זה - הגדרה → יצירה → בדיקה → תיקון - חוזר על עצמו באיטרציות עד שהאיכות של הנתונים, בדיוק כמו במידול הנתונים, מגיעה לרמה הנדרשת עבור המערכת המיועדת - מסמכים, טבלאות או לוחות - מחוונים.

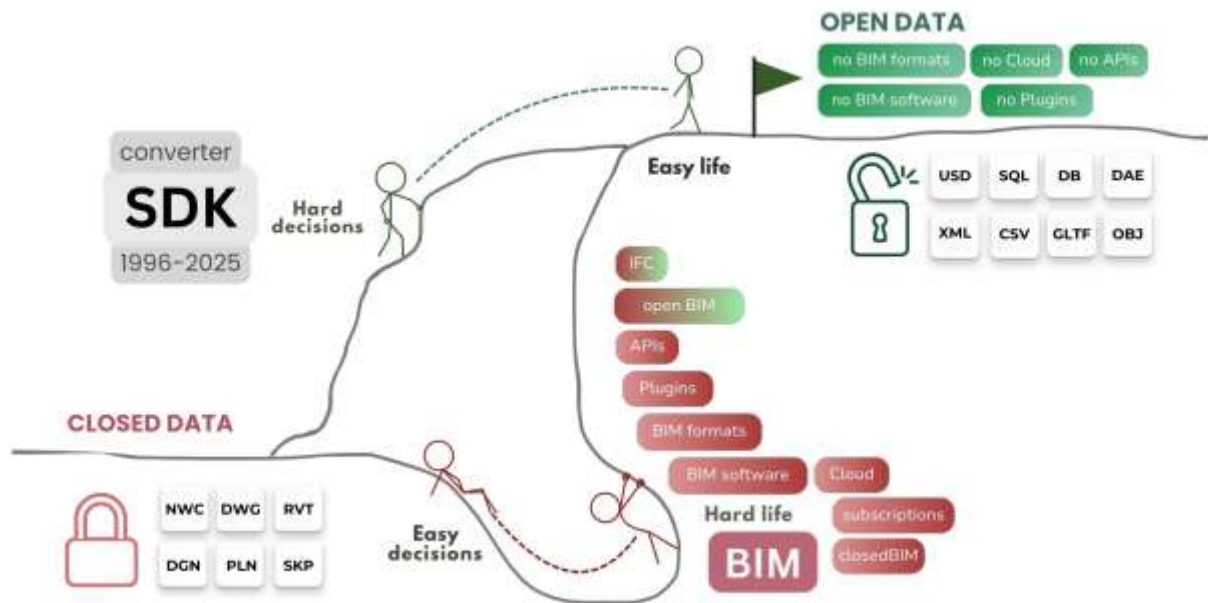


- ריש. 6.41. מחזור רוויה מידע בבסיסי נתונים עבור תהליכי עסקים בהוצאה לפועל של פרויקטים בבנייה

כמנגנון להעברת פרמטרים בצורה של קבוצת זוגות "מפתח-ערך", שנוצרו על בסיס (BIM) CAD אם נבחן את הדרישות שנקבעו מחוץ לסביבת התכנון, אז המוקד של הדיון יוסט מהפתרונות התוכניים הספציפיים ומגבלותיהם להיבטים יסודיים יותר – מבנה הנתונים, מודלים של נתונים ודרישותיהם. למעשה, מדובר ברוויה של בסיס הנתונים והמאפיינים CAD בפרמטרים ובתהליך הקלאסי של מידול נתונים. ההבדל הוא שבגלל הסגירות של בסיסי הנתונים BIM, עולה השאלה: מהי הייחודיות של BIM של הפורמטים בשימוש, תהליך זה מלווה בשימוש בכלים מיוחדים של? – אם אין גישות דומות בתעשיות אחרות

CAD-BIM נתפס כמשהו יותר מאשר פשוט מקור נתונים יחיד. בשיווק, הקישור BIM בעשרים השנים האחרונות נמכר לעיתים קרובות ככלי פרמטרי עם בסיס נתונים משולב מראש, המסוגל לאוטומט את תהליכי התכנון, המידול הפך יותר לכלי לשימור משתמשים על הפלטפורמה של BIM וניהול מחזור חיי האובייקטים בבנייה. עם זאת, במציאות הספקים, ולא לשיטה נוחה לניהול נתונים ותהליכים

API מבודדים בתוך הפלטפורמות שלהם, מסתירים מידע על הפרויקט מאחורי (BIM) CAD בסופו של דבר, נתוני קנייניים וליבות גיאומטריות. זה שלל מהמשתמשים את האפשרות לגשת באופן עצמאי לבסיסי הנתונים ולחלץ, לנתח לאוטומט ולהעביר נתונים למערכות אחרות תוך עקיפת האקוסיסטמות של הספקים



שמתעכנים מדי שנה ורישיונות מיוחדים API, ריש 6.42. בבנייה, הפורמטים המודרניים דורשים ליבות גיאומטריות מורכבות CAD (BIM). לתוכניות

צריכות להשתמש באותו גישה לעבודה עם נתונים, אשר בפועל כל ספקי CAD חברות העובדות עם כלים מודרניים של CAD-להנדסה הפוכה, נגד הפצתם ספקי ה SDK ללא יוצא מן הכלל מיישמים: המרת נתונים באמצעות כלים של CAD-ה ושימוש בכלים להנדסה הפוכה, אנו יכולים לקבל סט CAD נאבקים מאז 1995. עם גישה מלאה למסד הנתונים של שטוח של ישויות עם תכונות ולייצא אותן לכל פורמט פתוח נוח, הכולל גם גיאומטריה וגם פרמטרים של רכיבי התכנון גישה זו משנה באופן עקרוני את הפרדיגמה של עבודה עם מידע - מארכיטקטורה ממוקדת קבצים לארכיטקטורה ממוקדת נתונים.

■ RVT, IFC, PLN, DB1, CP2, CPIXML, USD, SQLite, XLSX, PARQUET פורמטי נתונים כמו ואחרים מכילים מידע זהה על רכיבים של אותו פרויקט. זה אומר שידע על פורמט ספציפי ועל הסכימה שלו לא צריך להיות מכשול לעבודה עם הנתונים עצמם

■ ותכונות MESH נתונים מכל פורמט ניתן לשלב במבנה פתוח, מסודר ומפורט, המכיל גיאומטריה משולשת של כל ישויות האובייקטים, ללא מגבלות של ליבות גיאומטריות

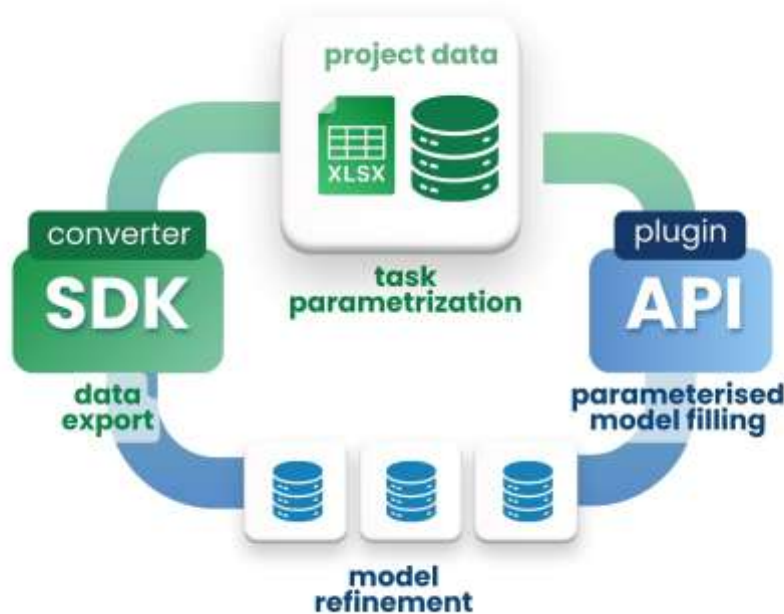
■ ניתוח נתונים שואף לאוניברסליות: באמצעות נתונים פתוחים, ניתן לעבוד עם נתוני פרויקטים ללא תלות בפורמט שבו נעשה שימוש

■ ובתוספים של ספקים: עבודה עם נתונים כבר לא תלויה בכישורים של שימוש ב API-צמצום, כמו גם תלות ב API.

מומרות לפורמטים נוחים לניתוח של הצגה מסודרת - המפתחים מפסיקים להיות CAD-כאשר הדרישות ונתוני ה תלויים בסכימות נתונים ספציפיות ובמערכות אקולוגיות סגורות

CAD ו-BIM תכנון באמצעות פרמטרים: עתיד

הם עוברים, CAD-לפני ששרטוט או מודל מקבלים צורה ב. CAD אף פרויקט בנייה בעולם לא התחיל אי פעם בתוכנית שלב של קונספטואליזציה, שבו תשומת הלב מתמקדת בפרמטרים המגדירים את הרעיון הבסיסי והלוגיקה של האובייקט העתידי. שלב זה תואם לרמה הקונספטואלית במידול נתונים. פרמטרים יכולים להתקיים אך ורק במחשבת המהנדס, אך במצב אידיאלי הם מעוצבים כרשימות מסודרות, טבלאות או נשמרים במסדי נתונים, מה שמאפשר שקיפות, שחזור --- ואוטומציה נוספת של תהליך התכנון



- במידע חיצוני באמצעות דרישות בשרשרת יצירת הערך CAD תהליך התכנון הוא תהליך איטרטיבי של מילוי מסד הנתונים של

עצמו) שלב הלוגי והפיזי של מידול נתונים, (חשוב לקבוע את הפרמטרים המגבילים CAD לפני שמתחילים במידול אשר משמשים כבסיס לפרויקט. תכונות אלו, כמו במקרה של דרישות אחרות, נאספות מהקצה של שרשרת השימוש בנתונים) למשל, מערכות (ודרכן כבר מוגדרות המגבלות, המטרות והמאפיינים המרכזיים של האובייקטים העתידיים בפרויקט.

המידול עצמו, כאשר ישנם דרישות מנוסחות היטב, יכול להיות אוטומטי לחלוטין ב-60%–100% בעזרת כלים פרמטריים למידול. ברגע שהפרויקט מתואר בצורה של פרמטרים, יצירתו הופכת לאפשרית טכנית, למשל בעזרת שפות תכנות מודרניות או פתרונות חינוניים במוצרים כמו CAD המובנות בסביבות Grasshopper ו-Dynamo ויזואליות, כמו Blender, UE, Omniverse.



רוב הפרויקטים המיועדים כבר היום נבנים באופן אוטומטי לחלוטין הודות לכלים של תכנות פרמטרי.

כבר היום פרויקטים תעשייתיים וסטנדרטיים גדולים נבנים לא על ידי ידיים של מחלקת המהנדסה, אלא דרך כלים פרמטריים ותכנות ויזואלי. זה מאפשר לבנות מודל, בהתבסס על נתונים, ולא על החלטות סובייקטיביות של מהנדס או מנהל ספציפי.

התוכן קודם לעיצוב. עיצוב ללא תוכן הוא לא עיצוב, אלא קישוט.

ג'פרי זלמן, מעצב אתרים ויום –

התהליך מתחיל לא מציור או מודל תלת-ממדי, אלא מהגדרת דרישות. הדרישות קובעות אילו רכיבים יישמשו בפרויקט אילו נתונים יש להעביר למחלקות ומערכות אחרות. רק קיום דרישות מובנות מאפשר לבדוק את המודלים באופן אוטומטי על בסיס קבוע (למשל, כל 10 דקות, מבלי להפריע לעבודה של המהנדס).

התהליך פשוט לממשק למילוי מסד נתונים, ולא משנה באיזה כלי (BIM) CAD ייתכן שבעתיד מערכת מתבצע המידול (ברמה הפיזית) – זה כבר לא יהיה חשוב CAD.

בדיוק כך, גם בהנדסת מכונות לעיתים קרובות נעשה שימוש במידול תלת-ממדי, אך זה לא תמיד אלמנט הכרחי או מחייב קלאסיים מספיקים – על בסיסם נבנית כל המודל המידע הנדרש. מודל זה מורכב בפרויקט. ברוב המקרים, מסמכי 2 מרכיבים, המובנים לפי תקנים תעשייתיים, ומכיל את כל המידע הנדרש להבנת המבנה וארגון הייצור. לאחר מכן, על בסיסו נבנית מודל המידע המפעלית, אליה מתווספים מוצרים ספציפיים וכרטיסי טכנולוגיה, המכוונים לצרכים של הטכנולוגים. כל התהליך יכול להתנהל ללא סיבוכים מיותרים, מבלי להעמיס על המערכת בגרפיקה תלת-ממדית היכן שזה לא נותן יתרונות ממשיים.

לא צריכים לשחק תפקיד מרכזי – הם רק כלי CAD חשוב להבין, שהמודל התלת-ממדי עצמו ומערכת לניתוח כמותי וגיאומטרי. כל שאר הפרמטרים, מלבד הגיאומטריה, שמתארים את המהות, צריכים ככל האפשר (BIM) CAD להישמר ולעבור עיבוד מחוץ לסביבת

תכנון באמצעות פרמטרים הוא לא רק טרנד, אלא עתיד בלתי נמנע של תעשיית הבנייה. במקום ליצור מודלים תלת-ממדיים מורכבים ידנית, המהנדסים יעבדו עם נתונים, יבדקו אותם ויאוטומטו תהליכים, מקרבים את הבנייה לעולם התכנות. עם הזמן, תהליכי התכנון יתבססו על עקרונות פיתוח תוכנה

■ יצירת דרישות → יצירת מודל → טעינה לשרת → בדיקת שינויים → בקשת משיכה

■ מתבצעות אוטומטית בדיקות מודל לפי הדרישות שנוצרו לפני (Pull request) במסגרת בקשת המשיכה תחילת או במהלך התכנון

■ לאחר בדיקת איכות הנתונים ואישור השינויים, הם מוטמעים בפרויקט, במסד הנתונים הכללי או מועברים אוטומטית למערכות אחרות

כבר עכשיו, בתחום ההנדסה המכאנית, שינויים בפרויקטים מתחילים עם יצירת הודעה על שינוי. תהליך דומה צפוי גם לתחום הבנייה: התכנון יהפוך לתהליך איטרטיבי, שבו כל שלב מגובה בדרישות פרמטריות. מערכת כזו תאפשר למעצבים ליצור בדיקות אוטומטיות ובקשות אוטומטיות להוספה-מיזוג בהתאם לדרישות ספציפיות

המהנדס של העתיד הוא, בראש ובראשונה, מפעיל נתונים, ולא מודליסט ידני. משימתו היא למלא את הפרויקט ישויות פרמטריות, כאשר הגיאומטריה היא רק אחד מהמאפיינים

תפקיד חשוב בשינוי יישא ההבנה של חשיבות המידול של נתונים, סיווג וסטנדרטיזציה, אשר נדונו בפירוט בפרקים XLSX הקודמים של הספר. התקנות המסדירות את התכנון בעתיד יוגדרו בצורה של זוגות פרמטרים מפתח-ערך בפורמט XML או

העתיד של תחום הבנייה הוא איסוף נתונים, ניתוחם, בדיקתם ואוטומציה של תהליכים באמצעות כלים אינה המטרה הסופית, אלא שלב באבולוציה. כאשר המומחים יבינו שהם (CAD או BIM). אנליטיים יפנה בהדרגה "BIM" המונח, CAD יכולים לעבוד ישירות עם הנתונים, תוך דילוג על כלים מסורתיים של את מקומו לקונספטים של שימוש בנתונים מובנים ומפורטים של פרויקט הבנייה

וכלים המבוססים עליהם (LLM) אחד הגורמים המרכזיים המזרזים את השינוי הוא הופעת מודלים לשוניים גדולים API טכנולוגיות אלו משנות את הגישה לעבודה עם נתוני פרויקטים, ומאפשרות גישה למידע מבלי צורך בידע מעמיק על הופך לאינטואיטיבי ונגיש CAD תהליך יצירת הדרישות והאינטראקציה עם נתוני LLM, או פתרונות של ספקים. בזכות

פרויקטים CAD בתהליכי עיבוד נתוני LLM הופעת

שינויים מהפכניים, CAD ולפורמטים פתוחים ומפשטים של CAD בנוסף לפיתוח כלים לגישה לבסיסי נתונים של (מודלים לשוניים גדולים). (אם בעבר הגישה למידע LLM בעיבוד נתוני פרויקטים מביאים לידי ביטוי הופעת כלים של

כיום ניתן לקיים אינטראקציה עם הנתונים, API, התבצעה בעיקר דרך ממשקים מורכבים ודורשה ידע בתכנות ובבאמצעות שפה טבעית.

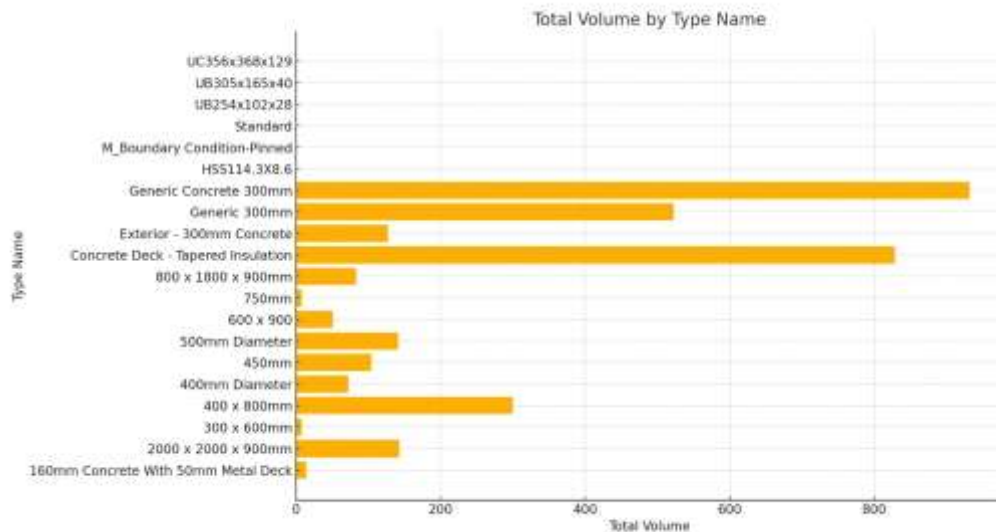
מהנדסים, מנהלים ומעצבים ללא רקע טכני יכולים לקבל את המידע הנדרש מנתוני הפרויקט על ידי ניסוח בקשות שאלה כמו: "הראה בטבלה עם LLM בשפה רגילה. בתנאי שהנתונים מובנים ונגישים, מספיק לשאול בצ'אט של קבוצות לפי סוג את כל הקירות עם נפח של יותר מ-10 מטרים מעוקבים" - והמודל יהפוך אוטומטית את הבקשה הזו לויצור את הטבלה הסופית, גרף או אפילו מסמך מוכן, Pandas או לקוד ב SQL.

מתקשרים עם נתוני פרויקטים המוצגים בפורמטים שונים של LLM להלן מספר דוגמאות מעשיות כיצד מודלים של CAD (BIM).

לאחר המרה לפורמט טבלאי (ראה CAD RVT לפרויקט בפורמט LLM דוגמת בקשה בצ'אט של 4.113):-

"לפי שם סוג תוך סיכום הפרמטר נפח RVT שהתקבל מקובץ Dataframe-סכם את הנתונים ב והראה את מספר הפריטים בקבוצה. והראה בבקשה את כל זה בצורה של היסטוגרמה אופקית ללא ערכים אפסיים.

(PNG פורמט) בצורה של היסטוגרמה אופקית LLM תשובת:



דרך בקשה QTO אנו מקבלים מיד טבלת LLM-במקום 17 לחיצות עכבר או 40 שורות קוד באמצעות שימוש בתוספים, ב. 6.45. רישקסטואלית.

Formulate של סוגי קירות עם שטח כולל ומספר מתוך הקטגוריה "קירות", QTO כדי ליצור טבלת LLM: בקשה טקסטואלית עבור צ'אט

קח רק את האלמנטים מתוך מסגרת הנתונים של הפרויקט, שיש להם בפרמטר "קטגוריה" את קבץ אותם לפי "שם סוג", "סכום את ערך העמודה" שטח, "הוסף את הכמות והצג", "OST_Walls" אותם בטבלה, מחק ערכים אפסיים

מוכנה QTO בצורה של טבלת LLM תשובת:

Type Name	Total Area	Count
CL_W1	393.12 sq m	10
Cavity wall_sliders	9.37 sq m	1
Foundation - 300mm Concrete	30.90 sq m	1
Interior - 165 Partition (1-hr)	17.25 sq m	3
Interior - Partition	186.54 sq m	14
Retaining - 300mm Concrete	195.79 sq m	10
SH_Curtain wall	159.42 sq m	9
SIP 202mm Wall - conc clad	114.76 sq m	4
Wall - Timber Clad	162.91 sq m	8

CAD- (BIM)- בשפה טבעית מספקת תוצאה באיכות דומה כמו בשימוש בכלים QTO יצירת טבלת 6.46. ריש

לאחר המרה למסגרת נתונים טבלאית ונכניס בקשה טקסטואלית דומה IFC נעשה בקשה לפרויקט בפורמט LLM: לכל צ'אט

וקח, "Parent" בפרמטר Level 1 ו-Level 2 קח רק את האלמנטים מהפרויקט, שיש להם ערכים בפרמטר "קטגוריה", ואז קבץ את הפריטים הללו לפי IfcSlab את האלמנטים שיש להם ערכים "PSet_RVT_Dimensions Area" סכום את הערכים בפרמטר, "ObjectType" פרמטר והצג אותם בצורה של תרשים עוגה

IFC: בצורה של תרשים עוגה מוכנה של קבוצת אלמנטים מנתוני LLM תשובת

For Floor:127mm Slab on Grade, Floor:150mm Exterior Slab on Grade,
Floor:Finish Floor - Ceramic Tile, Floor:Finish Floor - Wood, and
Floor:Residential - Wood Joist with Subflooring



בפורמט מובנה יכולה להיות כל סוג של גרף שנוח להבנת הנתונים IFC תוצאת הבקשה לנתוני 6.47. ריס

באמצעות Python-מסתתרות עשרות קוד ב (ריס - 6.45 - 6.47). מאחורי כל אחת מהתשובות המוכנות מקומית או מקוונת כדי להשיג תוצאות IDE ולשימוש בכל LLM הקוד המתקבל ניתן להעתיקה מצ'אט Pandas ספריית LLM.- זהות מחוץ לצ'אט

אלא גם עם, D CAD (BIM)אנו יכולים לעבוד לא רק עם פרויקטים שהתקבלו בפורמטים 3 LLM באותו צ'אט כדי LLM אליהם לאחר המרה לצורת נתונים מובנית נוכל להניח בקשה בצ'אט, DWG שרטוטים שטוחים בפורמט להציג, למשל, נתונים על קבוצות אלמנטים בצורה של קווים או גיאומטריות תלת-ממדיות

Pandas ו-LLM עם DWG ניתוח אוטומטי של קבצי

בשל חוסר המובנות של המידע - תמיד היה משימה מורכבת, שדרשה תוכנה DWG תהליך עיבוד הנתונים מקבצי הפך להיות אפשרי, LLM מיוחדת ולעיתים ניתוח ידני. עם זאת, עם התפתחות האינטליגנציה המלאכותית וכלי-אמיתי של בקשות ל Pipeline לאוטומט את רוב השלבים בתהליך זה, אשר עד היום היה ברובו תהליך ידני. נבחן: המאפשרים בעבודה עם הפרויקט, DWG לעבודה עם שרטוטי (ChatGPT בדוגמה זו) LLM

■ וקואורדינטות ID, לפי שכבות DWG לסנן נתוני

■ להמחיש את הגיאומטריה של האלמנטים

■ לאנוט את השרטוטים אוטומטית על בסיס פרמטרים

■ לפרוס פוליילינים של קירות במישור אופקי

■ ליצור הדמיות תלת-ממדיות אינטראקטיביות של נתונים שטוחים

■ מורכבים CAD למבנה ולנתח נתוני בנייה ללא כלים

תחילה נוצר בקשה, המתארת את LLM מתחיל מהפקת קוד רציפה דרך Pipeline-במקרה שלנו, תהליך בניית המבצע ומנותח, ומציג את התוצאה בתוך הצ'אט. אם התוצאה אינה תואמת, Python מייצר קוד ChatGPT. המשימה את הציפיות, הבקשה מתוקנת, והתהליך חוזר על עצמו.

צינור – זו סדרת צעדים אוטומטיים המבוצעים לעיבוד ולניתוח נתונים. בתהליך זה, כל שלב מקבל נתונים כקלט, מבצע המרות ומעביר את התוצאה לשלב הבא.

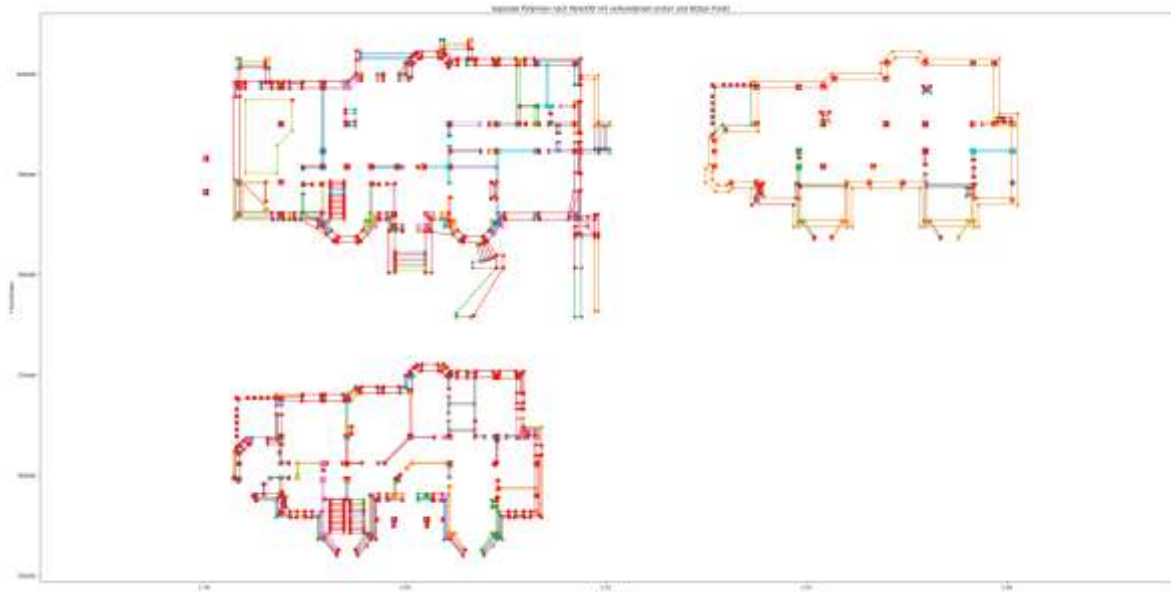
נוח, במקרה שלנו על IDE ומודבק לקוד בצורה של בלוקים בכל LLM-לאחר קבלת התוצאה הרצויה, הקוד מועתק מהקטעים שהתקבלו מאוחדים לצינור אחד, הממכן את כל התהליך – מהעלאת הנתונים. Kaggle.com הפלטפורמה ועד לניתוח הסופי שלהם. גישה כזו מאפשרת לפתח ולהרחיב תהליכי ניתוח במהירות, מבלי צורך במומחיות מעמיקה בתכנות. הקוד המלא של כל הקטעים המוזכרים למטה, יחד עם דוגמאות לשאלות, ניתן למצוא על הפלטפורמה "DWG Analyse with ChatGPT | DataDrivenConstruction" בחיפוש Kaggle.com.

– מהשלב הקלאסי (4.113 איור) לאחר המרה לצורת נתונים מובנית, DWG נתחיל את תהליך העבודה עם נתוני קיבוץ וסינון מתוך כל נתוני השרטוט, של האלמנטים הנדרשים עבור המשימה שלנו, במיוחד פוליגונים (הפרמטר מאפשר לקבץ קווים לקבוצות, (שיש להם בערך) בעמודת דאטה פריים" (שכבה "ערך מחרוזת, המכיל את 'ParentID' "wall" – (RegEx) השילוב הבא של אותיות

🔗 LLM-כדי לקבל קוד עבור משימה כזו ותוצאה בצורת תמונה, יש לכתוב את השאלתה הבאה ב

'Layer', 'ID', מכיל עמודות מסוימות DWG-קודם כל, בדוק אם דאטה פריים שהתקבל מ המכילים את המחרוזת, 'Layer' לאחר מכן, סנן את המזהים מעמודת 'Point' ו-'ParentID' את האלמנטים התואמים למזהים הללו. הגדר פונקציה לניקוי 'ParentID' מצא בעמודת 'wall' ו-'y', 'x' זה כולל הסרת סוגריים ופיצול הערכים לקואורדינטות. 'Point' ופיצול הנתונים בעמודת ייחודי, צייר פוליגון נפרד "ParentID" עבור כל matplotlib בנה גרף נתונים באמצעות 'z'. ודא שהנקודות הראשונות והאחרונות מחוברות, אם זה אפשרי. "Point" המחבר את קואורדינטות x-y קבע תוויות וכותרות מתאימות, תוך שמירה על קנה מידה אחיד בצירים.

🔗 Python: תספק תמונה מוכנה שמאחוריה מסתתר הקוד שנוצר בשפת LLM תשובת



ניקוי את הקואורדינטות שלהם ובנה פוליליינים, "wall" את כל הקווים של שכבת DWG-חילץ מקובץ ה LLM 6.48 איור
Python. בעזרת אחת מספריות

עכשיו נוסף לקווים את פרמטר השטח, שלכל פוליליין יש בתכונותיו (באחת מעמודות הדאטה פריים)

קח את ערך, "ID" אחד - מצא את המזהה הזה בעמודת "ParentID" עכשיו קבל מכל פוליליין רק
חלק ב-1,000,000 והוסף את הערך הזה לגרף, "Area".

תצוג גרף חדש, שבו לכל פוליליין תהיה תווית עם שטחו LLM תשובת

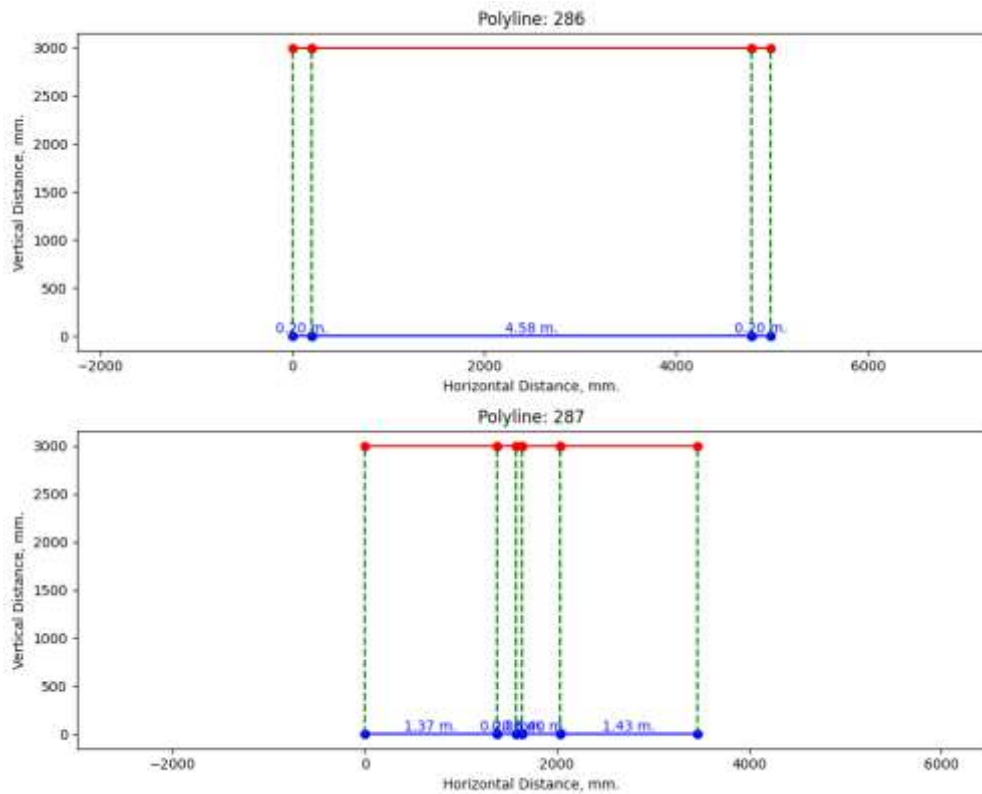


השלים את הקוד, שלוקח את ערכי השטח עבור כל פוליליין ומוסיף אותו לתמונה עם הוויזואליזציה של הקווים LLM 6.49 איור.

לאחר מכן, נמיר כל פוליליין לקו אופקי, נוסף קו מקביל בגובה 3000 מ"מ ונחבר אותם למישור אחד, כדי להציג כך את פריסת פני השטח של אלמנטים קירות.

ים הללו כרשימה ID-קחו את ה. "wall" עם הערך "Layer" יש לקחת את כל האלמנטים מעמודת כל האלמנטים. "ParentID" ים הללו בכל מסגרת הנתונים בעמודת ID-ומצאו את ה "ID" מעמודת של הנקודה הראשונה x, y הם קווים שמתאגדים לפוליליניה אחת. כל קו יש לו את הגיאומטריה יש לקחת כל פוליליניה בתור תור ולהתחיל מנקודה 0,0 אופקית לבנות את. "Point" בעמודת אורך כל קטע מהפוליליניה. אורך כל קטע פוליליניה בקו אחד. לאחר מכן, ציירו קווים זהים רק בגובה 3000 מ"מ, חברו את כל הנקודות למישור אחד.

תפיק קוד שמאפשר לבנות פריסות קירות במישור LLM תשובת

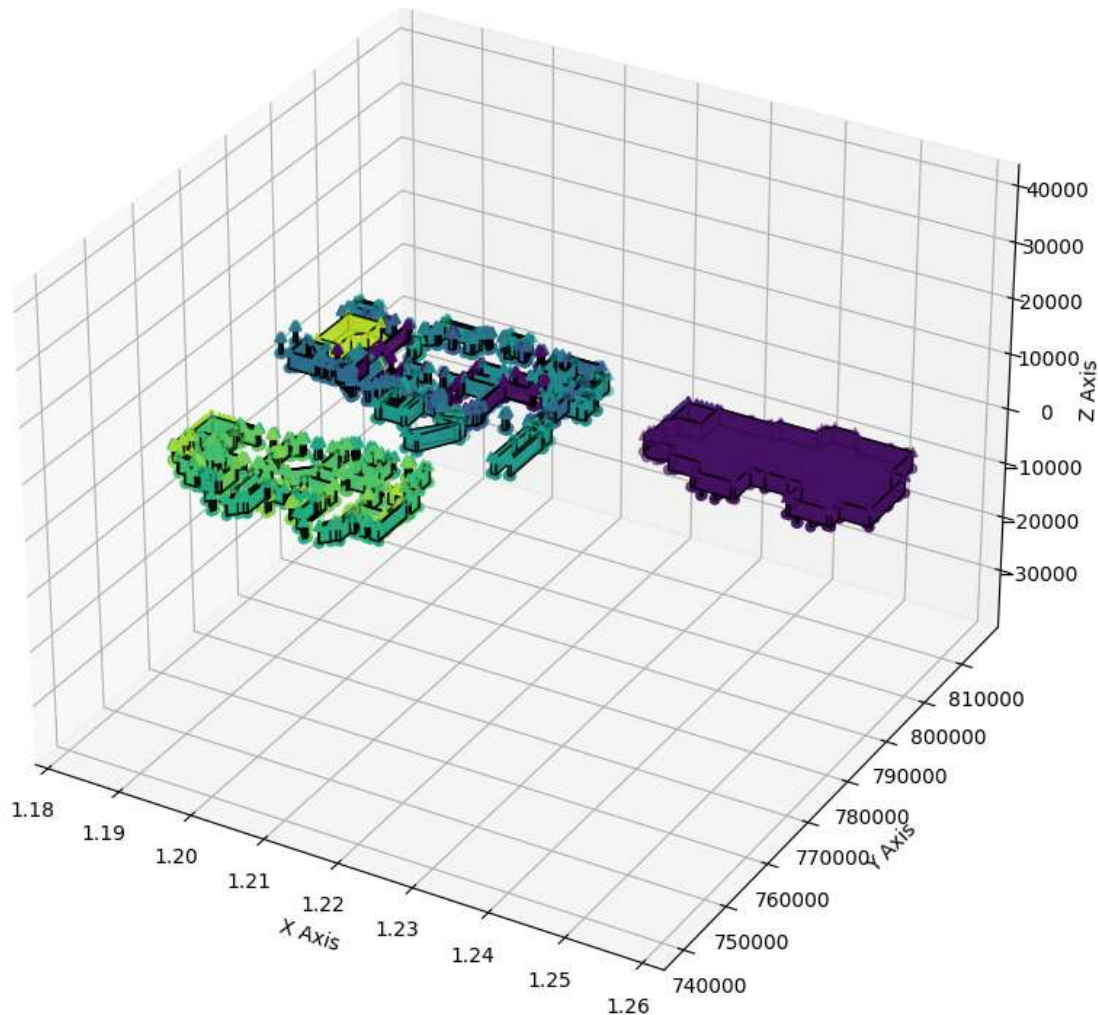


כל פוליליניה אנו הופכים בעזרת פקודות לפריסה, שמראה בצורה ברורה את המישורים של הקירות ישירות בצ'אט של 6.410 איור LLM.

עכשיו נעבור מהפרויקט הדו-ממדי למודל תלת-ממדי של קירות מקווים שטוחים על ידי חיבור השכבות העליונות והתחתונות של הפוליליניות:

ובגובה 3000 מ"מ $z = 0$ ויזואליזציה של אלמנטים הקיר בתלת-ממד, חיבור פוליליניות בגובה כדי ליצור גיאומטריה סגורה המייצגת את קירות הבניין. השתמש בגרף תלת-ממדי של Matplotlib.

ייצור גרף תלת-ממדי אינטראקטיבי, שבו כל פוליגוניה תוצג כסט של מישורים. המשתמש יוכל לנוע LLM בחופשיות בין האלמנטים בעזרת העכבר, לחקור את המודל במצב תלת-ממדי, על ידי העתקת הקוד מהצ'אט IDE-ל:



תלת-viewer-עזר לבנות קוד [129] עבור ויזואליזציה של קווים שטוחים של השרטוט בתצוגת תלת-ממד, שניתן לחקור ב LLM 6.411 איור IDE-ממדי בתוך ה

ועד קבלת התוצאה הסופית DWG הגיוני וחוזר על עצמו – מהמרה הראשונית והעלאת קובץ Pipeline כדי לבנות כך, לא רק שתבדקו את התוצאה בצ'אט IDE-ל LLM מומלץ לאחר כל שלב להעתיק את בלוק הקוד שנוצר על ידי – אלא גם תריצו אותה מיד בסביבת הפיתוח שלכם. זה מאפשר לבנות את התהליך בצורה רציפה, לתקן ולהתאים אותו לפי הצורך.

יחד עם דוגמאות לפקודות ניתן למצוא (איור - 6.411 איור - 6.48) לכל החלקים Pipeline-הקוד המלא של ה "DWG Analyse with ChatGPT | DataDrivenConstruction" בחיפוש Kaggle.com בפלטפורמת Pipeline-ניתן לא רק לצפות בקוד ובפקודות בשימוש, אלא גם להעתיק ולבדוק בחינם את כל ה Kaggle-ב [129]. --עצמו IDE-בסביבה עננית ללא צורך בהתקנת תוכנה נוספת או ה DWG עם מסגרות הנתונים המקוריות של

ה-DWG הגישה המוצגת בפרק זה מאפשרת אוטומציה מלאה של בדיקה, עיבוד והפקת מסמכים על בסיס פרויקט DWG שפותח מתאים לעיבוד שרטוטים בודדים כמו גם לעיבוד קבוצתי של עשרות, מאות ואלפי קבצי Pipeline הפקת דוחות וויזואליזציות נדרשות לכל פרויקט

לאחר, XLSX מומרות אוטומטית לפורמט CAD התהליך ניתן לבנייה בצורה רציפה ושקופה: תחילה הנתונים מקבצי מכן הם מועלים למסגרת נתונים, ולאחר מכן מתבצעת קיבוץ, בדיקה והפקת תוצאה - כל זה מתבצע באותו מחברת פופולרי. אם יש צורך, התהליך ניתן להרחבה בקלות על ידי אינטגרציה עם IDE בכל, Python או סקריפט Jupyter אוטומטית לפי קריטריונים שנקבעו, להחזיר את התוצאות חזרה CAD מערכות ניהול מסמכי פרויקטים: ניתן לחלץ קבצי למערכת האחסון ולהודיע למשתמשים על מוכנות התוצאות - באמצעות דוא"ל או דרך מסנג'רים

CAD לעבודה עם נתוני פרויקטים מפחית את התלות בתוכנות LLM השימוש בצ'אט ובסוכנים של מיוחדות ומאפשר לבצע ניתוח וויזואליזציה של פרויקטים ארכיטקטוניים מבלי צורך באינטראקציה ידנית עם הממשק - ללא לחיצות עכבר וזכירת ניווט מורכב בתפריטים

DataFrame, נתונים מובנים גרנולריים, LLM עם כל יום שעובר, יותר ויותר ניתן לשמוע בתעשיית הבנייה על הפכו, CAD דו-ממדיים מאוחדים, שנוצרו ממסדי נתונים שונים ומפורמטים DataFrame. ומסדי נתונים עמודיים לדלק אידיאלי עבור כלים אנליטיים מודרניים, אשר בהם עובדים באופן פעיל מומחים בתעשיות אחרות בכלכלה של מוצרים נישתיים סגורים ולכתוב סקריפטים API תהליך האוטומציה עצמו יתפשט באופן משמעותי - במקום ללמוד מורכבים לניתוח או טרנספורמציה של פרמטרים, עכשיו יהיה מספיק לנסח את המשימה בצורה של קבוצת פקודות הנדרש לשפת התכנות הנדרשת, אשר ניתן להריץ Workflow או תהליך Pipeline-טקסט נפרדות, אשר יתאגדו ל בחינם כמעט על כל מכשיר. לא יהיה צורך יותר לחכות להשקת מוצרים חדשים, פורמטים, תוספים או עדכונים מספקי מהנדסים ובונים יקבלו את האפשרות לעבוד באופן עצמאי עם נתונים, תוך שימוש בכלים פשוטים. CAD (BIM) כלי LLM חינוכיים וברורים, בעבודה עם עזרה של צ'אט וסוכנים של

צעדים נוספים: המעבר מפורמטים סגורים לנתונים פתוחים

בעבודה עם נתוני פרויקטים של העתיד, סביר להניח שאין צורך להעמיק בהבנה של ליבות גיאומטריות של כלים קנייניים או ללמוד מאות פורמטים לא תואמים, המכילים את אותה המידע. עם זאת, ללא הבנה מדוע המעבר לנתונים פתוחים ומובנים חשוב, קשה לטעון את הצורך בשימוש בכלים חדשים, חינוכיים ופתוחים ובגישות, אשר סביר להניח שלא יקודמו על ידי ספקי התוכנה

המגבלות והאפשרויות שלהם, ושמען, למרות ההבטחות, CAD (BIM) בפרק זה דנו בתכונות המרכזיות של נתוני השיעור של הספקים, מהנדסים ומעצבנים מתמודדים בכל יום עם קשיים בהפקה, העברה וניתוח של מידע פרויקט הבנת הארכיטקטורה של מערכות אלו והיכרות עם גישות חלופיות - המבוססות על פורמטים פתוחים ואוטומציה בעזרת יכולה לפשט באופן משמעותי את החיים אפילו למומחה אחד, שלא לדבר על חברות. לסיכום חלק זה, כדאי - LLM להדגיש את הצעדים המעשיים העיקריים שיכולים לעזור ליישם את הגישות שנדונו במשימות היומיומיות שלכם

■ הרחיבו את הכלים שלכם לעבודה עם נתוני פרויקטים

□ בהן אתם משתמשים CAD (BIM) חקרו את התוספים והכלים הזמינים להפקת נתונים ממערכות

☐ הזמינים, המאפשרים אוטומציה של הפקת נתונים מפורמטים סגורים מבלי API-וה SDK-חקרו את ה. הצורך לפתוח ידנית תוכנה מיוחדת

☐ (OBJ, glTF, USD, רכשו מיומנויות בסיסיות בעבודה עם פורמטים גיאומטריים פתוחים לא פרמטריים וספריות פתוחות מתאימות (DAE)

☐ CAD נסו לחשוב על מערכת לאחסון מטא-דאטה של הפרויקט בנפרד מהגיאומטריה מחוץ לפתרונות כדי לפשט את הניתוח והאינטגרציה עם מערכות אחרות (BIM)

☐ לאוטומציה של שאלות המרה בין פורמטים LLM-השתמשו ב

צרו תהליכים אישיים לעיבוד מידע פרויקטלי ■

☐ התחילו לתאר משימות ודרישות למידול דרך פרמטרים וערכיהם בפורמטים פשוטים ומסודרים

☐ צרו ספרייה אישית של סקריפטים או קטעי קוד לפעולות שבוצעות לעיתים קרובות

קידמו את השימוש בסטנדרטים פתוחים בעבודתכם ■

☐ הציעו לעמיתים ולשותפים להחליף נתונים בפורמטים פתוחים, שאינם מוגבלים לאקוסיסטמות של ספקי תוכנה

☐ הדגישו את היתרונות של שימוש בנתונים מסודרים בדוגמאות קונקרטיות

☐ יזמו דיונים על בעיות עם פורמטים סגורים ופתרונות אפשריים

הבנה אישית של עקרונות העבודה, (BIM-) CAD גם אם אינכם יכולים לשנות את מדיניות החברה בנוגע לפלטפורמות עם נתוני פרויקטים בפורמטים פתוחים תאפשר לכם לשפר משמעותית את היעילות בעבודתכם. על ידי יצירת כלים ושיטות אישיות להפקה והמרה של נתונים מפורמטים שונים, לא רק שתשפרו את תהליכי העבודה, אלא גם תקבלו גמישות שתאפשר לכם לעקוף את המגבלות של פתרונות תוכנה סטנדרטיים.



VII חלק

קבלת החלטות מבוססות נתונים, ניתוח, אוטומציה ולמידת מכונה

החלק השביעי מוקדש לניתוח נתונים ואוטומציה של תהליכים בתעשיית הבנייה. כאן נדונה כיצד נתונים הופכים לבסיס לקבלת החלטות ומוסברים העקרונות של (KPI) ויזואליזציה של מידע לניתוח יעיל. מתוארים בפירוט מדדי ביצוע מרכזיים ויצירת לוחות מחוונים למעקב אחר (ROI) שיטות להערכת החזר על השקעה ויצירת ETL (Extract, Transform, Load) פרויקטים. תשומת לב מיוחדת מוקדשת לתהליכי מה שמאפשר להפוך (Pipeline) ואוטומציה שלהם באמצעות צינורות נתונים מפורזים למידע מסודר לניתוח. נדונים כלים לאורקסטראציה של תהליכי המאפשרים לבנות Apache NiFi ו-Apache Airflow עבודה, כמו צינורות נתונים אוטומטיים ללא ידע מעמיק בתכנות. תפקיד משמעותי מוקצה וליישומם לפשט את ניתוח הנתונים ואוטומציה (LLM) לדגמים לשוניים גדולים של משימות שגרתיות.

פרק 7.1

ניתוח נתונים וקבלת החלטות מבוססות נתונים

לאחר שלבי האיסוף, הסידור, הניקוי והבדיקה של המידע, נוצר מערך נתונים שלם וראוי לניתוח. בחלקים הקודמים של ורשומות טקסט של פגישות PDF הספר נדונה הסיסטמטיזציה והסידור של מקורות מגוונים – מדוקומנטים בפורמט ונתונים גיאומטריים. מתואר בפירוט תהליך הבדיקה וההתאמה של המידע לדרישות של מערכות CAD ועד מודלים שונות ומסוגים, הסרת כפילויות וסתירות.

כל החישובים שבוצעו על בסיס נתונים אלה (החלק השלישי והרביעי של הספר) – החל מהמרות פשוטות ועד חישובי (החלק החמישי) – מהווים משימות מנותחות ברמה גבוהה. הם מהווים את הבסיס להבנת ESG מועדים, עלויות ומדדי המצב הנוכחי של הפרויקט, הערכת פרמטרים שלו וקבלת החלטות לאחר מכן. כתוצאה מכך, הנתונים, בעקבות החישובים, הופכים מקבוצת רשומות מנותקות למשאב מנוהל, המסוגל לענות על שאלות מפתח של העסק.

בפרקים הקודמים נדונו בפירוט תהליכי איסוף נתונים ובקרת איכותם לשימוש במקרי עסקים טיפוסיים ובתהליכים האופייניים לתעשיית הבנייה. האנליטיקה בהקשר זה דומה במידה רבה ליישומים בתעשיות אחרות, אך יש לה גם מספר מאפיינים ספציפיים.

בפרקים הבאים יידון תהליך ניתוח הנתונים ברמה גבוהה, כולל שלבי אוטומציה – מהשגת המידע הראשונית והמרתו ועד העברתו למערכות ולמסמכים היעודיים. תחילה תוצג החלק התיאורטי, המוקדש להיבטים שונים של קבלת החלטות על ETL-Pipeline בסיס נתונים. לאחר מכן, בפרקים הבאים, תתחיל החלק המעשי, הקשור לאוטומציה ולבניית

נתונים כמשאב בקבלת החלטות

תהליך קבלת ההחלטות על בסיס נתונים הוא לעיתים קרובות תהליך חזורי ומתחיל באיסוף שיטתי של מידע ממקורות – מידע מגוונים. בדומה למחזור הטבעי, אלמנטים בודדים של נתונים ומערכות מידע שלמות בהדרגה נופלים לאדמה מצטברים במחסני מידע של חברות. עם הזמן, נתונים אלה, בדומה לעלים וענפים שנפלו, מתמירים לחומר יקר ערך המצילום של מהנדסי נתונים ואנליטיקאים מארגן ומכין את המידע לשימוש עתידי והופך נתונים ומערכות שנפלו – לקומפוסט יקר ערך, לגידול נבטים חדשים ומערכות חדשות.

מגמות השימוש הרחב באנליטיקה בתעשיות שונות מסמנות את תחילתה של תקופה חדשה, שבה העבודה עם נתונים הופכת לבסיס לפעילות מקצועית. חשוב למומחים בתעשיית הבנייה להסתגל לשינויים אלה ולהיות מוכנים למעבר לעידן – חדש – עידן הנתונים והאנליטיקה.

העברת נתונים ידנית בין טבלאות וביצוע חישובים ידניים הולכים ונעלמים, מפנים מקום לאוטומציה לניתוח זרמי נתונים, לאנליטיקה וללמידת מכונה. כלים אלה הופכים לרכיבים מרכזיים במערכת המודרנית לתמיכה בקבלת החלטות.

בספר של מקינזי "רענון. מדריך מקינזי להתמודדות עם תחרות בעידן הדיגיטלי ובינה מלאכותית", מובאת מחקר שנערך בשנת 2022 עם השתתפות 1,330 מנהיגים בכירים מאזורי גיאוגרפיים, תעשיות וכיוונים פונקציונליים שונים. על פי תוצאותיו, 70% מהמובילים משתמשים באנליטיקה מתקדמת כדי לפתח רעיונות. משלהם, ו-50% מיישמים בינה מלאכותית לשיפור ואוטומציה של תהליכי קבלת החלטות.



ניתוח נתונים ואנליטיקה - כלי מרכזי להאצת קבלת החלטות בחברה.

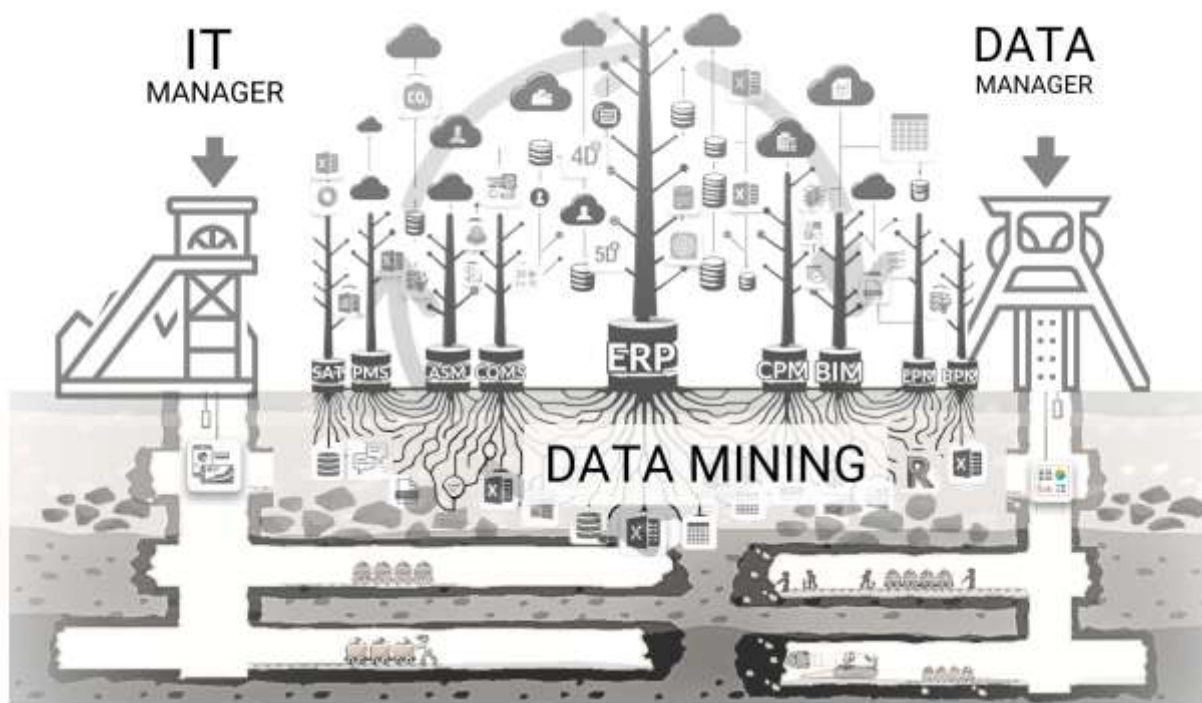
ניתוח נתונים, בדומה להתפשטות מיצילים, חודר דרך ההמוס של החלטות עבר, מסייע לחבר בין מערכות נפרדות ומנחה מנהלים לידע יקר ערך. ידע זה, בדומה לחומרים מזינים מעצים מתפוררים של מערכות נתונים, מזין החלטות חדשות בחברה, מה שמוביל לשינויים יעילים ולצמיחה איכותית של מידע, בדומה לירקות חדשים ולנבטים המופיעים מאדמה. -עשירה ובריא (איור 1.25)

מספרים יש היסטוריה חשובה שהם צריכים לספר. הם סומכים על כך שתיתן להם קול ברור ומשכנע.

סטיבן פייו, מומחה לויזואליזציה של נתונים -

בחברות בינוניות וקטנות, העבודה על חילוץ והכנה של מידע לניתוח נוסף מהווה כיום תהליך שדורש עבודה רבה (איור השווה לחציבת פחם במאה ה-18. עד לאחרונה, העבודה על חילוץ והכנה של נתונים הייתה מיועדת בעיקר, (7.12) להרפתקנים שעבדו בנישה צרה עם סט קטן ומוגבל של כלים לעבודה עם סוגים שונים של נתונים ממקורות לא מובנים חצי מובנים, מעורבים וסגורים.

מנהיגים ומנהלים המקבלים החלטות לעיתים קרובות אינם בעלי ניסיון מספיק בעבודה עם נתונים ומערכות מגוונות, אך הם נדרשים לקבל החלטות על בסיסם. כתוצאה מכך, קבלת החלטות מבוססת נתונים בתעשיית הבנייה המודרנית בעשורים האחרונים פחות דומה לתהליך אוטומטי ויותר לעבודה ידנית ממושכת של כורי פחם בשדות הפחם הראשונים.



איור 7.12 בתהליך חיפוש נתונים, המומחים עוברים דרך מורכבת של הכנת נתונים - מניקוי ועד מבנה לניתוח עתיד.

אף על פי ששיטות החילוץ המודרניות בתעשיית הבנייה הן ללא ספק מתקדמות יותר מאשר הטכניקות הפרימיטיביות של כורי פחם במאה ה-12, זה עדיין נשאר משימה מורכבת ומסוכנת מאוד, שדורשת משאבים משמעותיים ומומחיות שיכולה להרשות לעצמה רק חברות גדולות. תהליכי החילוץ והאנליזה של נתונים מהמורשת שנצברה מפרויקטים קודמים עד לאחרונה היו בעיקר בתחום החברות הגדולות והמתקדמות טכנולוגית, אשר אספו ושמרו נתונים במשך עשורים.

בעבר, חברות טכנולוגיות בוגרות שיחקו תפקיד מרכזי באנליזה, אוספות נתונים במשך עשורים. כיום המצב משתנה: הגישה לנתונים ולכלים לעיבודם הופכת לדמוקרטית - פתרונות שהיו מסובכים בעבר זמינים כעת לכולם וללא תשלום.

יישום אנליזה מאפשר לחברות לקבל החלטות מדויקות ומבוססות יותר בזמן אמת. להלן דוגמה מעשית המדגימה כיצד נתונים היסטוריים מסייעים בקבלת החלטות כלכליות מבוססות:

- 🔍 "מנהל פרויקט - "כעת המחיר הממוצע של בטון בעיר הוא 82 €/מ³, ובאומדן שלנו הוא 95 €/מ³
- 🔍 "שמאי - "בפרויקטים קודמים ההוצאות היו כ-15%, ולכן אני נזהר
- 🔍 "מנהל נתונים או מהנדס בקרה מצד המזמין - "בואו נבדוק את האנליזה של שלושת המכרזים האחרונים

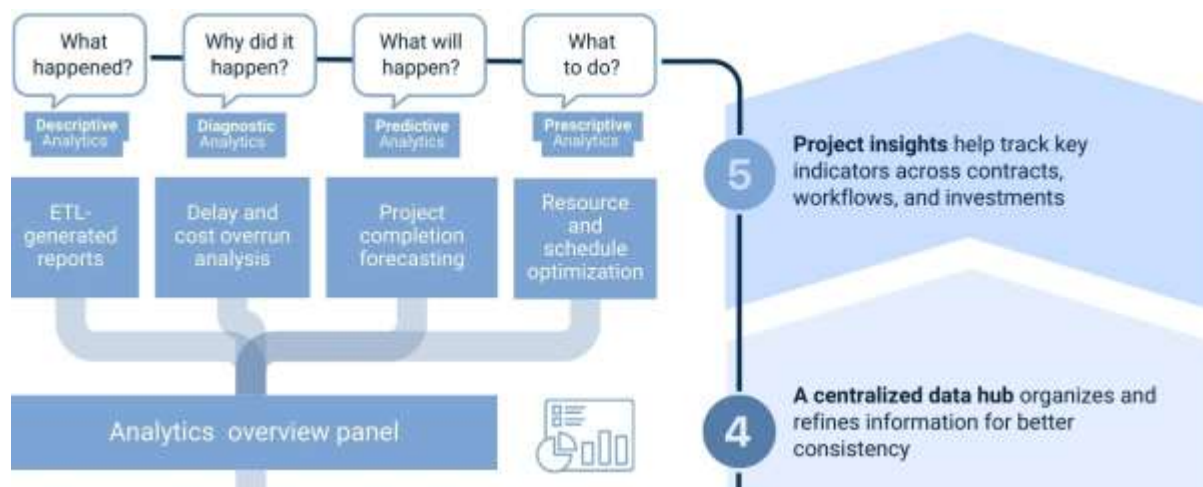
מפרויקטים קודמים אנו מקבלים DataFrame-לאחר ניתוח ה

■ מחיר הרכישה הממוצע בפועל: 84.80 €/מ³

■ שיעור ההוצאות הממוצע: +4.7%

שיעור מומלץ בהצעת המחיר: 85 מ€/מ³

פתרון כזה יתבסס לא על תחושות סובייקטיביות, אלא על סטטיסטיקה היסטורית קונקרטית, מה שמאפשר להפחית סיכונים ולהגביר את ההצדקה של הצעת המחיר. ניתוח נתונים מפרויקטים קודמים הופך להיות מעין "דשן אורגני", ממנו צומחים פתרונות חדשים ומדויקים יותר.



אנליטיקת נתונים עונה על שלושה שאלות מפתח: מה קרה, מדוע זה קרה ומה יש לעשות הלאה 7.13. ריש.

מנהיגים ומנהלי החלטות לעיתים קרובות מתמודדים עם הצורך לעבוד עם נתונים ומערכות מגוונות, מבלי שיש להם הכשרה טכנית מספקת. במצבים כאלה, אחד העוזרים המרכזיים בתהליך הבנת הנתונים הוא הוויזואליזציה - אחד השלבים הראשונים והחשובים בתהליך האנליטי. היא מאפשרת להציג מידע בצורה ברורה ומובנת.

וויזואליזציה של נתונים: מפתח להבנה ולקבלת החלטות

בתעשיית הבנייה המודרנית, שבה נתוני הפרויקטים מאופיינים במורכבות ובמבנה רב-שכבתי, הוויזואליזציה משחקת תפקיד מרכזי. וויזואליזציה של נתונים מאפשרת למנהלי פרויקטים ומהנדסים להמחיש דפוסים ומגמות מורכבות, החבויות בנפחים גדולים ומגוונים של נתונים.

וויזואליזציה של נתונים מקלה על הבנת מצב הפרויקט: הפצת משאבים, דינמיקת הוצאות או שימוש בחומרים. בעזרת גרפים ודיאגרמות, מידע מורכב ויבש הופך לנגיש ומובן, מה שמאפשר לזהות במהירות אזורים מרכזיים שדורשים תשומת לב ולגלות בעיות פוטנציאליות.

וויזואליזציה של נתונים לא רק מקלה על פרשנות המידע, אלא היא שלב קריטי בתהליך האנליטי ובקבלת החלטות. (2.25. ריש) "ניהוליות מבוססות, ועוזרת לענות על השאלות "מה קרה?" ו"איך זה קרה

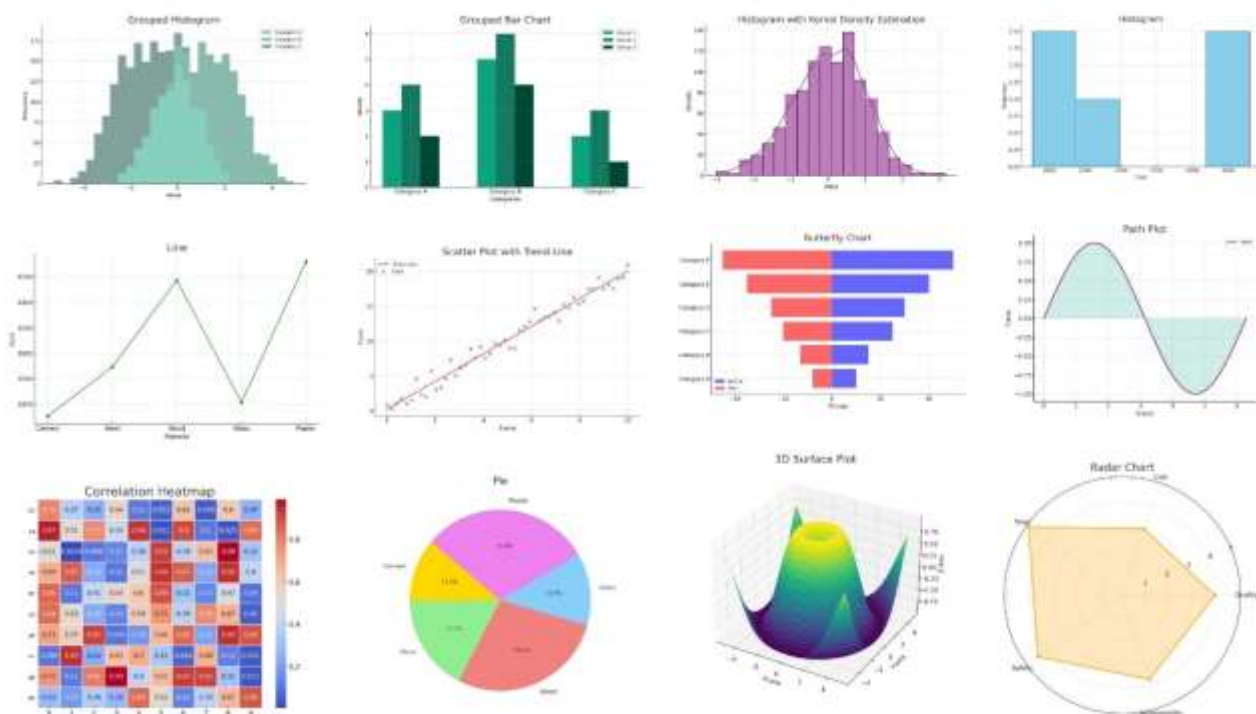
לפני קבלת החלטות מרכזיות, מנהלי פרויקטים יהיו סבירים יותר להשתמש בייצוגים חזותיים של נתונים, ולא במספרים יבשים וקשים לפרשנות מתוך גיליונות אלקטרוניים או הודעות טקסט.

נתונים שאינם מגובים בוויזואליזציה דומים לחומרי בניין מפוזרים באתר הבנייה: הפוטנציאל שלהם אינו ברור. רק כאשר נוצרת מהם ויזואליזציה ברורה, כמו מבית שנבנה מלבנים ובטון, מתברר איזו ערך הם מייצגים. עד שהבית לא נבנה, אי אפשר לדעת אם ערימת החומרים תהפוך לאוהל קטן, וילה מפוארת או גורד שחקים.

בעסקאות פיננסיות ובנתונים טקסטואליים, (2.110 איור - 1.24 איור) חברות מחזיקות בנתונים ממערכות שונות, רחבים. עם זאת, השימוש בנתונים אלה לטובת העסק מהווה לעיתים קרובות משימה מורכבת. במצבים כאלה, הוויזואליזציה הופכת לכלי חשוב להעברת משמעות הנתונים, המסייעת להציג את המידע בפורמטים המובנים לכל מומחה. לדוגמה, בצורה של לוחות מחוונים, גרפים ודיאגרמות.

מה צריך לסטודנטים כדי להצליח בעולם העסקים המשתנה במהירות) "2015 (מדגיש 9 [כי" PwC מחקר חברות מצליחות אינן מוגבלות לניתוח נתונים, אלא משתמשות באופן פעיל בכלים אינטראקטיביים של - ויזואליזציה, כמו גרפים, אינפוגרפיקה ולוחות מחוונים, כדי לתמוך בקבלת החלטות. על פי הדו"ח ויזואליזציה של נתונים מסייעת ללקוחות להבין את הסיפור שהנתונים מספרים, באמצעות גרפים, דיאגרמות. לוחות מחוונים ומודלים אינטראקטיביים של נתונים.

תהליך המרת מידע לצורות גרפיות ויזואליות, כמו דיאגרמות וגרפים, משפר את ההבנה והפרשנות של הנתונים) איור על ידי המוח האנושי. זה מאפשר למנהלי פרויקטים ואנליסטים להעריך במהירות תרחישים מורכבים ולקבל (7.14 החלטות מושכלות, בהתבסס לא על אינטואיציה, אלא על מגמות ודפוסים המוכרים ויזואלית.



סוגי ויזואליזציה שונים נוצרו כדי לסייע למוח האנושי להבין ולהשיג את המידע היבש של מספרים 7.14 איור.

נושאים נוספים הקשורים ליצירת ויזואליזציות מנתונים, ולשימוש בספריות ויזואליזציה חזותיות שונות, יידונו בפרק ETL, הבא, המוקדש לתהליכי

ויזואליזציה הופכת להיות מרכיב בלתי נפרד בעבודה עם נתונים בתעשיית הבנייה - היא מסייעת לא רק "לראות" את הנתונים, אלא גם להבין את משמעותם בהקשר של משימות ניהוליות. עם זאת, כדי שהויזואליזציה תהיה באמת מועילה יש לקבוע מראש מה יש לויזואליזציה ומה המדדים החשובים באמת להערכת היעילות של הפרויקט. כאן נכנסים לתמונה "בלעדיהם, אפילו הלוחות המחוונים היפים ביותר עלולים להישאר "רעש מידע. KPI ו-ROI מדדי ביצוע, כמו

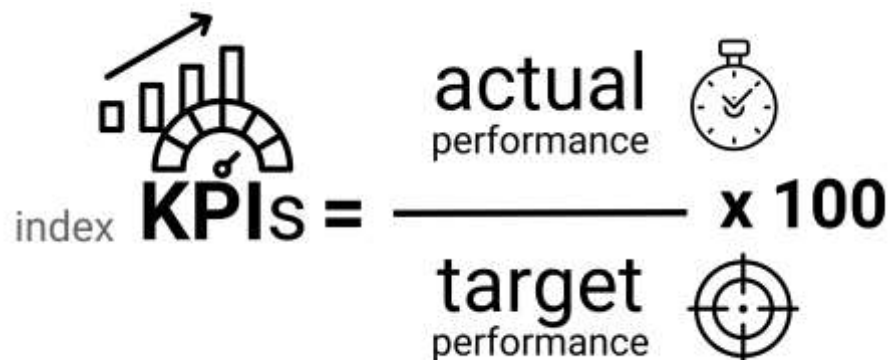
KPI ו-ROI מדדי ביצוע

כמו גם הויזואליזציה שלהם באמצעות דוחות ולוחות (KPI ו-ROI) בתעשיית הבנייה המודרנית, ניהול מדדי ביצוע משחקים תפקיד מרכזי בשיפור הפרודוקטיביות והיעילות של ניהול פרויקטים (dashboards), מחוונים

כמו בכל עסק, גם בבנייה יש לקבוע במדויק את המדדים שעל פיהם מעריכים הצלחה, החזר השקעה ופרודוקטיביות. כאשר מקבלים נתונים על תהליכים שונים, ארגון העובד על בסיס נתונים צריך בראש (מדדי ביצוע מרכזיים -) מדדים כמותיים המייצגים את מידת השגת KPI-ובראשונה ללמוד לקבוע את המטרות האסטרטגיות והמבצעות.

אישי KPI הכוללת מדדים בפועל ובתכנון. לדוגמה, כדי לחשב, (7.15 איור) בדרך כלל משתמשים בנוסחה KPI לחישוב

עבור פרויקט, עובד או תהליך, יש לחלק את המדדים בפועל במדדים המתוכננים ולהכפיל את התוצאה המתקבלת ב-100%.

$$\text{index KPIs} = \frac{\text{actual performance} \times 100}{\text{target performance}}$$


משמשים למדידת הצלחת הפרויקט או התהליך בהשגת המטרות המרכזיות KPI 7.15 איור

מפורטים יותר KPI ברמת אתר הבנייה ניתן להשתמש במדדי

■ זמני ביצוע שלביים מרכזיים (יסוד, התקנה, גימור) – מאפשרים לשלוט על עמידה בתוכניות העבודה

■ אחוז חריגות בחומרים – מסייע בניהול הרכישות ומזעור הפסדים

■ מספר השבתות הלא מתוכננות של ציוד – משפיע על הפרודוקטיביות וההוצאות

לדוגמה, אם. (2.25 איור) "בחירה במדדים לא נכונים עלולה להוביל להחלטות שגויות לגבי השאלה "מה לעשות החברה מתמקדת רק בעלות למטר רבוע, מבלי לקחת בחשבון את ההוצאות על תיקונים, חיסכון בחומרים עלול להוביל להידרדרות באיכות ולעלייה בהוצאות בפרויקטים עתידיים.

כאשר קובעים מטרות, חשוב להגדיר בבירור מה בדיוק נמדד. ניסוחים מעורפלים מביאים למסקנות לא נכונות ומקשים מוצלחים ולא מוצלחים בבנייה KPI על הבקרה. נבחן דוגמאות למדדי

טובים KPI מדדי

👉 "עד סוף השנה להפחית את אחוז התיקונים בעבודות הגימור ב-10%"

👉 "להגביר את מהירות ההתקנה של חזיתות ב-15% מבלי להפחית באיכות עד לרבעון הבא"

👉 "להפחית את זמן השבתת הציוד ב-20% באמצעות אופטימיזציה של לוחות הזמנים עד סוף השנה"

מדדים אלו ניתנים למדידה ברורה, יש להם ערכים ספציפיים ומסגרות זמן

רעים KPI מדדי

👉 "נבנה מהר יותר" כמה מהר יותר? מה זה אומר "מהר יותר"?"

👉 "נשפר את איכות עבודות הבטון" איך בדיוק נמדדת האיכות?"

👉 "נשפר את שיתוף הפעולה בין הקבלנים באתר" אילו קריטריונים יראו שיפור?"

טוב הוא כזה שניתן למדוד ולהעריך באופן אובייקטיבי. בבנייה זה חשוב במיוחד, מכיוון שבלי KPI מדד מדדים ברורים אי אפשר לשלוט על היעילות של העבודה ולהשיג תוצאות יציבות.

(החזר על השקעה) – מדד המייצג את היחס בין הרווחים ROI: יש מדד נוסף להערכת יעילות ההשקעות, KPI-בנוסף ל מאפשר להעריך עד כמה ההשקעה בשיטות, טכנולוגיות או כלים חדשים הייתה מוצדקת: ROI-להוצאות שהושקעו. ה. ועד השימוש בחומרים בנייה חדשים. מדד זה מסייע בקבלת (7.32 למשל איור) מהפתרונות הדיגיטליים ואוטומציה -החלטות מבוססות לגבי השקעות עתידיות, בהתבסס על השפעתן הממשית על רווחיות העסק.

(החזר השקעה) יכול לשמש כאחד המדדים המרכזיים להצלחה ROI-בהקשר של ניהול פרויקטי בנייה, ה, אם המטרה של החברה היא להעריך את החזר ההשקעה בפרויקט, טכנולוגיה או שיפור תהליכים, (KPI) יכול להראות עד כמה היא שיפרה את הרווחיות ROI-לדוגמה, אם מיושמת שיטה חדשה לניהול הבנייה, ה.

בהתבסס על נתונים שנאספו ממקורות שונים, כגון הוצאות חומרים, שעות עבודה ROI ו-KPI מדידה סדירה של מדדי והוצאות, מאפשרת למנהלי הפרויקטים לנהל את המשאבים ביעילות ולקבל החלטות במהירות. שמירה על נתונים אלו בטווח הארוך מאפשרת לנתח מגמות עתידיות ולאופטימיזציה של תהליכים.

ומדדים אחרים נעשה שימוש בגרפים ודיאגרמות שונות, אשר בדרך כלל מאוגדות ROI, KPI לצורך ויזואליזציה של בלוחות מחוונים.

לוחות מידע ודשבורדים: ויזואליזציה של מדדים לניהול יעיל

לצורך ויזואליזציה של מדדים ומטריקות נעשה שימוש בגרפים ודיאגרמות מגוונות, אשר בדרך כלל מאוגדות בויטריות נתונים ולוחות מחוונים. לוחות מחוונים אלו מספקים ייצוג מרכזי של מצב הפרויקט או חלקיו השונים, מציגים את המדדים המרכזיים (באידיאלי בזמן אמת). (לוחות מחוונים עדכניים ומעודכנים באופן קבוע מאפשרים לצוות להגיב במהירות לשינויים.

לוחות מחוונים הם כלים הממחישים הערכות כמותיות, מה שהופך אותן לנגישות וברורות לכל המשתתפים בפרויקט.



והוויזואליזציה שלהם באמצעות לוחות מחוונים - מפתח לשיפור הביצועים והיעילות של הפרויקט **KPI** ניהול **7.16** ריש

ההנהגה כשהיא דוגמא לכלים פופולריים בהם ניתן ליצור לוחות מחוונים

כלי של מיקרוסופט ליצירת דוחות אינטראקטיביים ולוחות מחוונים - Power BI

כלים רבי עוצמה לוויזואליזציה של נתונים וליצירת לוחות מחוונים - Google Data Studio ו- Tableau
ללא צורך בכתיבת קוד

זו ספרייה ליצירת גרפים אינטראקטיביים, ודש - מסגרת ליצירת - (7.16, ריש, 7.212 ריש) Plotly
 --אפליקציות אינטרנט לניתוח נתונים. ניתן להשתמש בהם בשילוב ליצירת לוחות מחוונים אינטראקטיביים

ישנן רבות ספריות פתוחות והינמיות Python-ב (7.211 ריש - 7.29 ריש) Python-ספריות רבות ב ואחרות. ניתן להשתמש בהן Matplotlib, Seaborn, Plotly, Bokeh לויזואליזציה של נתונים, כגון Django או Flask ליצירת גרפים ולשלבם באפליקציית אינטרנט באמצעות מסגרות כמו

JavaScript מאפשרות ליצור לוחות מחוונים אינטראקטיביים באמצעות ספריות JavaScript ספריות. ולשלבם באפליקציית אינטרנט, Chart.js או D3.js קוד פתוח, כגון

ויצירת לוחות מחוונים נדרשים נתונים עדכניים ולוח זמנים ברור לאיסוף וניתוח KPI לצורך הערכת המידע

ולוחות מחוונים בתעשיית הבנייה מהווים את הבסיס לגישה אנליטית לניהול פרויקטים. הם לא KPI, ROI, באופן כללי רק עוזרים לעקוב ולהעריך את המצב הנוכחי, אלא גם מספקים תובנות יקרות ערך לתכנון עתידי ואופטימיזציה של תהליכים - תהליכים התלויים ישירות בפרשנות הנתונים וביכולת לשאול את השאלות הנכונות ובזמן הנכון

ניתוח נתונים ואומנות השאלת שאלות

לים התשובות formפרשנות הנתונים היא השלב הסופי של הניתוח, שבו המידע מקבל משמעות ומתחיל "לדבר". כאן מת שלב זה מאפשר לסכם את התוצאות, לזהות דפוסים. (2.25. ריש) "לשאלות המרכזיות": מה לעשות? ו"איך לעשות לקבוע קשרים סיבתיים ולבצע מסקנות על סמך הוויזואליזציה והניתוח הסטטיסטי.

ייתכן שלא רחוק היום שבו תגיע ההבנה שכדי להתפתח כאזרח יעיל באחת מהמדינות הגדולות והמורכבות החדשות המתפתחות כיום, יש צורך גם לדעת לחשב, לחשוב בממוצעים, מקסימומים ומינימומים, כפי שנדרש כיום לדעת לקרוא ולכתוב.

סמואל ס. וילקס, ציטוט מנאום הנשיא בשנת 1951 לאגודה הסטטיסטית האמריקאית –

בהתאם לדו"ח שפורסם על ידי ממשלת בריטניה "אנליטיקה של נתונים ובינה מלאכותית ביישום פרויקטים ממשלתיים" (2024) [83], [יישום אנליטיקה של נתונים ובינה מלאכותית] בינה מלאכותית (מאפשר לשפר באופן משמעותי את תהליכי ניהול הפרויקטים, תוך שיפור דיוק התחזיות של לוחות זמנים ועלויות, וכן הפחתת סיכונים ואי ודאות. במסמך מודגש כי ארגונים ממשלתיים המשתמשים בכלים אנליטיים מתקדמים משיגים ביצועים גבוהים יותר בהשגת יוזמות תשתית.

עסק הבנייה המודרני, הפועל בתנאי תחרות גבוהה ורווחיות נמוכה במסגרת המהפכה התעשייתית הרביעית, ניתן להשוות לפעולות צבאיות. כאן ההישרדות וההצלחה של החברה תלויות במהירות השגת המשאבים ומידע איכותי – כלומר (7.17 איור) בקבלת החלטות בזמן ובצורה מבוססת.

אם אנליזת נתונים היא "מודיעין" המספק סקירה, אז אנליטיקה של נתונים היא "תחמושת" הנדרשת לפעולה. היא עונה על השאלות: מה לעשות? ואיך לעשות זאת, ומבנה את הבסיס להשגת יתרונות תחרותיים בשוק.

אנליטיקה הופכת נתונים מפוזרים למידע מובנה ומשמעותי, שעל בסיסו מתקבלות החלטות.

המשימה של אנליטיקאים ומנהלים היא לא רק לפרש מידע, אלא להציע החלטות מבוססות, לחשוף מגמות, לקבוע קשרים בין סוגי נתונים שונים ולמייין אותם בהתאם למטרות ולמאפייני הפרויקט. באמצעות כלים של ויזואליזציה ושיטות של ניתוח סטטיסטי, הם הופכים את הנתונים לנכס אסטרטגי של החברה.



בסופו של דבר, אנליזת נתונים הופכת את המידע שנאסף למקור לקבלת החלטות **7.17** איור.

כדי לקבל החלטות מבוססות באמת בתהליך האנליטיקה, יש ללמוד לנסח נכון את השאלות המופנות לנתונים. איכות השאלות הללו משפיעה ישירות על עומק התובנות המתקבלות וכך גם על איכות ההחלטות הניהוליות.

העבר קיים רק במידה שהוא נוכח ברשומות של היום. ומה שהרשומות הללו מייצגות נקבע על ידי השאלות שאנו [134] שואלים. אין היסטוריה אחרת מלבד זו

ג'ון ארצ'באלד וילר, פיזיקאי 1982 –

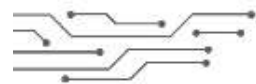
האמנות של לשאול שאלות עמוקות ולחשוב באופן ביקורתי היא מיומנות חשובה בעבודה עם נתונים. רוב האנשים נוטים לשאול שאלות פשוטות, שטחיות, שהתשובות עליהן אינן דורשות מאמצים משמעותיים. עם זאת, אנליזה אמיתית מתחילה בשאלות משמעותיות ומחושבות, המסוגלות לחשוף קשרים נסתרים ותלות סיבתית במידע, שעשויים להסתתר מאחורי מספר שכבות של טיעונים.

(BCG, 2017) "בהתאם למחקר "טרנספורמציה מונחית נתונים: האצה בקנה מידה כבר עכשיו טרנספורמציה דיגיטלית מוצלחת דורשת השקעות ביכולות אנליטיות, תוכניות ניהול שינוי והתאמת מטרות חברה שמביאה תרבות מונחית נתונים צריכה להשקיע ביכולות שימוש בנתונים. IT. עסקיות עם יוזמות אנליטיים ולהשיק תוכניות ניהול שינוי כדי להנחיל חשיבה, התנהגות ודרכי עבודה חדשות

ללא השקעות בפיתוח תרבות אנליטית, שיפור כלים לעבודה עם נתונים והכשרת מומחים, חברות בעתיד מסתכנות בקבלת - HiPPO. החלטות על בסיס מידע מיושן או לא שלם - או בהסתמכות על דעות סובייקטיביות של מנהלי

ההכרה בחשיבות ובצורך בעדכון מתמיד של אנליטיקה ולוחות מידע מובילה את ההנהלה להבנה של חשיבות אוטומציה של תהליכי אנליטיקה. אוטומציה מעלה את מהירות קבלת ההחלטות, מפחיתה את השפעת הגורם האנושי ומבטיחה את עדכניות הנתונים. בתנאים של צמיחה אקספוננציאלית של כמויות מידע, המהירות הופכת לא רק ליתרון תחרותי, אלא לגורם מפתח להצלחה מתמשכת.

ETL (Extract, Transform, Load) אוטומציה של תהליכי ניתוח ועיבוד נתונים קשורה באופן בלתי נפרד לנושא הנתונים נשלפים ממקורות שונים, מומרים ETL בדיוק כפי שבתהליך האוטומציה אנו צריכים להמיר נתונים, בתהליך בהתאם לדרישות הנדרשות ומועלים למערכות יעד לשימוש נוסף.



פרק 7.2. ETL זרימת נתונים ללא מאמצים ידניים: מדוע נדרש

הפחתת עלויות והאצת העבודה עם נתונים: ETL אוטומציה של

מפסיקים לגדול, למרות עלייה בכמויות הנתונים ובגודל הצוות, ההנהלה של חברות (KPI) כאשר מדדי ביצוע מרכזיים sooner or later, this understanding becomes a stimulus for launching comprehensive automation, the main goal of which is to reduce process complexity, accelerate processing, and decrease dependence on the human factor.

(2022), "בהתאם למחקר של מקינזי" איך לבנות ארכיטקטורת נתונים כדי לעודד חדשנות - היום ומחר חברות המשתמשות בארכיטקטורות נתונים זורמות מקבלות יתרון משמעותי, מכיוון שהן יכולות לנתח מידע בזמן אמת. טכנולוגיות זורמות מאפשרות לנתח הודעות בזמן אמת וליישם תחזוקה חיזוי במפעלים בזכות ניתוח נתוני חיישנים בזמן אמת.

פישוט התהליך הוא אוטומציה שבה פונקציות מסורתיות של ניהול ידני מוחלפות באלגוריתמים ובמערכות.

השאלה של אוטומציה, או ליתר דיוק, מזעור תפקיד האדם בעיבוד נתונים, "היא תהליך בלתי הפיך ורגיש מאוד לכל חברה. מומחים בכל תחום מקצועי לעיתים קרובות לא מעזים לחשוף לחלוטין את השיטות והניואנסים של עבודת עמיתיהם, מודעים לסיכון של אובדן עבודה בסביבה טכנולוגית מתפתחת במהירות.

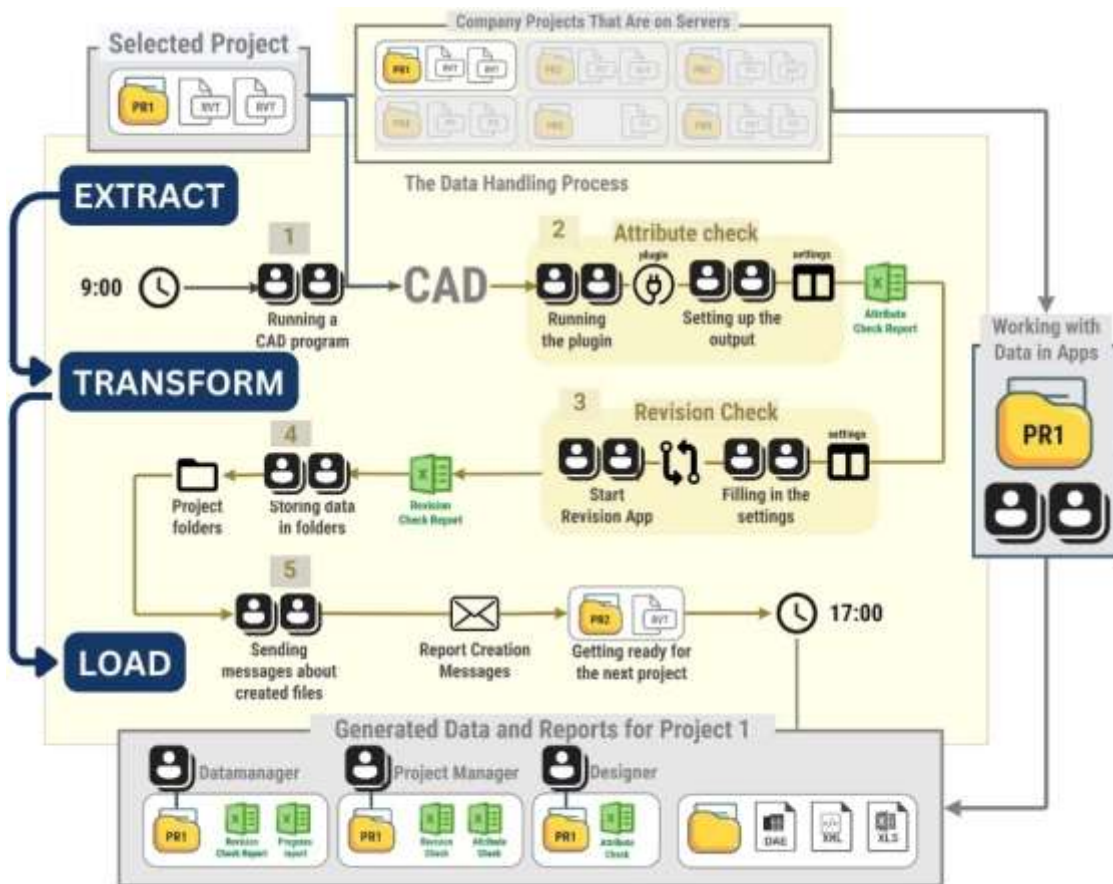
אם אתה רוצה לרכוש אויבים, נסה לשנות משהו

וודרו וילסון, נאום בכנס סוחרים, דטרויט, 1916 -

למרות היתרונות הברורים של אוטומציה, עדיין קיימת בשגרה של חברות רבות רמה גבוהה של עבודה ידנית, במיוחד בתחום העבודה עם נתוני הנדסה. כדי להדגים בצורה ברורה את המצב הנוכחי, נבחן דוגמה טיפוסית של עיבוד נתונים רציף במסגרת תהליכים כאלה.

העיבוד CAD סדר העבודה הידני עם נתונים ניתן להמחיש באמצעות אינטראקציה עם המידע המתקבל ממסדי נתונים ליצירת טבלאות מאפיינים או ליצירת תיעוד על CAD (BIM) במחלקות ("ידני" ETL תהליך) המסורתי של נתונים:-בסיס נתוני פרויקט מתבצע בסדר הבא

1. - CAD (BIM) המשתמש פותח ידנית את הפרויקט - על ידי הפעלת יישום: (Extract) חילוץ ידני
2. אימות: בשלב הבא בדרך כלל מופעלים ידנית מספר תוספים או יישומים עזר, המיועדים להכנת נתונים והערכת איכותם.
3. לאחר ההכנה מתחיל עיבוד הנתונים, שדורש ניהול ידני של כלים תוכנתיים שונים: (Transform) המרה ידנית - שבהם הנתונים מוכנים ליצוא
4. - יצוא ידני של הנתונים המומצים למערכות חיצוניות, פורמטים של נתונים ודוקומנטים: (Load) יצוא ידני



ידינו מסורתי מוגבל לרצונות וליכולות הפיזיות של טכנאי בודד ETL תהליך.

בניגוד לתעשיות אחרות, שבהן (ETL) קלאסי - חילוף, המרה וטעינה ETL תהליך עבודה כזה הוא דוגמה לתהליך אוטומטיים הפכו לסטנדרט, בתעשיית הבנייה עדיין שולטת העבודה הידנית, המאטה את התהליכים ETL צינורות ומגדילה את העלויות.

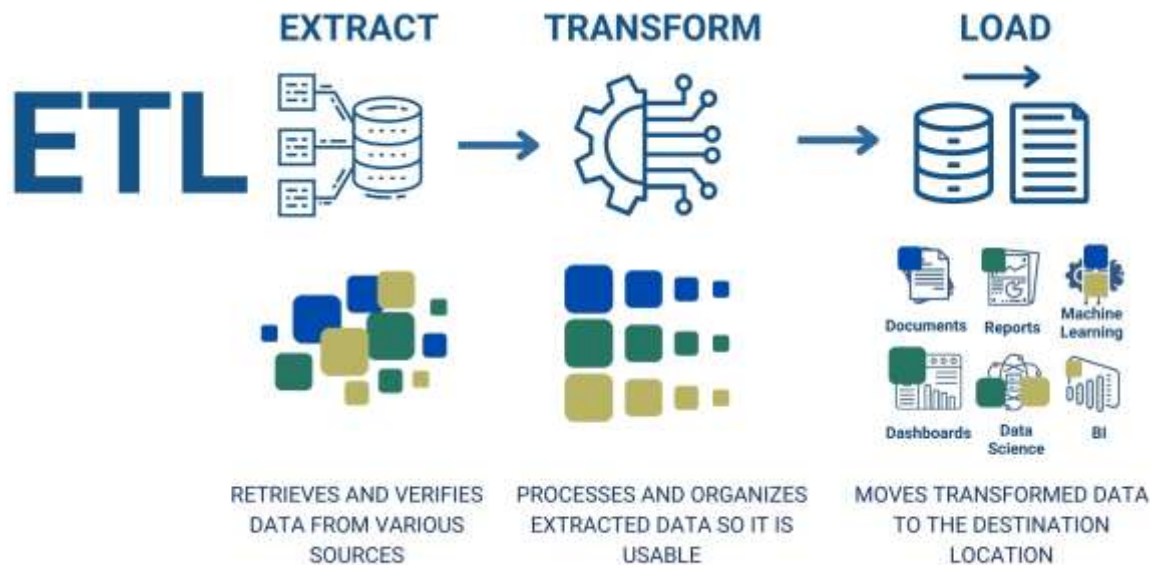
הוא תהליך של חילוף נתונים ממקורות שונים, המרתם לפורמט ETL (Extract, Transform, Load) הנדרש וטעינתם למערכת היעד לצורך ניתוח ושימוש נוסף.

-הוא תהליך המייצג שלושה מרכיבים מרכזיים בעיבוד נתונים: חילוף, המרה וטעינה ETL

- (API, קבצים, מסדי נתונים) חילוף - חילוף נתונים ממקורות שונים
- המרה - ניקוי, אגרגציה, נורמליזציה ועיבוד לוגי של נתונים
- BI טעינה - טעינת מידע מובנה לאחסון נתונים, דוח או מערכת

הוזכר רק באופן אקראי: בהמרת מסמך סרוק לא מובנה לפורמט טבלאי מובנה, בהקשר ETL קודם לכן בספר, המושג את התפיסה של תהליכים חיים ועיסקיים, וכן במסגרת systematize של פורמליזציה של דרישות, המאפשרת לס

בפירוט רב יותר בהקשר של תהליכי ETL כעת נבחן את CAD אוטומציה של בדיקת נתונים ועיבוד נתונים מפתרונות -עבודה טיפוסיים



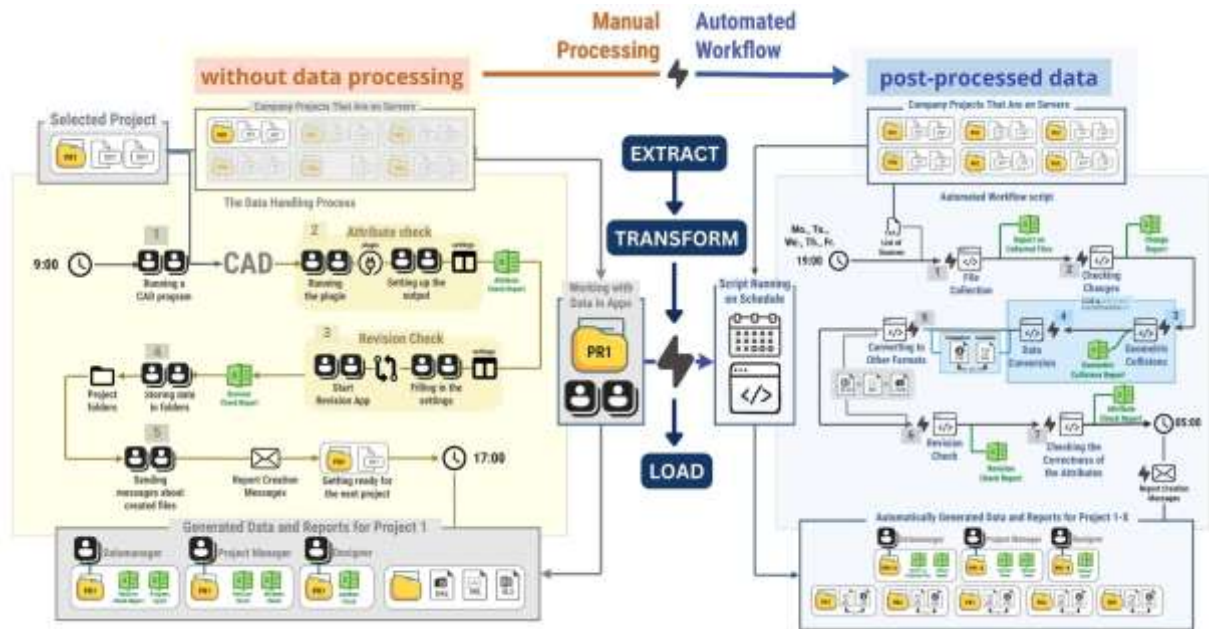
אוטומטית משפרת את המשימות החוזרות של עיבוד נתונים ETL.

ידני או חצי אוטומטי דורש נוכחות של מנהל או טכנאי, המנהל את כל השלבים ידנית - מאיסוף הנתונים ועד ETL תהליך יצירת הדוחות. תהליך כזה לוקח זמן משמעותי, במיוחד בתנאים של יום עבודה מוגבל (למשל, מ-9:00 עד 17:00)

לעיתים קרובות חברות שואפות לפתור את בעיית היעילות הנמוכה והמהירות הנמוכה על ידי רכישת פתרונות מודולריים אשר לאחר מכן משודרגים על ידי ספקים חיצוניים ויועצים. אך ספקים, (ועוד ERP, PMIS, CPM, CAFM) משולבים אלה ומפתחים חיצוניים לעיתים קרובות הופכים לנקודת תלות קריטית: המגבלות הטכניות שלהם משפיעות ישירות על הביצועים של כל המערכת ושל העסק כולו, כפי שתואר בפירוט בפרקים הקודמים על מערכות ופורמטים קנייניים. על "הבעיות שנוצרות מהפיצול ומהתלות, דיברנו בפירוט בפרק "איך עסק הבנייה שוקע בכאוס של נתונים

אם החברה אינה מוכנה ליישום פלטפורמה מודולרית גדולה מאחד הספקים, היא מתחילה לחפש דרכים פתוחים מודולריים, שבהם כל שלב (חילוץ ETL חלופיות לאוטומציה. אחת מהן היא פיתוח צינורות המרה, אימות, טעינה (מתבצע בצורה של סקריפטים המבוצעים לפי לוח זמנים

(איור 7.21) תהליך העבודה נראה כמו קוד מודולרי, שמתחיל בעיבוד נתונים ETL בגרסה האוטומטית של אותו תהליך והמרתם לצורת מבנה פתוחה. לאחר קבלת הנתונים המובנים, מופעלים אוטומטית, לפי לוח זמנים, תרחישים שונים או --מודולים לבדוק שינויים, להמיר ולשלוח הודעות) (איור 7.23)



איור 7.23 משמאל עיבוד ידני, מימין - תהליך אוטומטי, אשר בניגוד לעיבוד הידני המסורתי, אינו מוגבל ביכולות המשתמש.

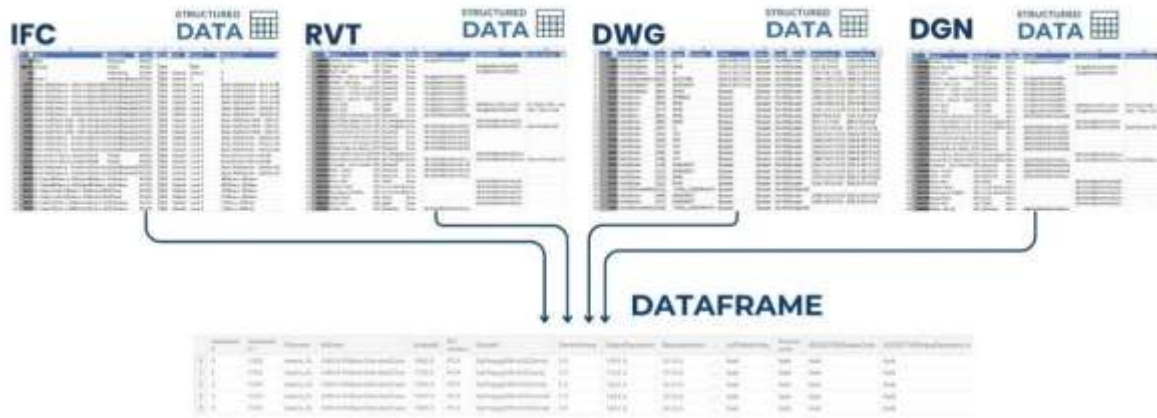
מבנה ואחידות: ET(L) בתהליך העבודה האוטומטי, עיבוד הנתונים מתפשט הודות לעיבוד מקדים של נתוני

בשיטות עיבוד מסורתיות, המומחים עובדים עם הנתונים "כפי שהם" - בצורה שבה הם נמשכים ממערכות ETL, או תוכנה. בתהליכים אוטומטיים, לעומת זאת, הנתונים לעיתים קרובות עוברים קודם כל דרך צינור שבו הם מובאים למבנה ולפורמט אחידים, המתאימים לשימוש ולניתוח נוסף.

המדגימה את תהליך בדיקת טבלאות הנתונים, המתוארת בפרק "בדיקת נתונים ותוצאות ETL, נבחן דוגמה מעשית של לתהליכי ניתוח ועיבוד נתונים LLM בשילוב עם Pandas הבדיקה) "איור 4.413. (לשם כך נשתמש בספריית -אוטומטיים

איסוף נתונים

מתחיל בכתיבת קוד לאיסוף מערכי נתונים, המיועדים לבדיקה ועיבוד - (Extract) חילוץ - ETL השלב הראשון בתהליך לאחר מכן. לשם כך נסרוק את כל התיקיות בשרת העבודה, נאסוף מסמכים בפורמט ובתוכן מסוימים, ולאחר מכן נמיר אותם לצורת מבנה. תהליך זה נבחן בפירוט בפרקים "המרת נתונים לא מובנים וסקסטואליים לצורת מבנה" ו"המרת נתוני --לצורת מבנה) "איור 4.11 - איור 4.112 CAD (BIM)



למסגרת נתונים גדולה אחת, שתכיל את כל חלקי הפרויקט (BIM) CAD איור 7.24 המרת נתוני

(איור 7.24) נעשה (BIM-) CAD- ובקבלת טבלת כל פרויקטי Extract כדוגמה מוחשית, בשלב טעינת הנתונים כדי לקבל טבלאות מובנות מכל, IFC-ו RVT שימוש בממירי נתונים עם תמיכה בהנדסה הפוכה [138] [לפורמטים DataFrame- הפרויקטים ולשלב אותן בטבלה גדולה אחת של

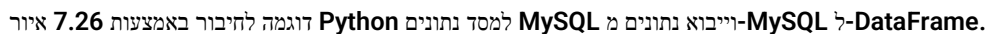


למסגרת נתונים גדולה ומסודרת IFC ו-RVT של כלי הנדסה הפוכה לקבצי Python ו-SDK המרה באמצעות קוד 7.25 איור (df).

קבצי JSON, קבצי Excel, גיליונות, CSV ממקורות שונים, כולל קבצי טקסט Pandas DataFrame-ניתן לטעון נתונים ל MySQL, וכן ממסדי נתונים כמו, HDF5 ו-Parquet פורמטים לאחסון כמויות גדולות של נתונים, כגון, XML ו-API-תומך בטעינת נתונים מ Pandas, ועוד. בנוסף, PostgreSQL, SQLite, Microsoft SQL Server, Oracle, Google BigQuery, Amazon Redshift ו-Snowflake. דפי אינטרנט, שירותי ענן ומערכות אחסון כמו

LLM כדי לכתוב קוד לחיבור ואיסוף מידע ממסדי נתונים, שלחו בקשה טקסטואלית דומה בצ'אט (או כל אחד אחר ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN):

☞ DataFrame-והמרת נתונים ל MySQL-בבקשה, כתוב דוגמה לחיבור ל



מוכן לעיבוד, (7.26 איור, 7.25 איור) ומבנה הנתונים, נוצר מערך מידע אחיד (Extract) לאחר שלב הוצאת הנתונים נוסף. עם זאת, לפני טעינת נתונים אלה למערכות היעד או השימוש בהם לניתוח, יש לוודא באיכותם, שלמותם והתאמתם שלב מפתח, המבטיח את אמינות – (Transform) לדרישות שנקבעו. בשלב זה מתבצע תהליך המרת הנתונים --המסקנות וההחלטות שלאחר מכן

מתבצע עיבוד והמרת הנתונים. תהליך זה יכול לכלול בדיקת תקינות, נורמליזציה, מילוי ערכים Transform-בשלב ה

חסרים ואימות באמצעות כלים אוטומטיים.

"מונע נתונים. מה הסטודנטים צריכים כדי להצליח בעולם העסקים המשתנה במהירות" PwC לפי מחקר חברות ביקורת מודרניות עוברות מבדיקת דגימות של נתונים לניתוח מערך המידע באמצעות [9] (2015) כלים אוטומטיים. גישה זו מאפשרת לא רק לזהות אי-התאמות בדיווחים, אלא גם להציע המלצות לאופטימיזציה של תהליכי עסק.

בבנייה ניתן ליישם שיטות דומות, למשל, לאימות אוטומטי של נתוני פרויקט, בקרת איכות הבנייה והערכת יעילות העבודה של קבלנים. אחד הכלים המאפשרים לאוטומט ולהאיץ את עיבוד הנתונים הוא השימוש בביטויים רגולריים, מאפשר לבדוק ביעילות מחרוזות נתונים RegEx. ETL בתהליך (Transform) בשלב המרת הנתונים (RegEx) (איור 4.47) דנו בפרק RegEx לזהות אי-התאמות ולהבטיח את שלמות המידע עם מינימום משאבים. פרטים נוספים על "-. תרגום דרישות לצורת מבנה"

מנהל קובע דרישות עבור מאפיינים מרכזיים של נכסים (RPM) נבחן דוגמה מעשית: במערכת ניהול נכסי נדל"ן -בשלב ההמרה יש לבצע אימות של הפרמטרים הבאים. (7.27)

- ("ID" מאפיין) בדיקת פורמטים של מזהי נכסים
- בקרת ערכי תקופת האחריות להחלפה (מאפיין "תקופת אחריות")
- בדיקת מחזור החלפת רכיבים (מאפיין "דרישות שירות")

Property Manager		Property Manager: Long-term Management		
ID	Element	Warranty Period	Replacement Cycle	Maintenance Requirements
W-NEW	Window	-	20 years	Annual Inspection
W-OLD1	Window	8 years	15 years	Biannual Inspection
W-OLD2	Window	8 years	15 years	Biannual Inspection
D-122	Door	15 years	25 years	Biennial Varnishing

איור 7.27 בדיקת איכות מתחילה בהגדרת דרישות למאפיינים ולערכים הגבוליים שלהם.

יכולים "ID" כדי לקבוע ערכים גבוליים לבדיקה של הפרמטרים, נניח מניסיונו כי הערכים המותרים עבור המאפיין, או ערכים דומים, כאשר התו הראשון הוא אות "D-122" או "W-OLD1", "W-NEW" לכלול רק ערכים מחרוזתיים או כל מספר בן שלוש ספרות (איור 7.27). (לצורך 'NEW', 'OLD' אחריו מקף ולאחר מכן שלושה תווים אלפביתיים -הבא (RegEx) אימות מזהים אלו ניתן להשתמש בביטוי רגולרי

`^W-NEW$|^W-OLD[0-9]+$|^D-1[0-9]{2}$`

תבנית זו מאפשרת לוודא שכל המזהים בנתונים עומדים בקריטריונים שנקבעו. אם ערך כלשהו לא עובר את הבדיקה,

להמרת נתונים ולהשתמש בנתונים המתקבלים ליצירת טבלת תוצאות Python המערכת רושמת שגיאה. כדי ליצור קוד LLM פשוט ננסה בקשה בצ'אט

LLM-הבקשה הטקסטואלית ב

באמצעות ביטויים רגולריים, הבדקים מזהים DataFrame כתוב קוד לבדוק את העמודות של תקופת 'G' עד 'A'-יעילות אנרגטית עם אותיות מ, RegEx דרך 'W-OLD' או 'W-NEW' בפורמט אחריות ומחזור החלפה עם ערכים מספריים בשנים

LLM: תשובת



```

1 # RegEx
2 import re
3
4 # Define regex patterns for each column
5 patterns = {
6     'Window_ID': r'^{W-NEW|W-OLD[0-9]+|D-1[0-9]{2}}$',
7     'Energy_Efficiency': r'^{A-G}\d*$',
8     'Warranty_Period': r'^{d+\syears}$',
9     'Replacement_Cycle': r'^{d+\syears}$'
10 }
11
12 # Define a verification function using regex
13 def verify_data(column, pattern):
14     # Apply the regex pattern to the series and return the boolean results
15     return column.apply(lambda x: bool(re.match(pattern, str(x))))
16
17 # Apply the verification function to the DataFrame for each column with its respective pattern
18 for column, pattern in patterns.items():
19     df['verified_' + column] = verify_data(df[column], pattern)
20

```

DataFrame. על עמודות הפרמטרים של RegEx איור 7.28 קוד לאוטומציה של תהליך הבדיקה באמצעות יישום תבניות

כדי להגדיר (RegEx ביטויים רגולריים) "re" (איור 7.28) (משתמש בספריית Python-הקוד שנוצר אוטומטית ב עבור כל עמודה) מאפיין (שצוינה, הפונקציה מיישמת את תבנית DataFrame-פונקציה, הבדוקת כל מאפיין של נתון ב כדי לבדוק את ההתאמה של כל רשומה לפורמט הצפוי ומוסיפה את התוצאות בצורה של ערכים חדשים RegEx DataFrame-לעמודת מאפיין חדשה ב (False/True))

בדיקה אוטומטית כזו מבטיחה התאמה פורמלית של הנתונים לדרישות שנקבעו ויכולה לשמש כחלק ממערכת בקרת איכות בשלב ההמרה

ובדיקת האיכות, הנתונים מוכנים לטעינה במערכות היעד. הנתונים שהומרו Transform-לאחר ביצוע מוצלח של שלב ה מסדי נתונים ופורמטים נוספים לשימוש עתידי CSV, JSON, Excel, ובבדקו יכולים להיות מיוצאים בפורמטים כמו. בנוסף, בהתאם למשימה, התוצאות יכולות להיות מוצגות בדוחות, גרפים או דשבורדים אנליטיים.

וויזואליזציה של התוצאות בצורה של דיאגרמות וגרפים ETL: טעינת

שבו ניתן, Load - כאשר הנתונים הובאו לצורת מבנה ובבדקו, מגיע השלב הסופי, Transform-לאחר סיום שלב ה, לטעון את הנתונים במערכת היעד או להציגם לניתוח. הייצוג הוויזואלי של הנתונים מאפשר לזהות במהירות חריגות. לנתח הפצות ולהעביר את המסקנות המרכזיות לכל המשתתפים בפרויקט, כולל אלו שאינם בעלי הכשרה טכנית.

(dashboards) במקום להציג מידע בטבלאות ובמספרים, אנו יכולים להשתמש באינפוגרפיקה, גרפים ולוחות מחוונים Matplotlib. הוא הספרייה Python אחד הכלים הנפוצים והגמישים ביותר לוויזואליזציה של נתונים מובנים בשפת -- היא מאפשרת ליצור גרפים סטטיים, מונפשים ואינטראקטיביים, ותומכת במגוון רחב של סוגי דיאגרמות.

ניתן להשתמש בשאילתא הבאה למודל RPM, כדי להציג את תוצאות בדיקת האטריבוטים ממערכת השפתי:

שהוזכרו לעיל, עם היסטוגרמה לתוצאות, כדי DataFrame כתוב קוד לוויזואליזציה של נתוני - להציג את תדירות השגיאות באטריבוטים

של תוצאות ביצוע הקוד LLM בקוד וביזואליזציה מוכנה ישירות בצ'אט LLM תשובת

Create in LLM Chat

Pipeline.py

```

1 # Re-importing necessary libraries for visualization
2 import matplotlib.pyplot as plt
3 import seaborn as sns
4
5 # Visualization 1: Bar Chart
6 plt.figure(figsize=(10, 6))
7 df_visual.plot(kind='bar', stacked=True, color=['green', 'red'])
8 plt.title('Data Verification Summary - Bar Chart')
9 plt.xlabel('Data Categories')
10 plt.ylabel('Count')
11 plt.xticks(rotation=45)
12 plt.tight_layout()
13 plt.show()

```



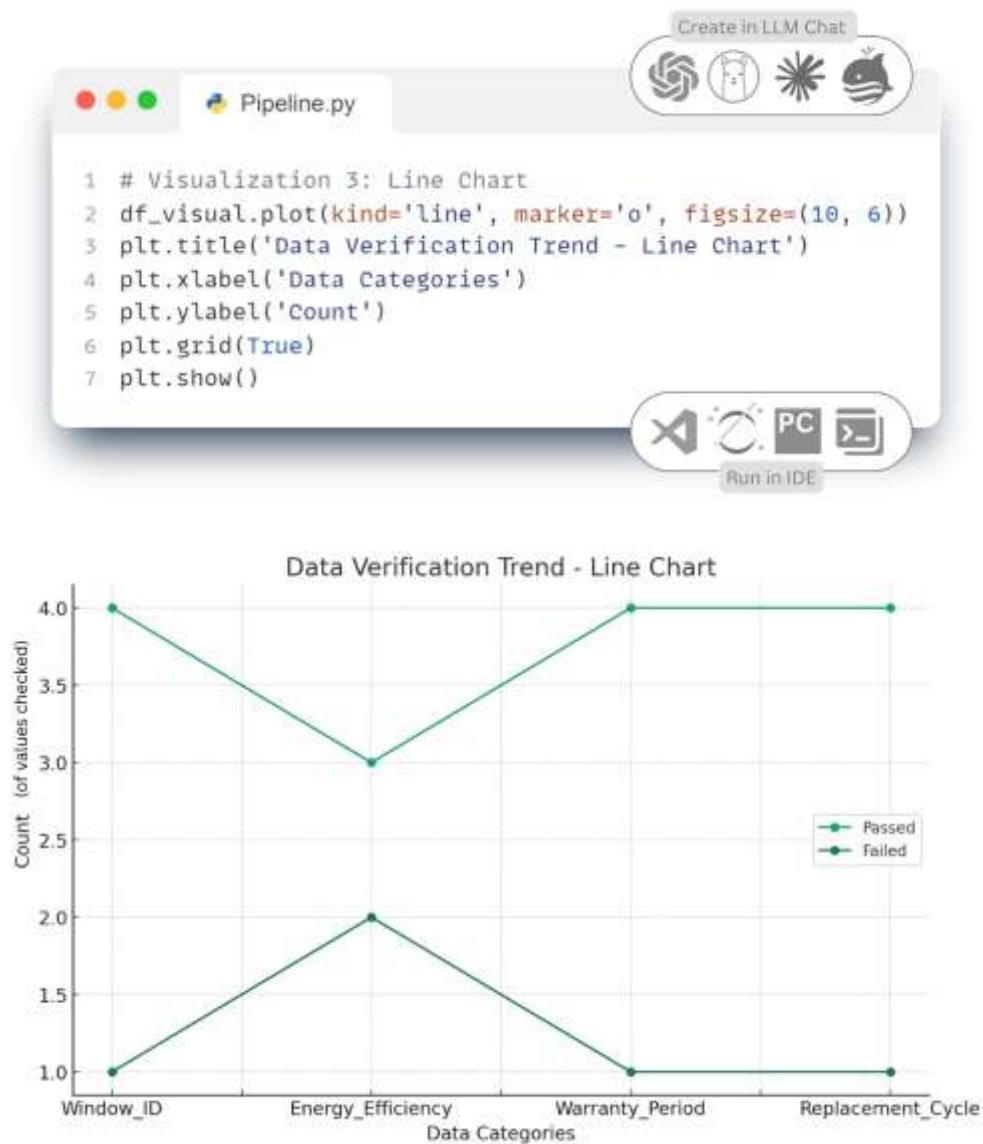
-Load-בהיסטוגרמה בשלב ה **RPM** בבדיקת ערכי האטריבוטים ממערכת **Transform**-ויזואליזציה של תוצאות שלב ה

ישנן רבות ספריות ויזואליזציה פתוחות וחינמיות המאפשרות להציג נתונים מובנים במגוון פורמטים

נמשיך בויזואליזציה של התוצאות בסוג גרף אחר בשאלתא הבאה בצ'אט:

הצג את אותם נתונים בגרף קווי

LLM: תשובת 🐙



Matplotlib - ויזואליזציה של נתוני הבדיקה בגרף קווי, שהושגה באמצעות ספריית

יישן רבות ספריות ויזואליזציה פתוחות וחינמיות, כגון:

- עיבור גרפים סטטיסטיים - **Seaborn**

■ --עבור ויזואליזציות אינטראקטיביות באינטרנט - Plotly

■ .עבור ויזואליזציה דקלרטיבית - Altair

■ .ליצירת דשבורדים שלמים - Dash או Streamlit

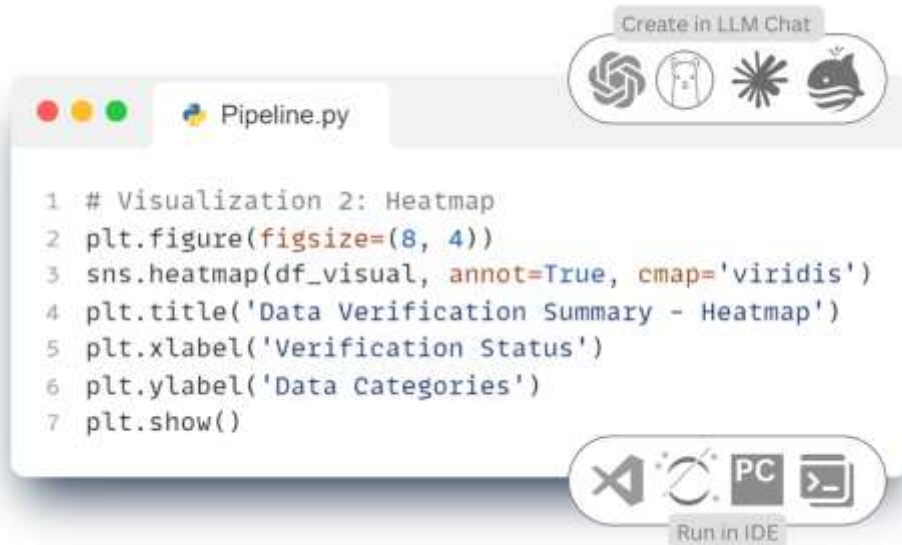
מאפשרים ליצור אוטומטית, LLM עם זאת, ידע בספריות ספציפיות לויזואליזציה אינו הכרחי - כלים מודרניים, כולל קוד לבניית גרפים ואפליקציות שלמות, בהתבסס על תיאור המשימה

Seaborn, הבחירה בכלי תלויה במשימות הפרויקט: בין אם מדובר בדו"ח, מצגת או לוח בקרה מקוון. לדוגמה, הספרייה עם קוד פתוח, מצוינת לעבודה עם נתונים קטגוריאליים, ועוזרת לזהות תבניות ומגמות

🗨 להשתמש ישירות בספרייה LLM-בפעולה, אתה יכול לבקש מ Seaborn כדי לראות את הספרייה LLM: הנדרשת או לשלוח שאילתא טקסטואלית דומה בהמשך העבודה עם

הצג מפה חמה עבור התוצאות

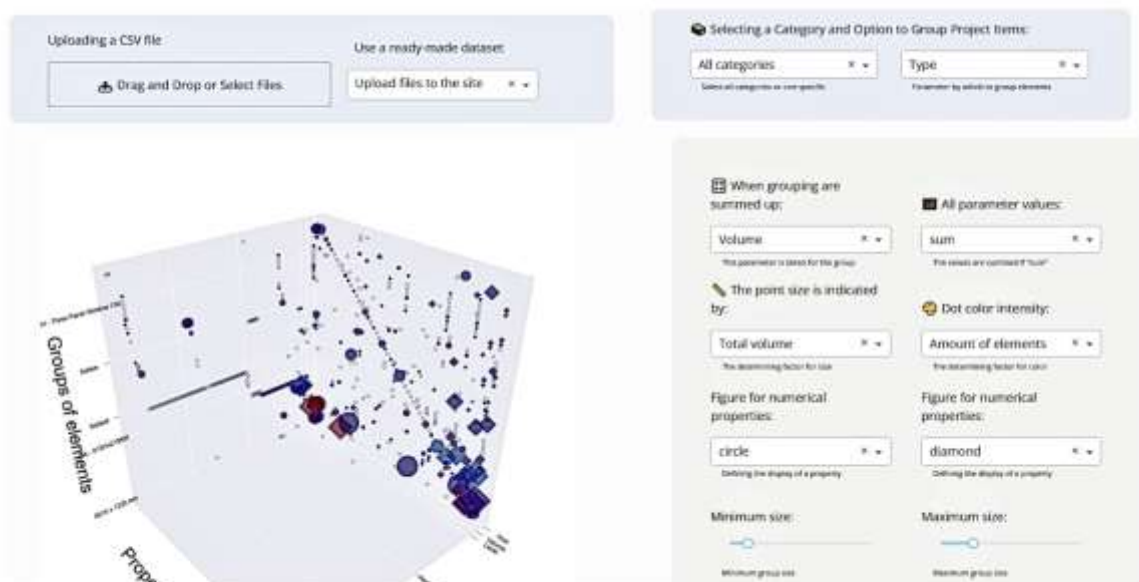
🗨 ואת הגרף ניתן, IDE-בצורה של קוד וגרף מוכן, כאשר ניתן להעתיק את הקוד לבנייתו ל LLM תשובת להעתיק או לשמור לצורך הוספה למסמך





Seaborn באמצעות ספריית (7.28 איור) ויזואליזציה של תוצאות בדיקת הנתונים 7.211 איור

עבור מי שמעדיף גישה אינטראקטיבית, קיימים כלים המאפשרים ליצור דיאגרמות ודשבורדים דינמיים עם אפשרות מציעה אפשרות ליצור דיאגרמות ודשבורדים (7.212 איור, 7.16 איור) Plotly אינטראקציה. ספריית --אינטראקטיביים מאוד, שניתן לשלב בדפי אינטרנט ולספק למשתמש אפשרות אינטראקציה עם הנתונים בזמן אמת.



Plotly באמצעות ספריית **CAD- (BIM-)** ויזואליזציה אינטראקטיבית תלת-ממדית של תכונות רכיבים מפריקט 7.212 איור

מספקות דרך נוחה להציג נתונים, מבלי לדרוש ידע מעמיק Streamlit ו-Dash, ספריית פתוחות מתמחות כמו משמשת לבניית דשבורדים אנליטיים Dash, מתאימה לדיאגרמות אינטראקטיביות מורכבות Bokeh. בפיתוח אתרים מאפשרת ליצור במהירות אפליקציות אינטרנט לניתוח נתונים Streamlit-מלאים, ו

בזכות כלים כאלה לויזואליזציה, מפתחים ואנליסטים יכולים להפיץ ביעילות את התוצאות בין עמיתים ובעלי עניין, תוך כדי הבטחת אינטראקציה אינטואיטיבית עם הנתונים ופישוט תהליך קבלת ההחלטות.

PDF יצירה אוטומטית של מסמכי ETL: טעינת

בשלב הטעינה ניתן לא רק לויזואליזציה של הנתונים, לייצא אותם לטבלאות או למסדי נתונים, אלא גם ליצור אוטומטית דוחות, הכוללים את הגרפים, הדיאגרמות והמדדים האנליטיים המרכזיים שצריך לקבל המנהל או המומחה הממתין לתוצאות הבדיקה. דוחות אוטומטיים יכולים לכלול גם הערות ופרשנות טקסטואלית של הנתונים, כמו גם אלמנטים ויזואליים ברורים – טבלאות, גרפים.

והיאור הניתוח על תוצאות הבדיקה שביצענו בפרקים (7.29 איור) עם היסטוגרמה PDF כדי ליצור דוח
:-לדוגמה, LLM, הקודמים, מספיק לנסח בקשה בהמשך השיחה עם

עם היסטוגרמה והיאור תוצאות בדיקת הנתונים שהתקבלו למעלה PDF כתוב קוד ליצירת קובץ
(בשיחה, (וכתוב אזהרה בטקסט על כך שכמה קטגוריות לא עברו את הבדיקה ושהצורך למלא את
הנתונים החסרים

מזכיר: מוכן עם התוצאות PDF בצורה של קוד וקובץ LLM תשובת

Pipeline.py

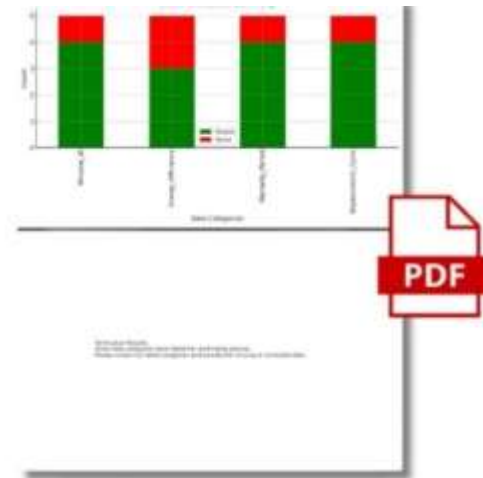
Create in LLM Chat

```

1 # Create a PDF document with the verification results
2 with PdfPages('verification_results.pdf') as pdf:
3     # Adding a histogram to the document
4     plt.figure()
5     df_pdf.plot(kind='bar', stacked=True, colour=['green', 'red'])
6     plt.title('Data Verification Summary')
7     plt.xlabel('Data Categories')
8     plt.ylabel('Count')
9     plt.tight_layout()
10    pdf.savefig()
11    plt.close()
12
13    # Adding a text page with a description
14    plt.figure()
15    plt.axis('off')
16    text_str = ("Verification Results:\n"
17               "Some data categories have failed the verification process.\n"
18               "Please review the failed categories and provide the missing data.")
19    plt.text(0.1, 0.5, text_str, wrap=True)
20    pdf.savefig()
21    plt.close()
22

```

Run in IDE



המכיל היסטוגרמה עם נתוני בדיקה וטקסט עם תוצאות הבדיקה, PDF קוד אוטומטי יוצר מסמך 7.213 איור

הנדרש (DOC או) PDF-יוצר מיד את קובץ ה LLM פתרון שנכתב אוטומטית, המורכב מ-20 שורות קוד באמצעות המראה את כמות הנתונים שעברו ולא עברו את (7.213 איור) עם ויזואליזציה בצורה של היסטוגרמה של תכונות הבדיקה, כמו גם הוספת בלוק טקסט עם תיאור קצר של התוצאות והמלצות לפעולות נוספות יצירת מסמכים אוטומטית היא מרכיב מרכזי בשלב הטעינה, במיוחד בתנאים של פעילות פרויקטים, שבהם מהירות הכנת הדו"ח ודיוקן הם קריטיים

FPDF יצירה אוטומטית של מסמכים עם ETL טעינת

היא שלב חשוב בעיבוד נתונים, במיוחד כאשר התוצאות של הניתוח צריכות ETL Load אוטומציה של דיווחים בשלב להיות מוצגות בפורמט נוח להעברה ולהבנה. בענף הבנייה, זה לעיתים קרובות רלוונטי לדיווחים על התקדמות העבודה סטטיסטיקות של נתוני פרויקטים, דיווחים על בדיקות איכות או מסמכים פיננסיים הזמינה גם לפיתוח וגם לפי.אצ'פי, FPDF, אחד הכלים הנוחים ביותר למשימות כאלה הוא הספרייה עם הקוד הפתוח

מספקת דרך גמישה ליצירת מסמכים באמצעות קוד, ומאפשרת הוספת כותרות FPDF הספרייה הפתוחה טקסט, טבלאות ודימויים. השימוש בקוד במקום עריכה ידנית מפחית את מספר השגיאות ומאיץ את תהליך PDF הכנת הדוחות בפורמט

הוא הוספת כותרות וטקסט עיקרי בצורה של הערות או תיאורים. עם PDF אחד מהשליבים המרכזיים ביצירת מסמך זאת, כאשר מכינים דוח, חשוב לא רק להוסיף טקסט, אלא גם למבנו בצורה נכונה. כותרות, רווחים, מרווחים בין השורות ניתן לקבוע פרמטרי עיצוב, לנהל את מיקום האלמנטים, FPDF כל אלה משפיעים על קריאות המסמך. באמצעות - ולהתאים את סגנון המסמך

כדי FPDF-יכול בקלות להשתמש ב HTML מי שכבר מכיר את HTML-דומה מאוד בעקרון העבודה שלו ל FPDF הוספת: HTML בכל רמה של מורכבות, מכיוון שהמבנה של הקוד מזכיר במידה רבה את הסימון של PDF לייצר מסמכי לא צריך לדאוג - ניתן להשתמש HTML, כותרות, טקסט, תמונות וטבלאות מתבצעת בצורה דומה. מי שאינו מכיר את

שיכול מיד לעזור לנסח את הקוד ליצירת העיצוב הנדרש של המסמך, LLM-ב.

📎 התומך IDE הדוגמה הבאה מדגימה כיצד ליצור דוח עם כותרת וטקסט עיקרי. הרצת הקוד הזה בכל המכיל את הכותרת והטקסט הנדרשים PDF בפייתון יוצרת קובץ.

```
PDF יוצרים מסמך # FPDFpdf = FPDF() # מייבאים את ספריית FPDF מייבאים את ספריית
pdf.הוספת_עמוד() # להוסיף עמוד

קובעים גופן: אריאל, מודגש, גודל # pdf.set_font("Arial", style='B', size=16)
יוצרים כותרת ומרכזים # pdf.cell(200, 10, "דו"ח פרויקט", ln=True, align='C')
רגיל, גודל 12 המסמך הזה מכיל נתונים על תוצאות ARIAL-משנים את הגופן ל # pdf.set_font("Arial", 12, גודל)
PDF-שומרים את קובץ ה # pdf.output(r"C:\reports\report.pdf") בדיוק קבצי הפרויקט
```



הנדרש לנו עם הטקסט PDF-איור 7.214 בעזרת מספר שורות קוד בפייתון אנו יכולים לייצר באופן אוטומטי את מסמך ה.

בעת הכנת דוחות חשוב לקחת בחשבון שהנתונים שממנו הם נוצרים, מסמך נדיר נשאר סטטי. כותרות, בלוקים בתהליך Transform-לעיתים קרובות נוצרים באופן דינמי, מקבלים ערכים עוד בשלב ה (7.214 איור) טקסטואליים ה-ETL.

שימוש בקוד מאפשר ליצור מסמכים המכילים מידע עדכני: שם הפרויקט, תאריך יצירת הדו"ח, כמו גם מידע על המשתתפים או הסטטוס הנוכחי. השימוש במשתנים בקוד מאפשר להכניס אוטומטית את הנתונים הללו במקומות הנדרשים בדו"ח, ובכך לחלוטין למנוע את הצורך בעריכה ידנית לפני שליחתו.

בנוסף לטקסט פשוט ולכותרות, בטכנולוגיית התכנון טבלאות תופסות מקום מיוחד. כמעט בכל מסמך יש נתונים מובנים מאפשרת Transform-מתיאור אובייקטים ועד תוצאות בדיקות. יצירת טבלאות אוטומטית על בסיס נתונים משלב ה מאפשרת להוסיף טבלאות לקבצי FPDF. לא רק לזרז את תהליך הכנת המסמכים, אלא גם למזער טעויות בהעברת מידע זה במיוחד נוח. (7.215 איור) (כטקסט או כתמונות, תוך הגדרת גבולות תאים, גדלים של עמודות וגופנים PDF -בעבודה עם נתונים דינמיים, כאשר מספר השורות והעמודות יכול להשתנות בהתאם למשימות המסמך.

📎 הדוגמה הבאה מראה כיצד לאוטומט את יצירת הטבלאות, לדוגמה, עם רשימת חומרים, חישובי עלויות או תוצאות בדיקות פרמטרים.

```
data = [ ["רכיב", "כמות", "מחיר", "כותרות העמודות"], # מחיר, "כמות", "רכיב",
        ["$500.", # נתוני השורה הראשונה

        ["", "ברזל", "2 טון", "$600", # "נתוני השורה השנייה
        ["$750"] # נתוני השורה השלישית

pdf = FPDF() # יוצרים מסמך PDFpdf.add_page() # מוסיפים עמודpdf.set_font("Arial",
        size=12) # קובעים גופן

for row in data: # מעבירים על פני שורות הטבלהfor item in row: # מעבירים על
        פני תאים בשורה
pdf.cell(60, 10, item, border=1) # יוצרים תא עם גבול, רוחב 60 וגובה 10pdf.ln() #
        שומרים את קובץ ה-PDFpdf.output(r"C:\reports\table.pdf") # עוברים לשורה הבאה
```

Item	Quantity	Price
Concrete	10 m ³	\$500
Rebar	2 t.	\$600
Brick	5000 pcs.	\$750



Transform-לא רק טקסט, אלא גם כל מידע טבלאי שהתקבל בשלב ה-PDF ניתן ליצור אוטומטית ב-7.215 איור

בתרחישים אמיתיים של דיווח, טבלאות בדרך כלל מייצגות מידע שנוצר באופן דינמי, שהתקבל בשלב טרנספורמציה של (השורה הראשונה) **data** בצורה סטטית: במילון PDF הטבלה מוכנסת למסמך (7.215 איור) הנתונים. בדוגמה המוצגת מתמלא אוטומטית לאחר, לדוגמה, קיבוץ של **data** של הקוד (הוכנסו נתונים לדוגמה, בתנאים אמיתיים משתנה דומה **DataFrame**.

בפועל, טבלאות כאלה לרוב נבנות על בסיס נתונים מובנים המתקבלים ממקורות דינמיים שונים: מסדי נתונים, קבצי הנתונים מאוגדים, מקובצים (ETL) **Transform** או תוצאות חישובים אנליטיים. לרוב בשלב ה-API ממשקי Excel, או מסוננים — ורק לאחר מכן מומרות לתוצר הסופי בצורה של גרפים או טבלאות דו-ממדיות, המוצגות בדיווחים. זה אומר שהתוכן של הטבלה יכול להשתנות בהתאם לפרמטרים שנבחרו, תקופת הניתוח, מסנני פרויקטים או הגדרות משתמש.

עושה את תהליך יצירת **Transform**-דינמיים ובסטי נתונים בשלב ה-**DataFrames** השימוש ב-גמיש, נרחב וחוזר בקלות ללא צורך בהתערבות ידנית **Load**-הדיווחים בשלב ה-

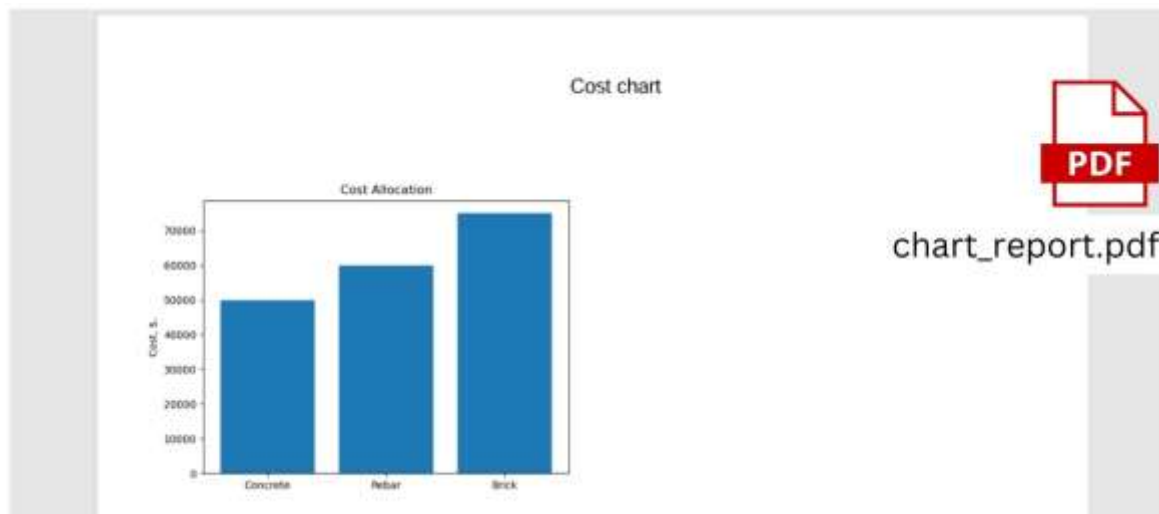
תומכת גם בהוספת גרפים של נתונים טבלאיים, מה שמאפשר לשלב בדו"ח תמונות **FPDF**, בנוסף לטבלאות וטקסט או ספריות ויזואליזציה אחרות, אשר דנו בהן לעיל. ניתן להוסיף למסמך באמצעות קוד **Matplotlib** שנוצרו באמצעות

כל גרף, דיאגרמה או תרשים

Matplotlib גרף, שנוצר מראש באמצעות PDF נוסף למסמך, Python FPDF באמצעות ספריית

```
import matplotlib.pyplot as plt # ליצירת גרפים מדמות ציר Matplotlib מייבאים את
plt.subplots() # [שמות הקטגוריות ערכים] # "לבנים", "ברזל", "בטון" = יוצרים דמות וצירי גרף קטגוריות ["בטון", "ברזל", "לבנים"] # שמות הקטגוריות ערכים
# ערכים לפי קטגוריות

plt.ylabel("$") # מסמנים את ציר y (קטגוריות, ערכים) # יוצרים תרשים עמודות ציר
plt.title("מחירי חומרים") # מוסיפים כותרת
plt.savefig(r"C:\reports\chart\chart.png") # שומרים את הגרף כקובץ תמונה
pdf = FPDF() # מוסיפים עמוד # pdf.add_page() # יוצרים מסמך # FPDF()
pdf.cell(200, 10, "גרף הוצאות", ln=True, align='C') #
pdf.image(r"C:\reports\chart\chart.png", x=10, y=30, w=100) # מוסיפים כותרת
pdf.output(r"C:\reports\chart_report.pdf") # שומרים את
# PDF (x, y - קואורדינטות, w - רוחב) PDF קובץ ה
```



PDF בעזרת עשרה שורות קוד ניתן לייצר גרף, לשמור אותו, ולאחר מכן להכניס אותו למקום הנדרש במסמך 7.216. ריש.

תהליך הכנת המסמכים והלוגיקה הופכים להיות שקופים, מהירים ונוחים. תבניות המובנות בקוד מאפשרות FPDF בזכות לייצר מסמכים עם נתונים עדכניים, תוך כדי הימנעות מהצורך במילוי ידני.

במקום עיצוב ידני של דוחות, המומחים יכולים להתמקד בניתוח נתונים – ETL באמצעות אוטומציה של ובקבלת החלטות, ולא בבחירת הכלי המתאים לעבודה עם סילואי נתונים שונים עם ממשק משתמש ברור.

מספקת כלי גמיש ליצירת מסמכים אוטומטית בכל דרגת מורכבות – מדוחות טכניים קצרים ועד FPDF כך, ספריית סיכומים אנליטיים מורכבים עם טבלאות וגרפים, מה שמאפשר לא רק לזרז את מחזור המסמכים אלא גם להפחית באופן

משמעותי את הסיכון לטעויות הקשורות להזנה ידנית ולעיבוד נתונים.

הכנת דוחות והעלאה למערכות אחרות: ETL טעינת

סופיים, שהוכנו בהתאם לדרישות שנקבעו PDF נוצרו התוצאות בצורה של טבלאות, גרפים ודוחות Load-בשלב ה מה שנדרש לשילוב עם מערכות, (למשל, CSV), לאחר מכן ניתן לייצא את הנתונים הללו לפורמטים קריאים למכונה הייצוא, CSV-ועוד פתרונות ארגוניים או תעשייתיים. בנוסף ל BI פלטפורמות, ERP, CAFM, CPM, חיצוניות – כמו או ישירות לבסיסי נתונים התומכים בהחלפת מידע אוטומטית XML, JSON, XLSX יכול להתבצע לפורמטים

למשל, LLM מספיק להגדיר בקשה בממשק, Load-כדי ליצור קוד מתאים לאוטומציה של שלב ה ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude או QWEN:

שבו העמודות עם הקידומת, DataFrame-כתוב קוד ליצירת דוח על תוצאות בדיקת נתונים ב, ערכים חסרים מוחלפים ב-0, 'Failed' ו-'Passed'-נספרות, משנות את שמן ל 'verified_' רק את השורות שעברו את כל הבדיקות CSV ולאחר מכן מייצאים לקובץ

LLM: תשובות



לשילוב עם CSV הסופי, מייצאים לקובץ DataFrame-מה Transform-נתונים מאומתים, שהתקבלו בשלב ה 7.217. ריש מערכות אחרות.

טעינה, במהלכו נשמרים הנתונים שנבדקו בפורמט - ETL (בקוד המוצג) איור 7.217 (מתבצע השלב הסופי של תהליך המלא, הכולל חילוף ETL התואם לרוב המערכות והמסדי נתונים החיצוניים. כך, סיימנו את מחזור תהליך, CSV, טרנספורמציה, ויזואליזציה, תיעוד וייצוא נתונים למערכות ולפורמטים הנדרשים לנו, מה שמבטיח שחזור, שקיפות ואוטומציה בעבודה עם מידע

שפותח ניתן לשימוש הן לעיבוד פרויקטים בודדים והן ליישום רחב היקף - בעת ניתוח מאות ואלפי נתונים ETL-צינור ה

או מקורות אחרים המגיעים ממערכות PDF ענני נקודות, קבצי CAD נכנסים בצורת מסמכים, תמונות, סריקות, פרויקטי לא רק לכלי לעיבוד טכני, אלא ליסוד התשתית ETL-מפוזרות. האפשרות לאוטומציה מלאה של התהליך הופכת את ה המידע של הבנייה הדיגיטלית.

PDF ויזואליזציה של נתונים מתוך מסמכי LLM באמצעות ETL

- מלא, המכסה את כל השלבים המרכזיים בעבודה עם נתונים בתסריט אחד ETL הגיע הזמן לעבור להרכבת תהליך ללא עבודה ידנית - לחלץ נתונים PDF אוטומטי, המאפשר לעבד קבצי ETL חילוץ, טרנספורמציה וטעינה. נבנה צינור מהמסמכים, לויזואליזציה, לניתוח ולהעביר למערכות אחרות.

ETL-את כל תהליכי ה (LLM) בדוגמה שלנו יתואר באמצעות פקודות, שצריכות להסביר למודל השפה ETL-תהליך ה בתיקייה ובתיקיות המשנה PDF-עם תיאור התוצאה הסופית שיש להשיג. במקרה זה, המשימה היא למצוא את כל קבצי ה שלה, לחלץ מהם מידע רלוונטי - לדוגמה, שמות חומרים, כמות ועלות - ולהציג את התוצאה בצורה של טבלה מסודרת לניתוח עתידי (DataFrame).

ויצירת PDF לחילוץ אוטומטי של נתונים ממספר קבצי LLM-הבקשה הטקסטואלית הראשונה ב נתונים לשלב החילוץ DataFrame:

בתיקייה הנתונה ובתיקיות המשנה שלה. הנתונים PDF כתוב קוד לחילוץ מידע על חומרים מקבצי DataFrame-כוללים את שם החומר, כמות ועלות. התוצאה צריכה להישמר ב PDF-ב

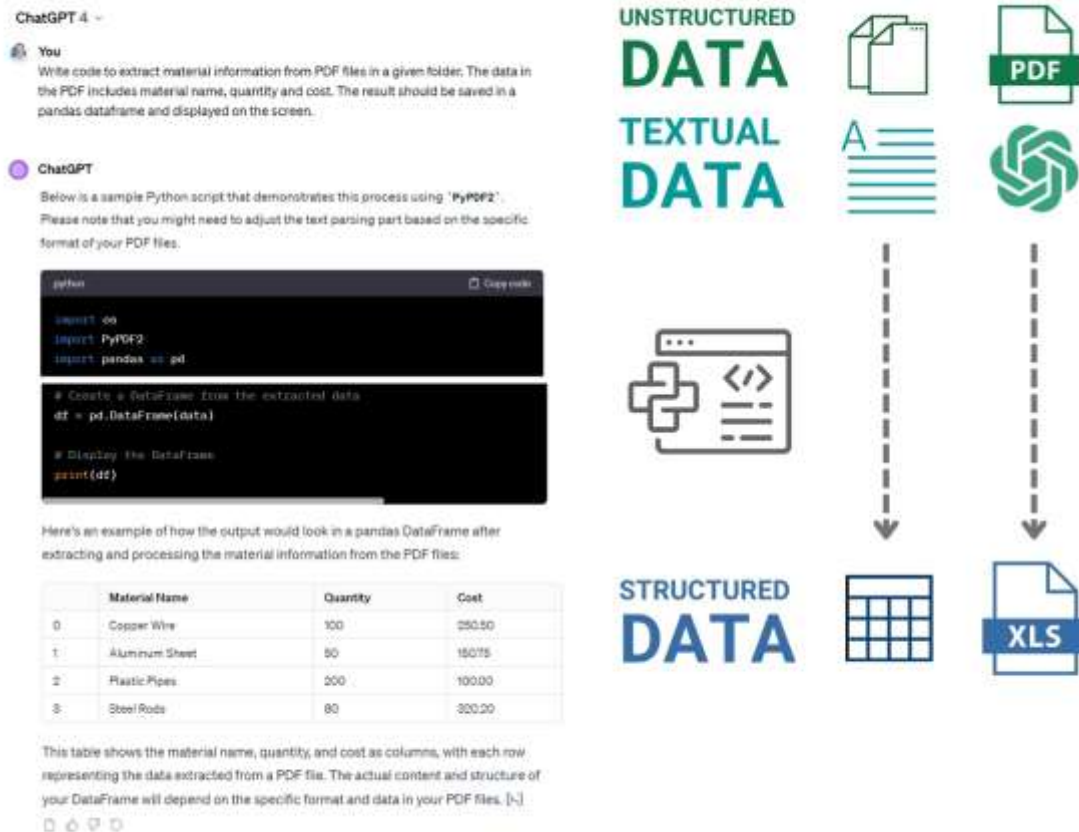
LLM: תשובות



בתיקיה מסוימת ובכל תיקיות המשנה שלה PDF לחילוץ נתונים מקבצי Python יוצר קוד LLM איור 7.218.

מוכן, שמבצע אוטומטית סריקה של כל התיקיות, פותח את Python (איור 7.218) (היא סקריפט LLM התשובה של שנמצאו, מחלץ מהם מידע טקסטואלי וממיר אותו לטבלה. הקוד שהתקבל בשיחה ניתן להרצה באחת מה PDF-קבצי ה-PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse עם תוסף PyDev, Thonny, Wing IDE, IntelliJ IDEA עם תוסף Python, Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker.

אנו יכולים גם להעלות IDE-ולהשתמש בו ב-LLM-כאופציה, כדי לפשט את תהליך האיסוף, במקום להעתיק את הקוד מ-7.219 (ולקבל טבלה, מבלי צורך לראות את הקוד או להריץ אותו LLM ישירות לצ'אט של PDF עשרה קבצי התוצאה של הרצת הקוד הזה תהיה טבלה עם האטריבוטים שבחנו.

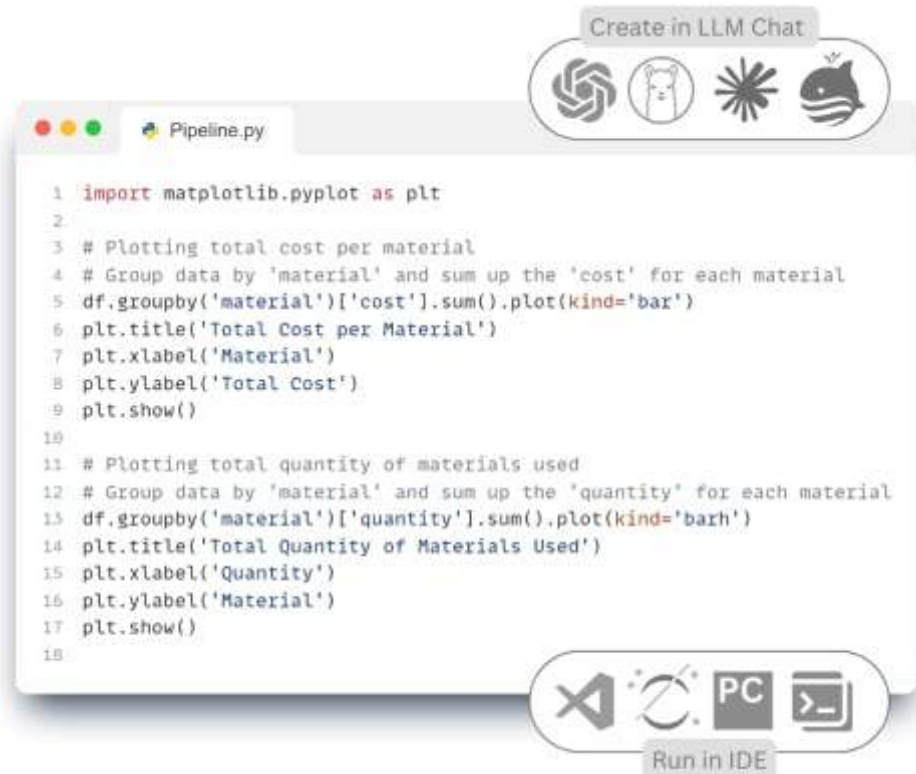


עם האטריבוטים **DataFrame** בצורה מסודרת של **PDF** המחלצת נתונים מקבצי **LLM-איור 7.219** תוצאת הרצת הקוד ב שבהרנו.

בשלב הבא, אנו מבקשים מהמודל הלשוני על סמך הנתונים שהתקבלו - לדוגמה, להשוות בין עלויות ונפחי השימוש בחומרים וליצור מספר דוגמאות לויזואליזציה, שישמשו כבסיס לניתוח עתידי.

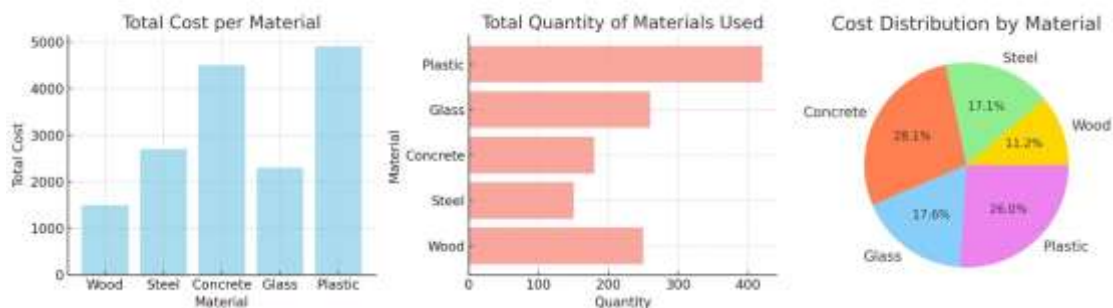
🔗 **Transform**-לבנות מספר גרפים מתוך הטבלאות שהתקבלו בשלב ה **LLM** בקש בהמשך השיחה עם - (איור 7.218)

🔗 - (איור 7.218) **DataFrame**-ויזואליזציה של העלות הכוללת וכמות כל חומר מתוך ה



לצורך ויזואליזציה של הנתונים מתוך מסגרת הנתונים באמצעות ספריית Python בקוד LLM איור 7.220 תשובת מודל matplotlib.

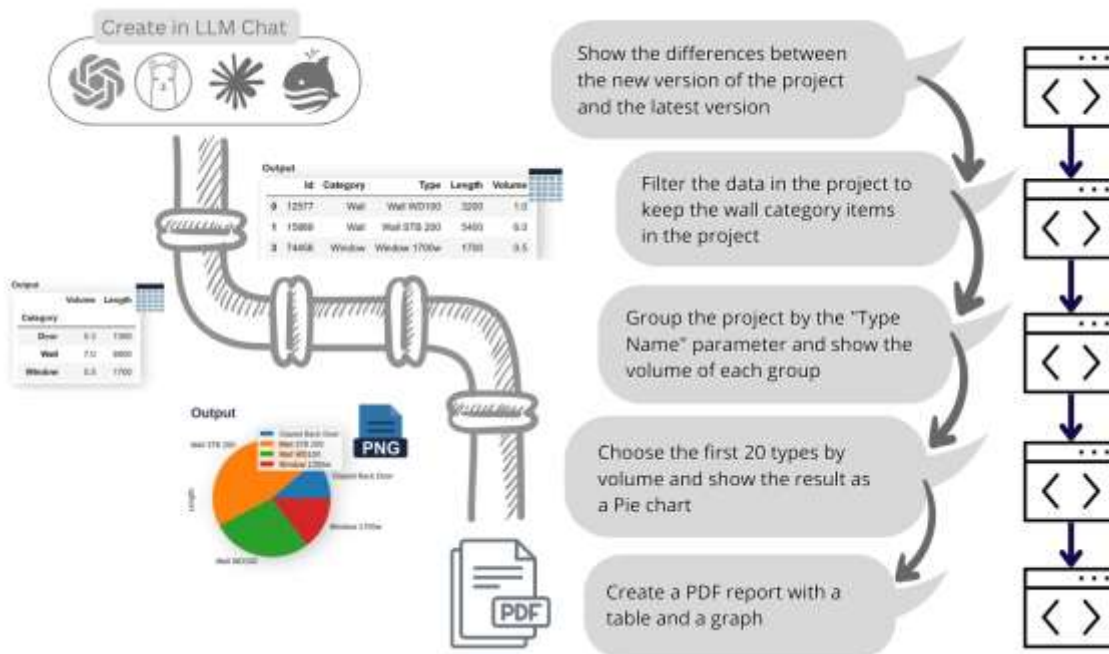
לאחר ביצוע הקוד matplotlib (איור 7.220) באמצעות ספריית Python מייצרת ומבצעת אוטומטית קוד LLM הזה, אנו מקבלים גרפים של עלויות ונפחי השימוש בחומרים בפרויקטים בניה ישירות בשיחה (איור 7.221, מה) --שמשפט באופן משמעותי את העבודה האנליטית



DataFrame-בצורת גרפים על בסיס הנתונים שנאספו ב LLM איור 7.221 ויזואליזציה של תגובת

ניתוח וביצוע קוד, ויזואליזציה של תוצאות זמינה דרך בקשות טקסט פשוטות, ETL תמיכה בפיתוח רעיונות לכתיבת קוד בהחלט משנה את הגישה לתכנות ואוטומציה של LLM ללא צורך בלימוד יסודות התכנות. הופעת כלים כמו, LLM-ב - (עיבוד נתונים) איור 7.222)

מה הערך האמיתי של אינטליגנציה מלאכותית לעסק שלך ואיך אתה יכול להפיק ממנה "PwC לפי דוח תועלת" (2017? [139], אוטומציה של תהליכים ושיפור פרודוקטיביות יהיו המניעים העיקריים לצמיחה כלכלית. צפוי כי שיפור הפרודוקטיביות יספק יותר מ-55% מכלל הגידול בתוצר המקומי הגולמי בזכות אינטליגנציה מלאכותית בתקופה 2017-2030.



מסייעת ביצירת טיוטת קוד, אשר מיועדת לשימוש בפרויקטים עתידיים ללא צורך LLM איור 7.222 אינטליגנציה מלאכותית LLM-בשימוש ב

כמו גם ChatGPT, LLaMa, Mistral, Claude, DeepSeek, QWEN, Grok באמצעות כלים כמו נתונים פתוחים ותוכנה בקוד פתוח, אנו יכולים לאוטומט את התהליכים שבוצעו בעבר רק באמצעות מערכות מודולריות פרופריטריות יקרות ומסובכות לתחזוקה

אוטומטיים לעיבוד נתונים, ייהנו מיתרונות Pipeline בהקשר של בנייה, זה אומר שהחברות שיאמצו ראשונות תהליכי משמעותיים – משיפור היעילות בניהול פרויקטים ועד הפחתת הפסדים כספיים והסרת יישומים מפוצלים ואחסון נתונים מבודד.

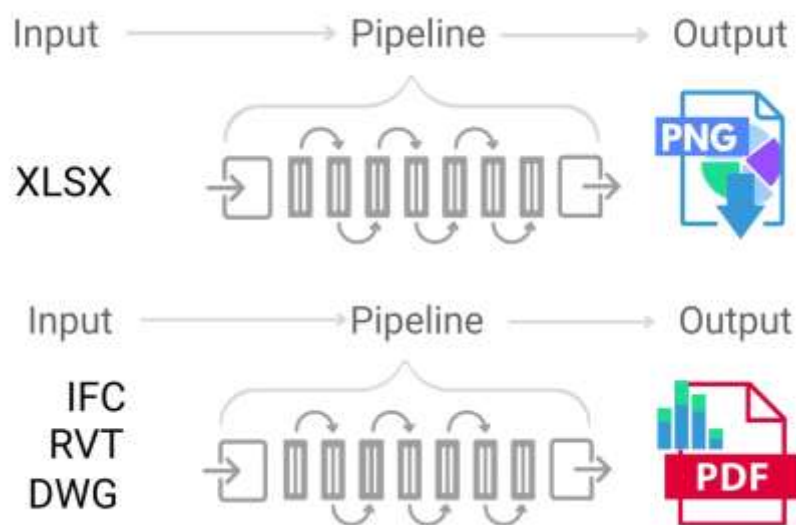
היא חלק חשוב מאוטומציה של תהליכי אנליטיקה ועיבוד ETL הלוגיקה המתוארת לביצוע משימות עסקיות בתהליך (Pipelines) רישם נתונים, שהיא סוג ספציפי של המושג הרחב יותר – קונ

פרק 7.3. אוטומטי ETL צינור

ETL צינור: קונוויור נתונים אוטומטי של

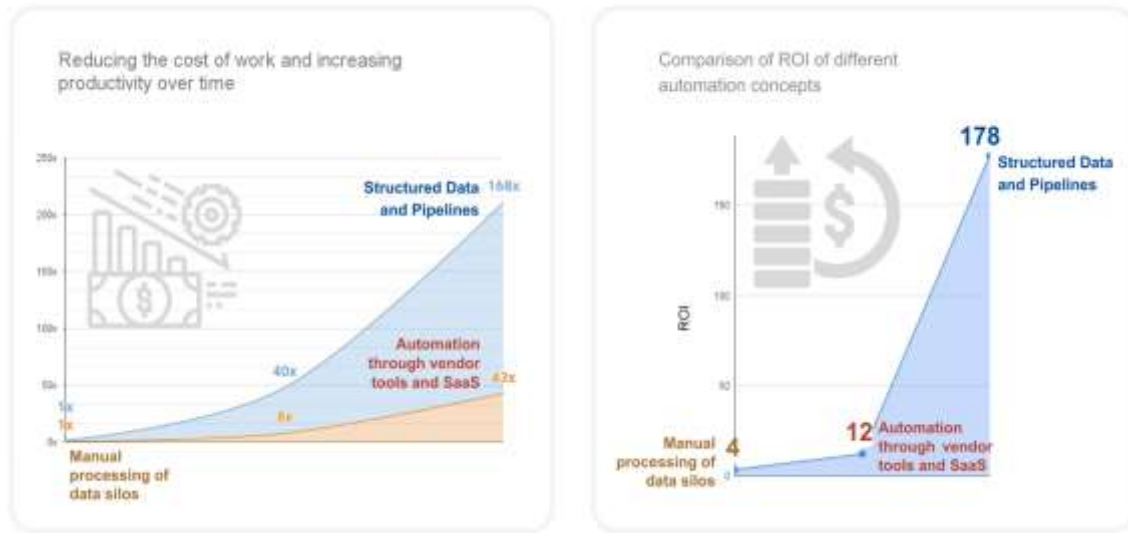
משמש באופן מסורתי לעיבוד נתונים במערכות אנליטיות, מכסה מקורות נתונים מובנים ולא מובנים. עם ETL תהליך נמצא בשימוש תדיר יותר, המתאר – (Рейс) Pipeline – זאת, בסביבה הדיגיטלית המודרנית, המונח הרחב יותר. כל שרשרת עיבוד רציפה, שבה הפלט של שלב אחד הופך לקלט של השלב הבא

גישה זו רלוונטית לא רק לנתונים, אלא גם לסוגים אחרים של אוטומציה: עיבוד משימות, בניית דוחות, אינטגרציה עם -תוכנה ומסמכים דיגיטליים) איור 7.31



7.31 Pipeline איור – Pipeline של השלב הבא –

הוא אחד המרכיבים המרכזיים של אוטומציה, במיוחד בתנאים של עבודה עם כמות גדולה של Pipeline יישום של נתונים מגוונים. הארכיטקטורה של הקו מאפשרת לארגן שלבים מורכבים של עיבוד בפורמט מודולרי, רציף ומנוהל, מה שמגביר את הקריאות, מפשט את תחזוקת הקוד, ומאפשר דיבוג שלב-שלב ובדיקות ניתנות להרחבה.



לבדוק נתונים מקטין בעשרות ובמאות פעמים את זמן הביצוע בהשוואה לעיבוד באמצעות Pipeline של תהליך ROI 7.32 איור כלים קלאסיים.

עיבוד נתונים באמצעות קו מאפשר להגדיל באופן ('וכו ERP, PMIS, CAD) בניגוד לעבודה ידנית במערכות קנייניות --משמעותי את מהירות ביצוע המשימות, להימנע מעבודה חוזרת ולאוטומט את השקת התהליכים בזמן הנדרש

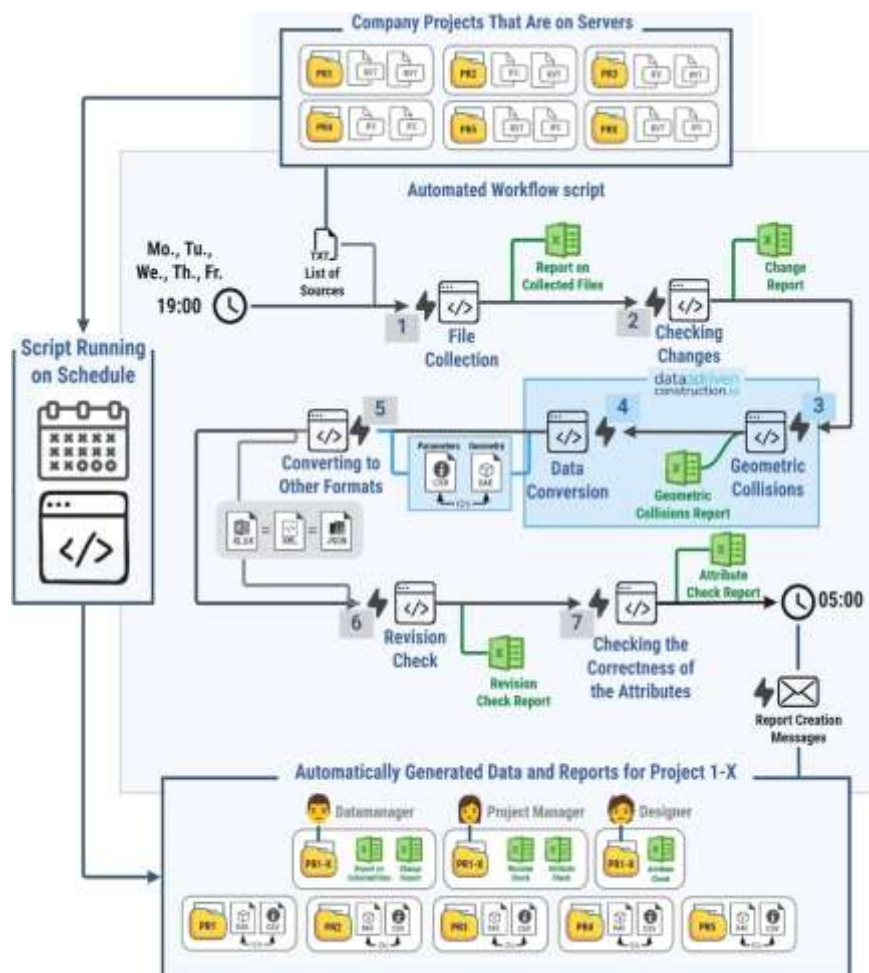


Excel מבלי לפתוח את XLSX לקבלת אוטומטית של גרף מנתונים טבלאיים בקובץ ETL Pipeline-דוגמה ל 7.33 איור

יש לקבוע מראש את מקורות הנתונים, כמו ETL אוטומטי, בדומה לתהליך Pipeline כדי לעבד נתוני זרימה ולבנות גם את מסגרות הזמן לאיסוף הנתונים - הן עבור תהליך עסקי ספציפי, והן במסגרת כל החברה

בפרויקטים בנייה, הנתונים מגיעים ממקורות רבים ושונים עם תדירות עדכון שונה. כדי ליצור תצוגת נתונים אמינה, חשוב מאוד לתעד את רגע ההוצאה והעדכון של המידע. זה מאפשר לקבל החלטות בזמן ולשפר את היעילות בניהול הפרויקט.

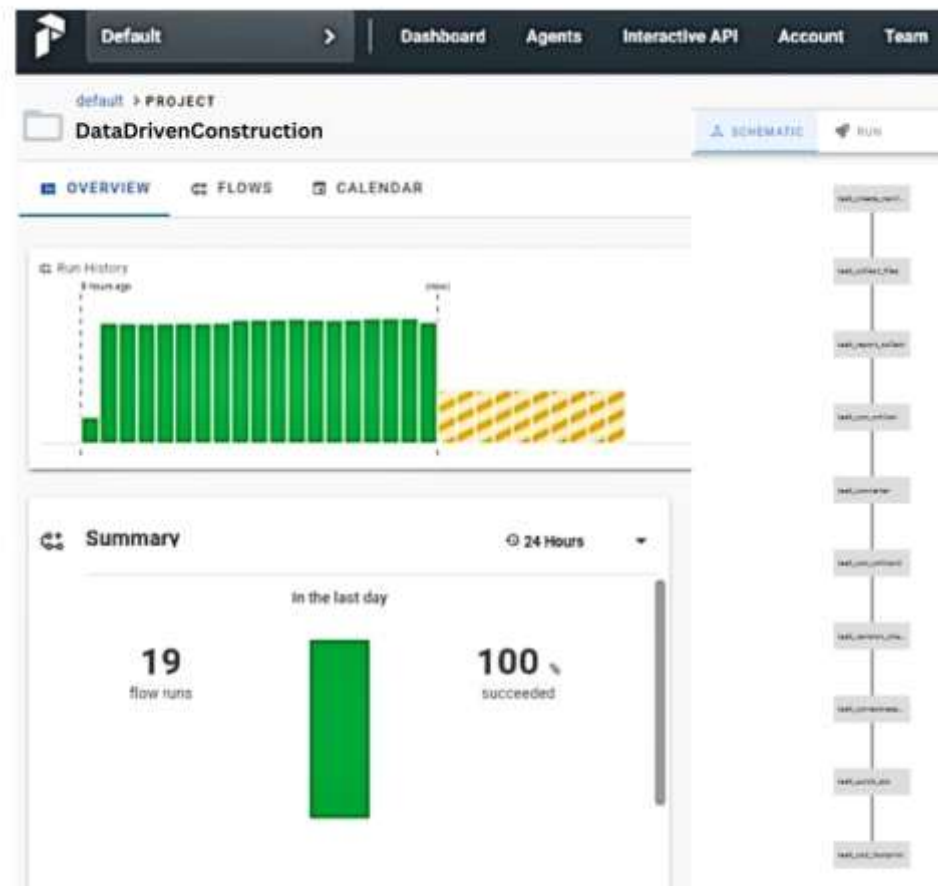
אחת האפשרויות היא להפעיל את תהליך האיסוף בשעה קבועה - למשל, בשעה 19:00, לאחר סיום יום העבודה. בשלב זה מופעל הסקריפט הראשון, שאחראי על אגרגציה של נתונים ממערכות ומאגרי מידע שונים. לאחר מכן מתבצע עיבוד אוטומטי והמרה של הנתונים לפורמט מובנה, המתאים לניתוח. בשלב הסופי, באמצעות הנתונים המוכנים, נוצרות אוטומטית דוחות, לוחות מחוונים ומוצרים אחרים, כפי שמתואר בפרקים הקודמים. כתוצאה מכך, עד השעה 05:00 --- בבוקר, למנהלים כבר יש דוחות עדכניים על מצב הפרויקט בפורמט הנדרש.



שנאספו אוטומטית בערב, מעובדים בלילה, כך שלמנהלים יהיו דוחות עדכניים ודוחות חדשים, Pipeline-נתונים ב 7.34 איור בבוקר.

אוטומציה של תהליכי המרה והמחשה באמצעות לוחות מידע - הם, KPI איסוף נתונים בזמן, קביעת מרכיבים מרכזיים בקבלת החלטות מוצלחות על בסיס נתונים.

תהליכים אוטומטיים כאלה יכולים להתבצע עם אוטונומיה מלאה: הם מופעלים לפי לוח זמנים, מעבדים נתונים ללא כאלה ETL מעורבות של מפעיל ויכולים להיות מיושמים הן בענן והן על השרתים של החברה. זה מאפשר לשלב את קווי הקיימת, תוך שמירה על שליטה על הנתונים והבטחת גמישות בהרחבה IT-בתשתית.



אוטומטי של תהליכים, שבו 10 סקריפטי פייתון מופעלים בזה אחר זה לאחר השעה 19:00 בכל יום ETL קו 7.35 איור -עבודה.

אוטומציה של תהליכי עבודה לא רק מעלה את הפרודוקטיביות של הצוות, משחררת זמן למשימות משמעותיות ופחות בתהליכי (AI) שגרתיות, אלא גם מהווה את הצעד הראשון החשוב בדרך ליישום טכנולוגיות אינטליגנציה מלאכותית "עסקים, על כך נדבר בהרחבה בפרק "תחזיות ולמידת מכונה

LLM באמצעות Pipeline-ETL תהליך בדיקת נתונים של

פירקנו שלב אחר שלב את תהליך הכנת ETL בפרקים הקודמים, שהוקדשו ליצירת דרישות נתונים ואוטומציה של המרה, אימות והצגת הנתונים. פעולות אלו בוצעו בצורה של בלוקים קוד נפרדים (איור 7.218 - איור 7.220) (שכל אחד מהם ביצע משימה ספציפית

כעת המטרה הבאה שלנו היא לשלב את האלמנטים הללו לקו ייצור אחיד, מחובר ואוטומטי של עיבוד נתונים - קו ייצור (שבו כל השלבים) טעינה, בדיקה, הצגה, ייצוא (מתבצעים באופן רציף בסקריפט אחד המופעל - ETL-Pipeline

אוטומטית.

המקורי → עד בדיקת המבנה והערכים באמצעות CSV בדוגמה הבאה ייושם מחזור מלא של עיבוד נתונים: מטעינת קובץ PDF. ביטויים רגולריים → חישוב התוצאות → יצירת דוח חזותי בפורמט

🔗 LLM-כדי לקבל את הקוד המתאים ניתן להשתמש בבקשה טקסטואלית הבאה ל

DataFrame-בודקת את נתוני ה CSV, בבקשה כתוב דוגמת קוד שמבצעת טעינת נתונים מקובץ יעילות אנרגטית עם, 'W-OLD' או 'W-NEW' באמצעות ביטויים רגולריים, בודקת מזהים בפורמט תקופת אחריות ומחזור החלפה עם ערכים מספריים בשנים ובסוף יוצרת דוח, 'G' עד 'A'-אותיות מ עם היסטוגרמה של התוצאות PDF עם ספירת הערכים שעברו ולא עברו את הבדיקה, מייצרת ומוסיפה תיאור טקסטואלי

LLM תשובת: 



7.36 אוטומטיזציה של מחזור מלא של עיבוד נתונים: מטעינה ובדיקה ועד יצירת דוח מסודר בפורמט **Pipeline (ETL)** איור 7.36 PDF.

CSV לאחר העתקת הקוד, יבצע אימות נתונים מקובץ, DIE-או ב LLM הקוד האוטומטי (איור 7.36) (בתוך צ'אט באמצעות הביטויים הרגולריים שנקבעו, ייצור דוח על מספר הרשומות שעברו ולא עברו את הבדיקה, ולאחר מכן ישמור PDF. את תוצאות הבדיקה בפורמט

שבו כל שלב - מטעינת נתונים ועד יצירת דוח - מיושם כמודול נפרד, מבטיח שקיפות, ETL-Pipeline מבנה דומה של עושה את התהליך שקוף וברור Python-יכולת הרחבה וחזרתיות. הצגת הלוגיקה של הבדיקה בצורה של קוד קריא ב לא רק למפתחים, אלא גם למומחים בתחום ניהול הנתונים, איכותם וניתוחם

לאוטומציה של עיבוד נתונים מאפשר לסטנדרטיזציה של התהליכים, מעלה Pipeline השימוש בגישת את חזרתם ומפשט את ההתאמה לפרויקטים חדשים. בזכות זאת מתפתחת מתודולוגיה אחידה לניתוח נתונים, ללא קשר למקור או לסוג המשימה - בין אם מדובר בבדיקת התאמה לסטנדרטים, יצירת דוחות או העברת נתונים למערכות חיצוניות

אוטומציה כזו מפחיתה את השפעת הגורם האנושי, מצמצמת את התלות בפתרונות קנייניים ומעלה את הדיוק והאמינות של התוצאות, מה שהופך אותן לשימושיות הן לניתוח תפעולי ברמת הפרויקטים והן לניתוח אסטרטגי ברמת החברות

CAD (BIM)-בדיקת נתונים ומידע של רכיבי הפרויקט ב: ETL-צינור

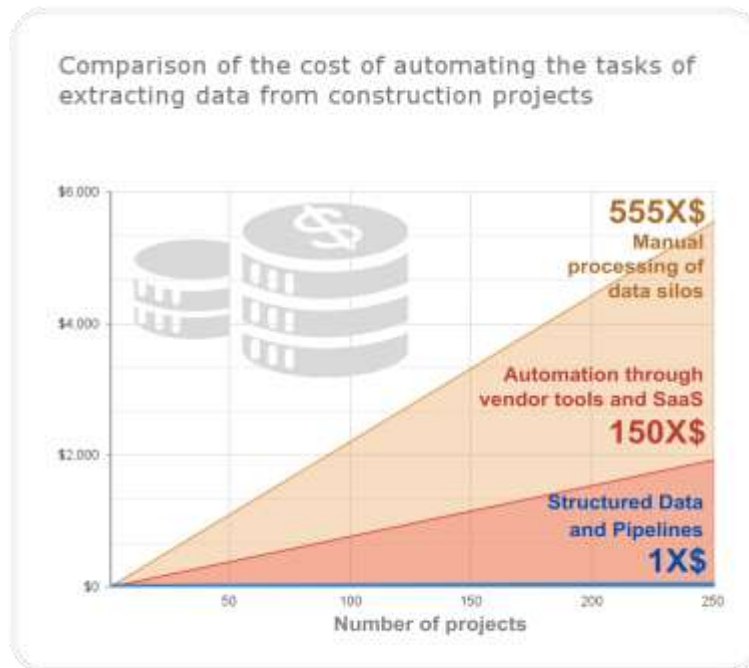
הם אחד ממקורות הנתונים המורכבים והמתעכנים ביותר בעסקי חברות CAD (BIM) נתונים ממערכות ומסדי נתונים הבנייה. יישומים אלה לא רק מתארים את הפרויקט באמצעות גיאומטריה, אלא גם משלימים אותו במידע טקסטואלי רב-שכבתי: כמויות, תכונות חומרים, ייעודי חדרים, רמות יעילות אנרגטית, סובלנות, תקופות שימוש ואטריבוטים נוספים

נוצרים בשלב התכנון והופכים לבסיס לתהליכי עסקים CAD האטריבוטים המוקנים ליישומים במודלים ו-ERP נוספים, כולל חישובי עלויות, הכנת לוחות זמנים, הערכת מחזור חיים ואינטגרציה עם מערכות. כאשר היעילות של התהליכים תלויה במידה רבה באיכות הנתונים המתקבלים ממחלקות התכנון, CAFM,

-. , , (7.21). CAD-(BIM-) הגישה המסורתית לבדוק אטריבוטים במודלים

קבלנים כלליים ומנהלי פרויקטים מתמודדים עם הצורך לעבד כמות גדולה של נתוני פרויקטים, כולל RVT, DWG, גרסאות מרובות וקטעים של אותם מודלים. הנתונים מגיעים מארגוני תכנון בפורמטים ודורשים בדיקה קבועה לעמידה בסטנדרטים תעשייתיים (3.114). ואחרים DGN, IFC, NWD ותאגידיים.

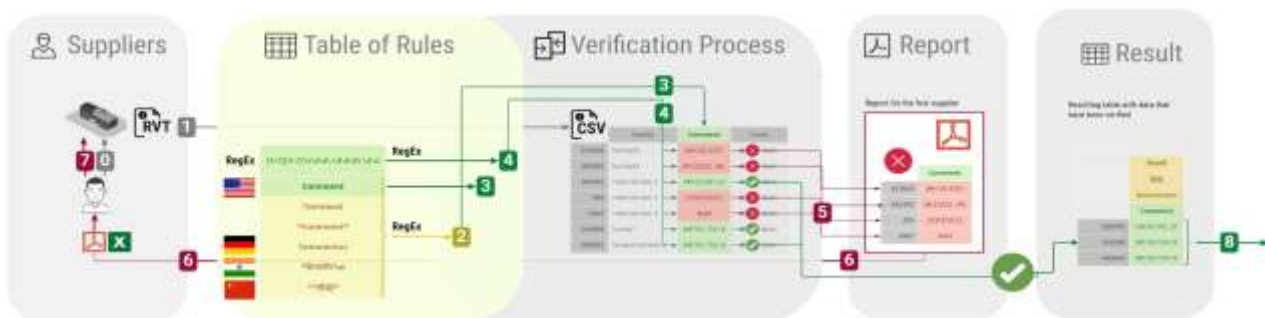
בשל התלות בפעולות ידניות ובתוכנות מיוחדות, תהליך האימות של הנתונים הופך לנקודת צוואר בקבוק בתהליכי העבודה הקשורים לנתונים מהמודלים עבור כל החברה. אוטומציה ויישום דרישות מובנות מאפשרים להפסיק את התלות (7.37). הזו, להגדיל באופן משמעותי את מהירות ואמינות בדיקת הנתונים.



[140] אוטומציה מגדילה את מהירות הבדיקה ועיבוד הנתונים, מה שמפחית את עלות העבודה בעשרות מונים 7.37 ..

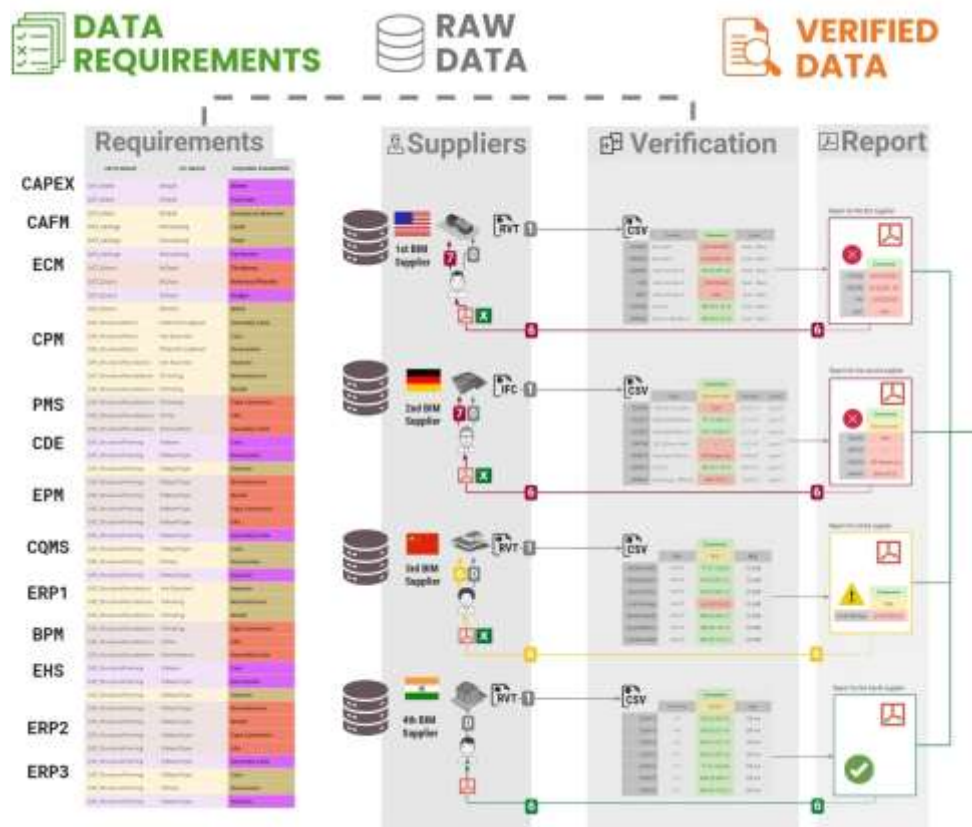
(RVT, DWG, DGN, NWS) ממגוון פורמטים סגורים (Extract של ETL) כולל הוצאת נתונים CAD תהליך בדיקת נתוני שבהם ניתן להחיל על כל (IFC, CPXML, USD) ואחרים (או פורמטים חצי-מובנים ופרמטריים פתוחים RegEx (7.38). באמצעות ביטויים רגולריים (Transform של) אטריבוט וערכיו טבלאות כללים תהליך שדנו בו, בפירוט בחלק הרביעי של הספר -

לפורמטים מובנים (Load של) ורשומות שנבדקו בהצלחה צריכה להסתיים ביצוא PDF יצירת דוח שגיאות בפורמט המתחשבים רק ביישויות שנבדקו, שיכולות לשמש לתהליכים נוספים



תהליך בדיקת הנתונים מספקי נתוני התכנון ועד לדו"ח הסופי, שנבדק באמצעות ביטויים רגולריים 7.38 .

כאשר יש דרישות מובנות ובזרימה מתמדת של נתונים חדשים CAD (BIM) אוטומציה של בדיקת נתונים ממערכות מפחיתה את הצורך בהתערבות ידנית בתהליך האימות (כל אחד מהתהליכים ETL-Pipelines (7.39). של בדיקה והכנת דרישות לנתונים נדונו בפרקים הקודמים -



מפשטת את ניהול פרויקטי הבנייה בזכות האצת התהליכים ETL אוטומציה של בדיקת נתונים דרך 7.39 ..

יכולה להימשך מספר ימים עד שבועות CAD (BIM) באופן מסורתי, בדיקת המודלים המוצעים על ידי קבלנים ומומחי אוטומטיים, ניתן לקצר את הזמן הזה למספר דקות. במצב טיפוסי, הקבלן מצהיר ETL עם זאת, עם יישום תהליכי המודל נבדק ועומד בדרישות. "הצהרה כזו מפעילה שרשרת של בדיקות לגבי טענת הקבלן על איכות הנתונים"

🔊 "מנהל הפרויקט - "הקבלן טוען": המודל נבדק, הכל בסדר

🔊 "מנהל הנתונים - "מעלים את האימות

■ מזהה הפרות בשניות. האוטומציה מונעת מחלוקות Pandas-סקריפט פשוט ב

▪ FireRating IS NULL: פרמטר, OST_StructuralColumns: קטגוריה

▪ Excel/PDF-מייצרים רשימת מזהה הפרות → מייצאים ל

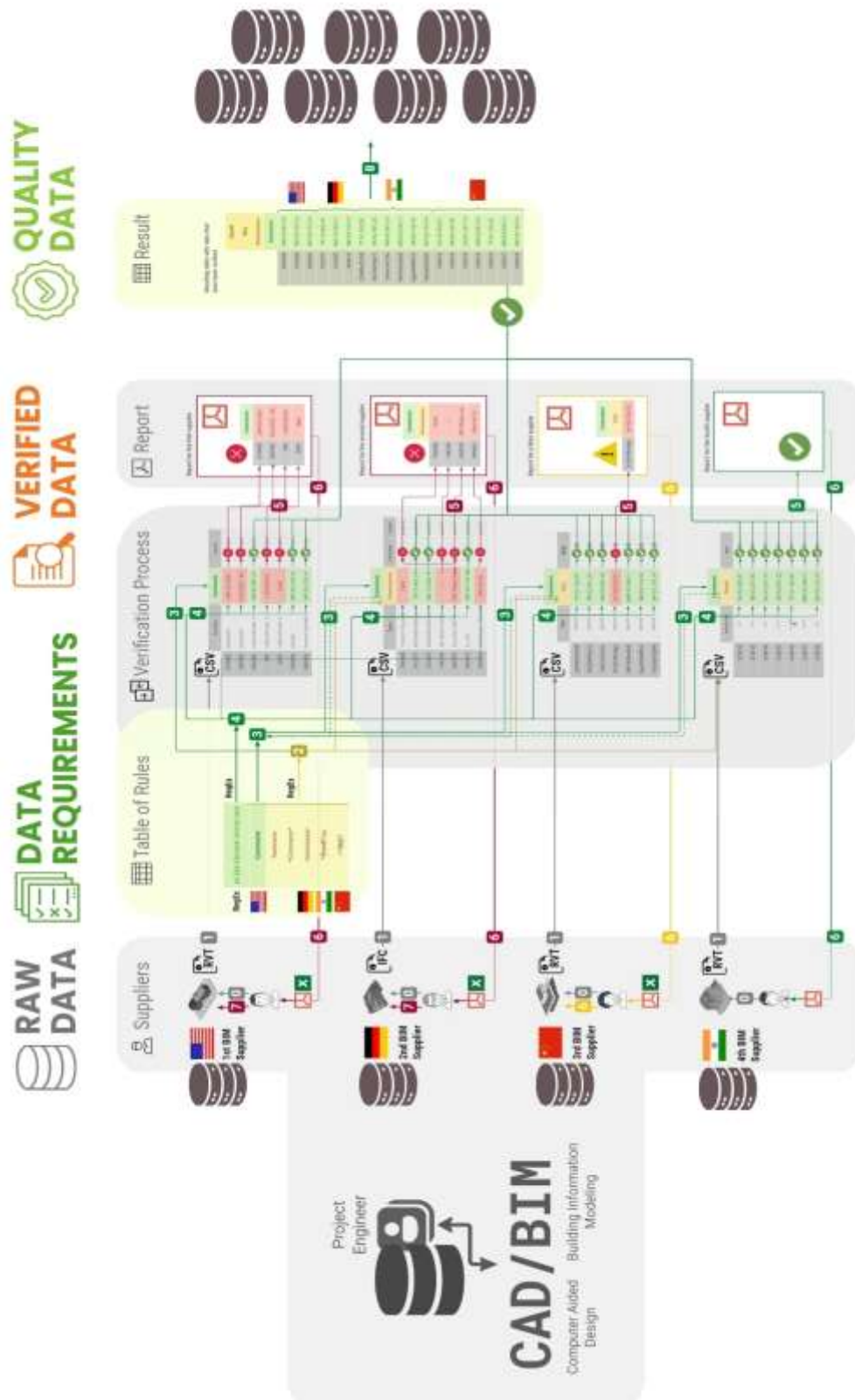
מזהה הפרות בשניות Pandas-סקריפט פשוט ב

```
df = model_data[model_data["Category"] == "OST_StructuralColumns"] # סינון issues
= df[df["FireRating"].isnull()] # ערכים ריקים
issues[["ElementID"]].to_excel("fire_rating_issues.xlsx") # ייצוא מזהי
```

"FireRating מנהל הנתונים למנהל הפרויקט - "הבדיקה הראתה כי ל-18 עמודים חסר פרמטר
הוא חובה, בלעדיו לא ניתן לקבל את FireRating מנהל הפרויקט לקבלן - "המודל מוחזר לתיקון: פרמטר
"המודל".

לא עובר את בדיקת האיכות, האוטומציה מונעת מחלוקות, והקבלן מקבל כמעט מיד דוח CAD-כתוצאה מכך, מודל ה
מובנה עם רשימת מזהי רכיבים בעייתיים. כך תהליך האימות הופך לשקוף, ניתן לשחזור ומוגן מפני גורמים אנושיים
(7.310 איור).

גישה כזו הופכת את תהליך בדיקת הנתונים לפונקציה הנדסית, ולא לבקרת איכות ידנית. זה לא רק מעלה
את הפרודוקטיביות, אלא גם מאפשר להחיל לוגיקה אחידה על כל הפרויקטים של החברה, מה שמבטיח
טרנספורמציה דיגיטלית מקיפה של התהליכים, מהתכנון ועד ההפעלה.



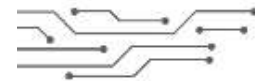
אוטומציה של בדיקת מאפייני רכיבים מונעת את הגורם האנושי ומפחיתה את הסיכון לטעויות 7.310 איור.

משתמשי המערכת, הממתינים לנתונים איכותיים ממערכות, (7.310 איור) בזכות השימוש בקווי ייצור אוטומטיים יכולים לקבל מיד את הפלט הנדרש להם - טבלאות, מסמכים, תמונות - ולשלב אותם במהירות במשימות, CAD- (BIM-), העבודה שלהם.

אוטומציה של בקרה, עיבוד וניתוח מביאה לשינוי בגישות לניהול פרויקטים בנייה, במיוחד בחלק של אינטראקציה בין מערכות שונות, ללא שימוש במערכות מודולריות מורכבות ויקרות או פתרונות סגורים מספקים.

בעוד שגישות ואביזרי שיווק באים והולכים, תהליכי בדיקת דרישות הנתונים יישארו תמיד חלק בלתי נפרד מהתהליכים העסקיים. במקום ליצור כל פעם פורמטים וסטנדרטים מיוחדים חדשים, תעשיית הבנייה צריכה להתמקד בכלים שכבר הוכיחו את יעילותם בתעשיות אחרות. כיום קיימות פלטפורמות חזקות לאוטומציה של עיבוד נתונים ואינטגרציה של Extract, Transform ו-Load. תהליכים, המאפשרות לחברות לצמצם משמעותית את הזמן על פעולות שגרתיות ולמזער טעויות בתהליכי

המאפשר, Apache Airflow הוא ETL אחד מהדוגמאות הפופולריות לפתרונות אוטומציה ואורקסטרציה של תהליכי, נעשה שימוש פעיל גם בפתרונות דומים אחרים, Airflow, לצד ETL. לארגן תהליכים חישוביים מורכבים ולנהל קווי לאוטומציה של תהליכי עסקים n8n למיפוי ועיבוד נתונים בזמן אמת, וכן Apache NiFi כמו



פרק 7.4. ותהליכי עבודה: פתרונות מעשיים ETL אורקסטריציה של

ואפאצ'י איירפלו: אוטומציה ואורקסטריציה של תהליכי עבודה DAG

היא פלטפורמה חנימית עם קוד פתוח, המיועדת לאוטומציה, אורקסטריציה ומעקב אחר תהליכי Apache Airflow (ETL קווי) עבודה.

בכל יום בעבודה עם כמויות גדולות של נתונים נדרשת

■ (למשל, מספקים או לקוחות) Extract - הורדת קבצים ממקורות שונים

■ (לארגן, לנקות ולבדוק) Transform - המרת הנתונים לפורמט הנדרש

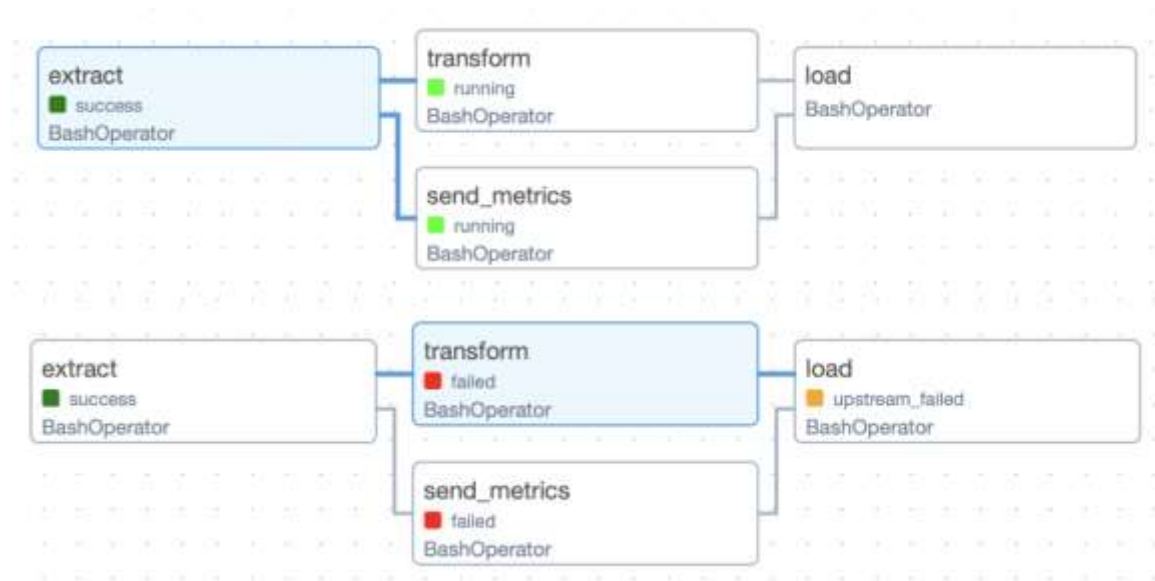
■ (לייצא למערכות, מסמכים, מסדי נתונים או לוחות מחוונים) Load - שליחת התוצאות לבדיקה ויצירת דוחות

כאלה לוקח זמן רב ומוביל לסיכון של טעויות הקשורות לגורם האנושי. שינוי במקור ETL ביצוע ידני של תהליכי הנתונים או כשל באחד מהשלבים עלול לגרום לעיכובים ולתוצאות לא נכונות.

אמין, למזער טעויות, לקצר את זמן עיבוד הנתונים ETL מאפשרים לבנות קו, Apache Airflow כלים לאוטומציה, כמו DAG (Directed Acyclic Graph) הוא המושג Apache Airflow ולהבטיח את נכונותם בכל שלב. הבסיס של שולל מעגלים DAG. שבו כל משימה (קשורה לתלויות אחרות ומבוצעת בסדר שנקבע, cyclic גרף מכוון א מה שמבטיח מבנה לוגי וצפוי לביצוע המשימות

לוקח על עצמו את האורקסטריציה - ניהול התלויות בין המשימות, שליטה על לוח הזמנים לביצוע, מעקב אחר Airflow המצב ותגובה אוטומטית לכשלים. גישה זו ממזערת את ההתערבות הידנית ומבטיחה את אמינות התהליך כולו.

אורקסטרטור משימות - כלי או מערכת המיועדת לניהול ולשליטה על ביצוע משימות בסביבות חישוביות ומידע מורכבות. הוא מקל על תהליך ההפצה, האוטומציה והניהול של ביצוע משימות, מה שמאפשר לשפר את היעילות בעבודה ולייעל את המשאבים



ריש 7.41 Apache Airflow DAG-ETL, מספקת ממשק נוח, שבו ניתן להמחיש את המשימה ועוד.

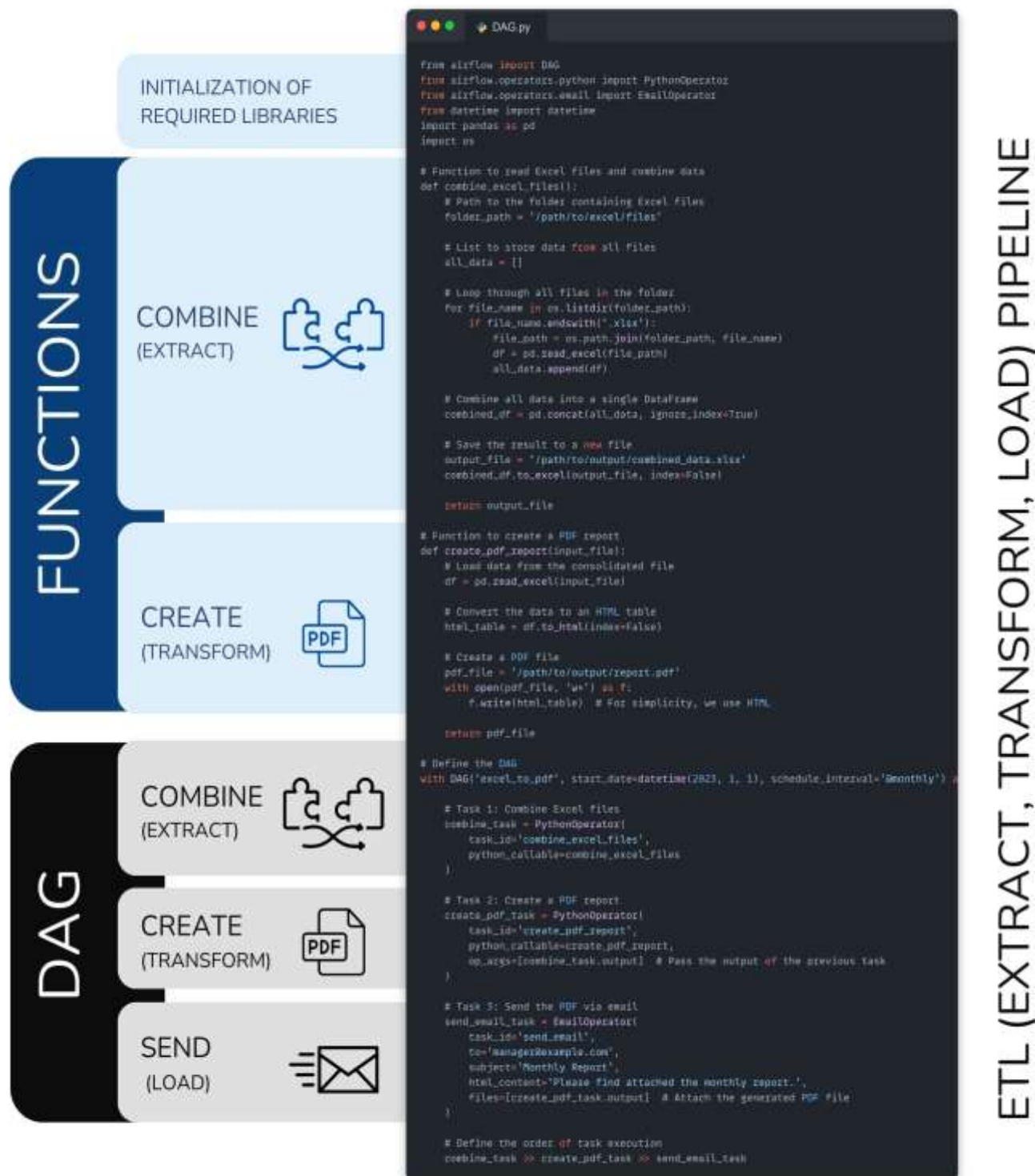
ETL בשימוש נרחב לאורקסטריציה ואוטומציה של חישובים מפורזים, עיבוד נתונים, ניהול תהליכי Airflow (Extract, Transform, Load), תכנון משימות ותסריטים נוספים לעבודה עם נתונים. כברירת מחדל, Apache Airflow משתמשת ב-SQLite.

בגרף, המנוהל דרך Extract, Transform ו-Load - מורכבת משימות ETL-פשוט, בדומה ל-DAG-דוגמה ל לאחר מכן, extract מוגדר סדר ביצוע המשימות (קטעי קוד: (למשל, קודם מתבצע, (ריש 7.41) הממשק המשתמש כאשר כל המשימות הושלמו, תהליך טעינת load (ושליחת מדדים, (וסוגרת את העבודה משימת transform הנתונים נחשב להצלחה.

ETL אפאצ'י איירפלו: יישום מעשי של אוטומציה בתהליכי

גמישים ETL אפאצ'י איירפלו משמש באופן נרחב לארגון תהליכי עיבוד נתונים מורכבים, ומאפשר בניית צינורות אפאצ'י איירפלו ניתן להרצה הן דרך ממשק אינטרנטי והן באופן תכנותי באמצעות קוד פייתון (איור 7.42). (בממשק ים, להריץ משימות ולנתח את תוצאות DAG האינטרנטי) (איור 7.43) (מנהלים ומפתחים יכולים לעקוב באופן חזותי אחרי --הביצוע.

ניתן לקבוע רצף ברור לביצוע משימות, לנהל את התלות ביניהן ולהגיב אוטומטית לשינויים בנתונים DAG באמצעות --המקוריים. נבחן דוגמה לשימוש באיירפלו לאוטומציה של עיבוד דוחות (איור 7.42).



לעיבוד נתונים באמצעות אפאצ'י איירפלו ETL איור 7.42: מושג צינור

ETL: שמבצע משימות מפתח במסגרת צינור DAG (דוגמה זו) איור 7.42 (נבחן

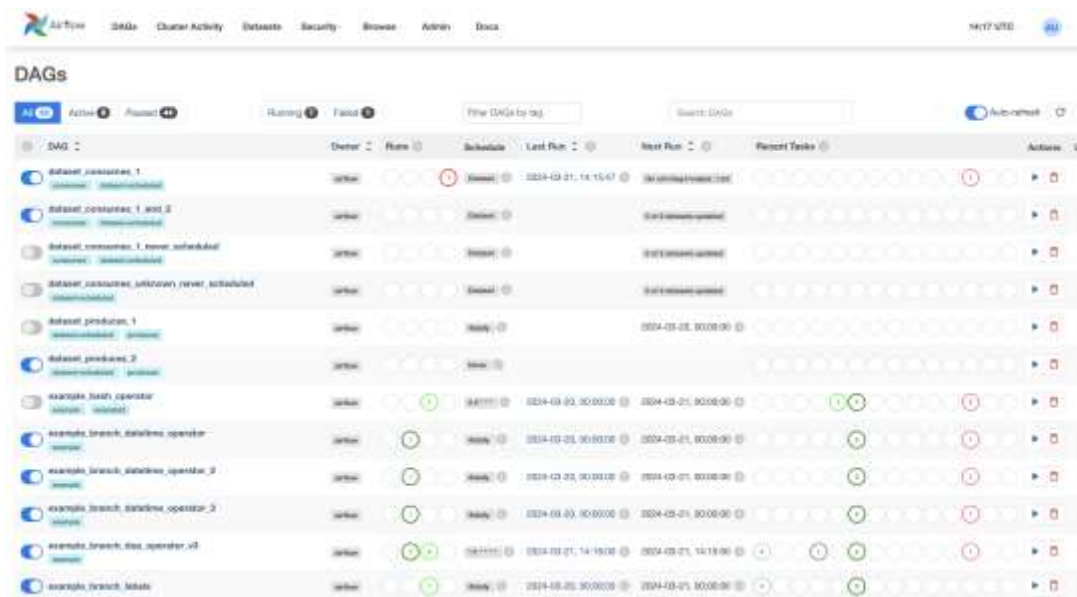
מעבר רציף על כל הקבצים בתיקייה שנבחרה • קריאת נתונים מכל קובץ • Excel (Extract) קריאת קבצי שילוב כל הנתונים במסגרת נתונים אחת • pandas באמצעות ספריית

שמירת הטבלה • HTML-המרת מסגרת הנתונים המאוחדת לטבלה ב • PDF (Transform) יצירת מסמך HTML בגרסה הדמיה - דרך) PDF בפורמט

בדואר PDF לשליחת מסמך EmailOperator-שימוש ב • (Load) שליחת הדו"ח בדואר אלקטרוני אלקטרוני

קביעת רצף ביצוע המשימות: חילוף נתונים → יצירת דו"ח → שליחה • DAG הגדרת (ביום הראשון של כל חודש - @monthly) קביעת לוח זמנים להרצה •

ולשלוח PDF ליצור מסמך, Excel האוטומטית) איור 7.42 (מוצג כיצד לאסוף נתונים מקבצי ETL בדוגמת אותו בדואר אלקטרוני. זו רק אחת מתוך מגוון תרחישים אפשריים לשימוש באיירפלו. דוגמה זו ניתנת להתאמה -לכל משימה ספציפית, כדי לפשט ולאטומט את תהליכי עיבוד הנתונים



DAG	Status	Last Run	Next Run	Percent Tasks
dataset_processor_1	Failed	2024-03-21, 14:15:47	2024-03-21, 14:15:47	0%
dataset_processor_1_and_2	Failed	2024-03-21, 14:15:47	2024-03-21, 14:15:47	0%
dataset_processor_1_and_2_scheduled	Failed	2024-03-21, 14:15:47	2024-03-21, 14:15:47	0%
dataset_processor_1_and_2_scheduled	Failed	2024-03-21, 14:15:47	2024-03-21, 14:15:47	0%
dataset_processor_1	Failed	2024-03-21, 14:15:47	2024-03-21, 14:15:47	0%
dataset_processor_2	Failed	2024-03-21, 14:15:47	2024-03-21, 14:15:47	0%
example_branch_operator	Running	2024-03-21, 14:15:47	2024-03-21, 14:15:47	100%
example_branch_operator_2	Running	2024-03-21, 14:15:47	2024-03-21, 14:15:47	100%
example_branch_operator_3	Running	2024-03-21, 14:15:47	2024-03-21, 14:15:47	100%
example_branch_operator_v2	Running	2024-03-21, 14:15:47	2024-03-21, 14:15:47	100%
example_branch_operator	Running	2024-03-21, 14:15:47	2024-03-21, 14:15:47	100%

בסביבה עם מידע על ההרצות האחרונות DAG-איור 7.43: סקירה של כל קבוצות ה

ממשק האינטרנט של אפאצ'י איירפלו) איור 7.43 (מספק סביבה חזותית מקיפה לניהול תהליכי נתונים. הוא מציג את ים בצורה של גרפים אינטראקטיביים, כאשר הצמתים מייצגים משימות והקצוות - את התלות ביניהן, מה-DAG-ה שמאפשר לעקוב בקלות אחרי תהליכי עיבוד נתונים מורכבים. הממשק כולל לוח בקרה עם מידע על מצב ביצוע המשימות, היסטוריית הרצות, יומני פרטים ומדדי ביצועים. מנהלים יכולים להריץ משימות ידנית, להפעיל מחדש

ים ולהגדיר משתני סביבה - כל זאת דרך ממשק משתמש אינטואיטיבי DAG פעולות שנכשלו, להשהות.

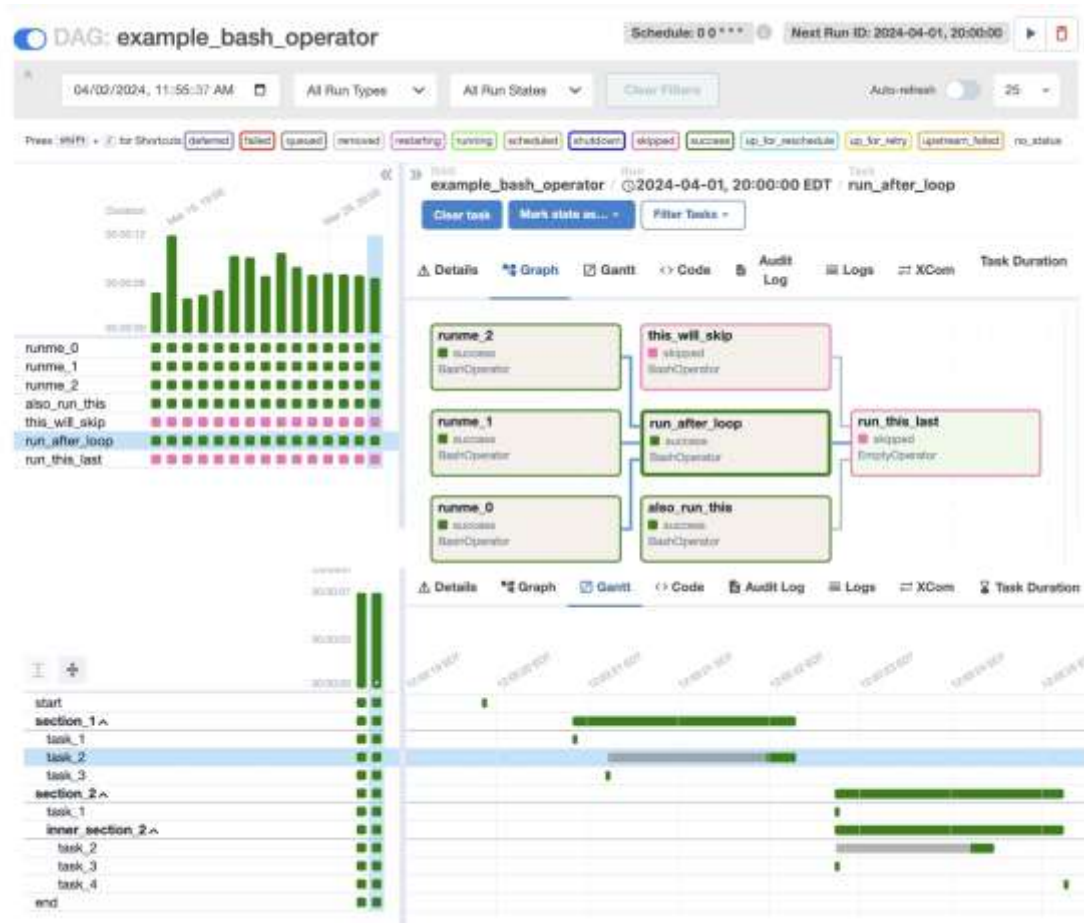
חיצוניים API ארכיטקטורה דומה יכולה להיות משודרגת עם אימות נתונים, התראות על מצב הביצוע, אינטגרציה עם הוספת משימות חדשות, שינוי הסדר שלהן, שילוב DAG-או מסדי נתונים. איירפלו מאפשר גמישות בהתאמת ה בממשק DAG שרשראות - מה שהופך אותו לכלי יעיל לאוטומציה של תהליכי עיבוד נתונים מורכבים. כאשר מפעילים האינטרנט של איירפלו) איור 7.43, איור 7.44 (ניתן לעקוב אחרי מצב ביצוע המשימות. המערכת משתמשת בקוד צבע:

■ ירוק - המשימה הושלמה בהצלחה.

■ צהוב - התהליך מתבצע.

■ אדום - שגיאה בביצוע המשימה.

(במקרה של תקלות) למשל, קובץ חסר או מבנה נתונים פגום (המערכת מפעילה אוטומטית שליחת התראה



מפשט באופן משמעותי את האבחון של בעיות, אופטימיזציה של תהליכים ועבודה משותפת של **Apache Airflow 7.44** איור צוותים על צינורות מורכבים לעיבוד נתונים.

נוח בכך שהוא אוטומט את המשימות השגרתיות, מה שמפחית את הצורך לבצע אותן Apache Airflow ידנית. הוא מספק אמינות באמצעות ניטור ביצוע תהליכים והתראה מיידית על שגיאות. הגמישות של המערכת מאפשרת להוסיף בקלות משימות חדשות או לשנות את הקיימות, תוך התאמת זרימות העבודה לדרישות המשתנות.

הפתוח Prefect, קיימים כלים דומים לאורקסטריציה של זרימות עבודה. לדוגמה, Apache Airflow-בנוסף ל Spotify, שפותח על ידי Python, Luigi, מציע תחביר פשוט יותר ומשתלב טוב יותר עם (7.35 איור) והחינמי המציעים גישות, Kronos ו-Dagster מספק פונקציונליות דומה ועובד היטב עם נתונים גדולים. כמו כן, יש לציין את מודרניות ליצירת צינורות עם דגש על מודולריות וסקלריות. הבחירה בכלי אורקסטריציה תלויה בצרכים. -מורכבים לעיבוד נתונים ETL הספציפיים של הפרויקט, אך כולם מסייעים לאוטומט את תהליכי ראויה להזכיר בנפרד – פלטפורמה עם קוד פתוח, המיועדת לעיבוד ומסלול נתונים בזמן Apache NiFi פלטפורמת מתמקדת בזמן אמת, בעיבוד נתונים בזמן NiFi, המתמקד בעיבוד אצווה ובניהול תלותים, Airflow-אמת. בניגוד ל אמת ובמסלול גמיש בין מערכות

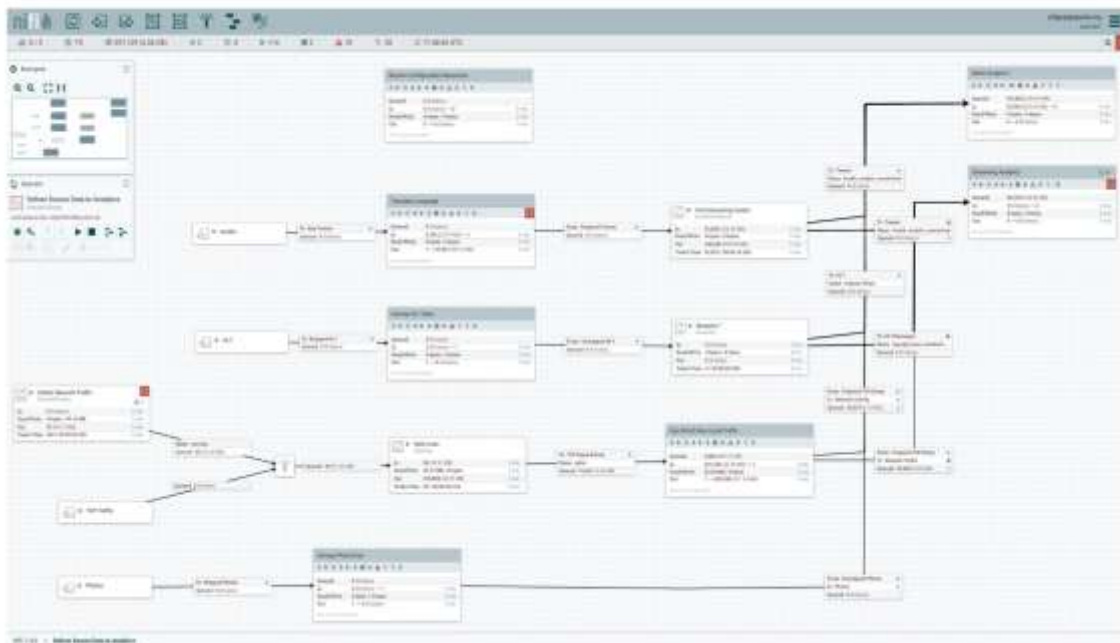
אפאצ'י ניפי למיפוי והמרת נתונים

היא פלטפורמה חזקה עם קוד פתוח, המיועדת לאוטומט את זרימת הנתונים בין מערכות שונות. היא Apache NiFi "Niagara Files" בשם (NSA) פותחה במקור בשנת 2006 על ידי סוכנות הביטחון הלאומי של ארצות הברית והפך לחלק מהיוזמות שלהם, Apache Software, לצרכים פנימיים. בשנת 2014 הפרויקט נפתח והועבר לקרן להעברת טכנולוגיות

העובד עם משימות אצווה ודורש, Airflow-מתוכנן לאיסוף, עיבוד והעברת נתונים בזמן אמת. בניגוד ל Apache NiFi פועל במצב עיבוד רציף, המאפשר העברת נתונים ללא הפסקה בין שירותים שונים NiFi, לוחות זמנים ברורים

חיישנים של אתרי בנייה, מערכות ניטור, IoT מתאים במיוחד לאינטגרציה עם מכשירי Apache NiFi בזרימה על השרת, שם עשויה להיות דרישה לתגובה מיידית לשינויים CAD ולמשל, לבדוק פורמטים. בנתונים.

מאפשר לסטנדרטיזציה של נתונים) שלב העיבוד (לפני העברתם NiFi, בזכות הכלים המובנים לסינון, עיבוד ומסלול (טעינה) (לאחסון או למערכות אנליטיות. אחד היתרונות המרכזיים שלו הוא התמיכה המובנית באבטחה ובקרת גישה, מה שהופך אותו לפתרון אמין לעיבוד מידע רגיש



Apache NiFi ייצוג גרפי של זרימת הנתונים בממשק **7.45** איור

סוגר ביעילות את המשימות של העברת נתונים, סינון ומסלול בזמן אמת. הוא מתאים במיוחד לתרחישים **Apache NiFi** טכנולוגיים מורכבים, שבהם חשוב להעביר מידע בין מערכות עם קיבולת גבוהה

עם זאת, במקרים שבהם המטרה העיקרית היא אינטגרציה של שירותים מגוונים, אוטומציה של פעולות שגרתיות והגדרה מהירה של תהליכי עבודה ללא ידע מעמיק בתכנות, יש צורך בפתרונות עם סף כניסה נמוך וגמישות מקסימלית. אחד המוקדות באוטומציה עסקית ואורקסטריציה ויזואלית **Low-Code/No-Code** פלטפורמת **n8n** הכלים הללו הוא של תהליכים

אורקסטראציה של תהליכים ללא קוד, עם קוד נמוך **n8n**

לבניית תהליכי עבודה אוטומטיים, המובחנת **Open Source Low-Code/No-Code** היא פלטפורמת **n8n** בפשטות השימוש שלה, גמישותה ואפשרות האינטגרציה המהירה עם מגוון רחב של שירותים חיצוניים

הוא שיטה ליצירת מוצרים דיגיטליים ללא כתיבת קוד. כל האלמנטים בתהליך - מהלוגיקה **No-Code** מיועדות למשתמשים **No-Code** ועד הממשק - מתבצעים אך ורק באמצעות כלים ויזואליים. פלטפורמות ללא הכשרה טכנית ומאפשרות ליצור במהירות אוטומציות, טפסים, אינטגרציות ואפליקציות אינטרנט דרך ממשק **Google Sheets** דוגמה: משתמש מגדיר שליחה אוטומטית של התראות או אינטגרציה עם גרירה ושחרור ללא ידע בתכנות

מעניקה לחברות **ETL Pipelines** בתהליכי אוטומציה ובניית **n8n**, בזכות הקוד הפתוח ואפשרות ההתקנה המקומית שליטה מלאה על הנתונים שלהן, מבטיחה אבטחה ועצמאות מספקי שירותי ענן

n8n-Python הממוקד במשימות חישוביות עם אורקסטרציה נוקשה ודורש ידע ב-Apache Airflow-בניגוד ל אף על פי שהממשק שלה (7.46 ראה) מספקת עורך ויזואלי המאפשר ליצור תסריטים ללא צורך בידע בשפות תכנות בתסריטים מורכבים יותר משתמשים יכולים להוסיף, (No-Code) מאפשר ליצור תהליכים אוטומטיים ללא כתיבת קוד (Low-Code) משלהם להרחבת האפשרויות Python ו-JavaScript פונקציות.

הוא גישה לפיתוח תוכנה, שבה הלוגיקה העיקרית של האפליקציה או התהליך נוצרת Low-Code באמצעות ממשק גרפי ואלמנטים ויזואליים, והקוד התכנותי משמש רק להתאמה או הרחבת הפונקציות מאפשרות להאיץ משמעותית את פיתוח הפתרונות, תוך מעורבות לא רק של Low-Code פלטפורמות מתכנתים אלא גם של משתמשים עסקיים עם מיומנויות טכניות בסיסיות. דוגמה: משתמש יכול להרכיב Python או JavaScript-תהליך עסקי מחלקים מוכנים, ואם יש צורך, להוסיף סקריפט משלו ב

ממוקמת כפלטפורמה עם סף כניסה נמוך, ליצירת תסריטים מורכבים מועילים ידע בסיסי בתכנות n8n-אף על פי ש הגמישות של המערכת מאפשרת להתאים אותה למגוון רחב. API הבנה של טכנולוגיות אינטרנט ומיומנויות עבודה עם ושירותי ענן IoT של משימות - מאוטומציה של עיבוד נתונים ועד אינטגרציה עם מסג'רים, מכשירי

n8n-תכנות עיקריות ויתרונות השימוש ב

■ קוד פתוח ואפשרות התקנה מקומית מספקים שליטה מלאה על הנתונים, עמידה בדרישות אבטחה ועצמאות מספקי שירותי ענן.

■ מסחר אלקטרוני, פלטפורמות ענן, מסג'רים, CRM, ERP אינטגרציה עם יותר מ-330 שירותים, כולל ומסדי נתונים.

■ לוגיקת קבלת API, גמישות התסריטים: מהודעות פשוטות ועד שרשראות מורכבות עם עיבוד בקשות AI. החלטות וחיבור לשירותי

■ במקרה הצורך, משתמשים יכולים לשלב קוד מותאם אישית, מה: Python ו-JavaScript-תמיכה ב שמרחיב את אפשרויות האוטומציה

■ ממשק חזותי אינטואיטיבי: מאפשר להגדיר ולחזות במהירות את כל שלבי התהליך

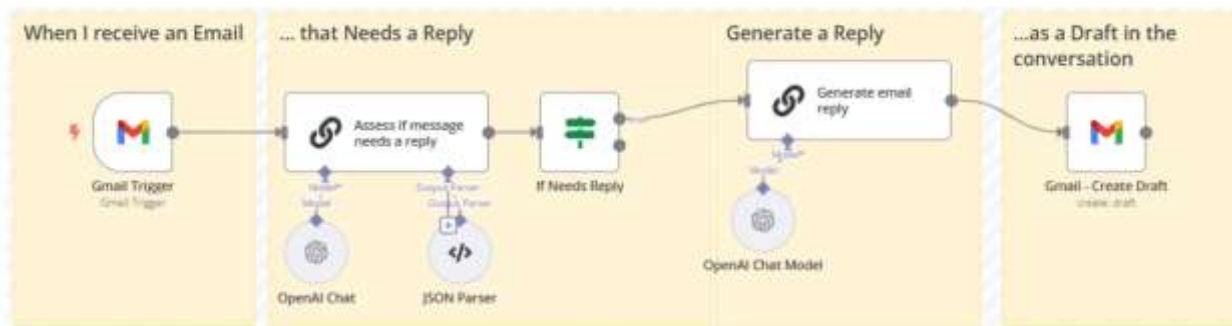
מספקות כלים ליצירת פתרונות דיגיטליים עם מינימום קוד, מה שהופך אותן לאידיאליות Low-Code פלטפורמות מסוג עבור צוותים שאין להם מומחיות טכנית מעמיקה, אך זקוקים לאוטומציה של תהליכים

יכול לשמש לאוטומציה של תהליכים שונים, כגון אינטגרציה עם מערכות ניהול פרויקטים, בדיקות n8n, בתחום הבנייה זרימה, כתיבת דוחות ומכתבים מוכנים, עדכון אוטומטי של נתוני מלאי חומרים, שליחת הודעות לצוותים על מצב מאפשר להפחית באופן משמעותי את מספר הפעולות הידניות, להפחית את n8n-מותאם ב Pipeline. המשימות ועוד הסיכון לטעויות ולהאיץ את קבלת ההחלטות לביצוע פרויקטים

הזמינים באתר n8n, מוכנים, חינוניים ופתוחים של Pipeline ניתן לבחור אחד מתוך כמעט אלפיים כדי לאוטומט את תהליכי העבודה בתחום הבנייה כמו גם משימות אישיות, ובכך, n8n.io/workflows, להפחית פעולות שגרתיות

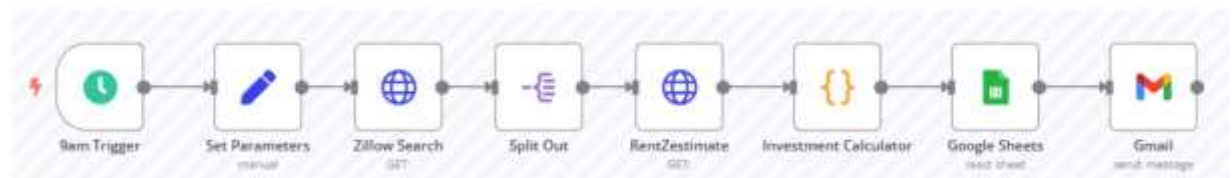
אשר יוצרות אוטומטית טיוטות תגובות, n8n.io המוכנות, הזמינות בחינם באתר Pipeline-ניקה אחד מתוך תבניות ה- ועוזרות למשתמשים שמקבלים כמות גדולה של מיילים או מתקשים בניסוח תגובות, Gmail-ב.

צור טיוטות תגובות להודעות נכנסות "מנתחת את ההודעות הנכנסות בעזרת: Gmail-זו "מגיב אוטומטי ב n8n תבנית HTML-וממירה את הטקסט ל ChatGPT קובעת את הצורך בתגובה, יוצרת טיוטה עם, ChatGPT של LLM במקביל, המייל לא נשלח אוטומטית, מה שמאפשר לערוך ולאשר את Gmail-ומוסיפה אותו לשרשרת ההודעות ב OpenAI ואינטגרציה עם Gmail API ל-OAuth התגובה ידנית. ההגדרה לוקחת כ-10 דקות וכוללת קונפיגורציית בסופו של דבר, מתקבל פתרון נוח וחינמי לאוטומציה של תקשורת דוא"ל שגרתית מבלי לאבד שליטה על תוכן API. המיילים.



תהליך אוטומטי ליצירת תגובות על מיילים בעזרת n8n. ריש.

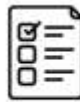
אוטומציה של" n8n של Pipeline. חיפוש עסקאות משתלמות בשוק הנדל"ן – n8n דוגמה נוספת לאוטומציה עם אוסף מדי יום הצעות עדכניות העומדות, "Zillow API, Google Sheets ו-Gmail", עסקאות יומיות בנדל"ן בעזרת (Cash on Cash) הוא מחשב אוטומטית מדדים השקעתיים מרכזיים. Zillow API בקריטריונים שנקבעו, באמצעות ושולח דוח סיכום בדוא"ל, מה Google Sheets מעדכן את (ROI, Monthly Cash Flow, Down Payment), שמאפשר למשקיעים לחסוך זמן ולהגיב במהירות על הצעות הטובות ביותר.



תהליך אוטומטי להערכת האטרקטיביות ההשקעתית של נדל"ן. ריש.

הופכת לכלי יקר ערך עבור חברות השואפות לדיגיטליזציה ולהגברת התחרותיות n8n, בזכות הגמישות וההרחבה שלה בשוק בעזרת כלים יחסית פשוטים וחינמיים עם קוד פתוח.

NiFi יכולים להתפרש כשלושה רמות של עיבוד נתונים) איור 7.48 (n8n ו-Airflow, Apache NiFi כלים כמו מתאם את ביצוע המשימות, מאחד את הנתונים בצינורות Airflow, מנהל את זרימת הנתונים, מבטיח את הגעתם והמרתם - אוטומט את האינטגרציה עם שירותים חיצוניים ומנהל את הלוגיקה העסקית n8n עיבוד, בעוד



	The main task	Approach
Apache NiFi	Streaming and data transformation	Real-time stream processing
Apache Airflow	Task orchestration, ETL pipelines	Batch planning, DAG processes
n8n	Integration, automation of business logic	Low-code visual orchestration

יכולים להתפרש כשלושה רמות משלימות בארכיטקטורת ניהול n8n ו-Apache Airflow, Apache NiFi איור 7.48
נתונים מודרנית.

יחד, הכלים החינמיים והפתוחים הללו יכולים ליצור דוגמה לאקוסיסטמה יעילה לניהול נתונים ותהליכים בתעשיית
הבנייה, המאפשרת לחברות לנצל את המידע בצורה יעילה לקבלת החלטות ואוטומציה של תהליכים.

צעדים נוספים: המעבר מפעולות ידניות לפתרונות מבוססי אנליטיקה

חברות הבנייה המודרניות פועלות בתנאים של חוסר ודאות גבוה: שינויי מחירים בחומרים, עיכובי אספקה, חוסר בכוח
מסייע לחברות לזהות BI ומערכות ETL אדם ולוחות זמנים קפדניים. השימוש בלוחות מחוונים אנליטיים, צינורות
במהירות בעיות, להעריך את היעילות של המשאבים ולחזות שינויים עוד לפני שהם גורמים להפסדים כספיים
לסיכום חלק זה, כדאי להדגיש את הצעדים המעשיים העיקריים שיכולים לסייע ביישום הטכנולוגיות הנדונות במשימות
היומיומיות שלכם:

■ יישמו ויזואליזציה של נתונים ולוחות מחוונים אנליטיים

☐ (KPI) למדו את תהליך יצירת לוחות מחוונים למעקב אחר מדדי ביצוע מרכזיים

☐ (Power BI, Tableau, Matplotlib, Plotly) השתמשו בכלי ויזואליזציה לנתונים שלכם

■ ETL אוטומטו את עיבוד הנתונים באמצעות תהליכי

☐ ETL באמצעות תהליכי (CAD, מסמכים, גיליונות) הגדרו איסוף אוטומטי של נתונים ממקורות שונים

☐ ארגנו את המרת הנתונים למשל, בדיקה באמצעות ביטויים רגולריים או חישוב (באמצעות סקריפטים
Python ב-

☐ תוך, FPDF באמצעות ספריית (DOC או) PDF נסו להגדיר יצירה אוטומטית של דוחות בפורמט
אחרים PDF או חילוף מידע ממסמכי Excel שימוש בנתונים מקבצי

■ לאוטומציה (LLM) השתמשו במודלים לשוניים

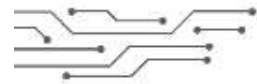
☐ ליצירת קוד שיסייע בחילוף וניתוח נתונים ממסמכים לא (LLM) השתמשו במודלים לשוניים גדולים

מובנים

□ ולמדו תבניות ומקרים מוכנים באתר שלהם. קבעו אילו תהליכים n8n הכירו את כלי האוטומציה No-Code/Low-Code בעבודתכם ניתן לאוטומט לחלוטין באמצעות גישה

הגישה האנליטית לנתונים ואוטומציה של תהליכים לא רק מקצרת את הזמן על פעולות שגרתיות, אלא גם משפרת את מקבלות את האפשרות להגיב ETL איכות ההחלטות המתקבלות. חברות המיישמות כלים לניתוח חזותי וצינורות במהירות לשינויים

היא רק הצעד הראשון לקראת בשלות n8n, Airflow ו-NiFi אוטומציה של תהליכי עסקים באמצעות כלים כמו דיגיטלית. השלב הבא הוא אחסון איכותי וניהול של הנתונים עצמם, אשר מהווים את הבסיס לאוטומציה. בחלק השמיני נבחן כיצד חברות בניה יכולות לבנות ארכיטקטורת אחסון נתונים יציבה, תוך מעבר מהכאוס של מסמכים וקבצים בפורמטים שונים לאחסונים מרכזיים ופלטפורמות אנליטיות





VIII חלק

אחסון וניהול נתונים בבנייה

החלק השמיני בוחן טכנולוגיות מודרניות לאחסון וניהול נתונים בתעשיית הבנייה. כאן ניתוחים CSV ו-XLSX פורמטים יעילים לעבודה עם כמויות גדולות של מידע - מפורמטים פשוטים כמו עם השוואה מפורטת של ORC ו-Apache Parquet ועד פורמטים יעילים יותר כמו (Data Lakes) אגני נתונים, (DWH) יכולותיהם ומגבלותיהם. נדונות מושגי מחסני נתונים (Data Governance) כמו גם עקרונות ניהול הנתונים, (Data Lakehouse) ופתרונות היברידיים שלהם נדונות בעיות "ביצת" (Data Minimalism) ומינימליזם בעבודה עם מידע (Data Governance) ואסטרטגיות למניעת כאוס במערכות מידע. מוצגים גישות חדשות (Data Swamp) "הנתונים Bounding-Box לעבודה עם נתונים, כולל בסיסי נתונים וקטוריים ויישומם בבנייה דרך מושג ה-DataOps ו-VectorOps בחלק זה נדונות גם מתודולוגיות. תהליכי עבודה עם נתונים

פרק 8.1. תשתית נתונים: מפורמטי אחסון ועד אחסונים דיגיטליים

אטומי נתונים: יסוד לניהול מידע יעיל

כל מה שיש ביקום מורכב מיחידות הבניין הקטנות ביותר - אטומים ומולקולות, ובסופו של דבר כל מה שחי ודומם חוזר בהכרח למצב זה. בתהליך זה מתרחש בטבע במהירות מדהימה, שאנו מנסים להעביר לתהליכים המנוהלים על ידי בני אדם.

ביער, כל האורגניזמים החיים הופכים עם הזמן לחומר מזין, שמשרת כבסיס לצמחים חדשים. צמחים אלה בתורם, הופכים למזון ליצורים חיים חדשים, המורכבים מאותם אטומים שיצרו את היקום לפני מיליוני שנים.

בעולם העסקים, חשוב גם לפרק מבנים מורכבים ורכ-שכבתיים ליחידות הבסיסיות ביותר, המינימליות לעיבוד - בדומה לאטומים ומולקולות בטבע. זה מאפשר לאחסן ולנהל את אטומי הנתונים ביעילות, והופך אותם לבסיס עשיר ופורה שמספק משאב מרכזי לצמיחה אנליטית ואיכות קבלת החלטות.



ניתוח וקבלת החלטות מתבססים על נתונים שניתן לעשות בהם שימוש חוזר, אשר פעם עובדו ונשמרו 8.11. ריס.

יצירות מוזיקליות מורכבות מהתווים, אשר מחוברים יחד ויוצרים יצירות מוזיקליות מורכבות, בעוד שמילים נוצרות מיחידה פרימיטיבית - האות-צליל. בין אם מדובר בטבע, מדע, כלכלה, אמנות או טכנולוגיה, העולם מציג אחדות והרמוניה מדהימות במאמציו להריסה, מבנה, מחזוריות ויצירה. באותו אופן, תהליכים במערכות חישוב עלויות מתפרקים ליחידות מובנות קטנות - פריטי משאבים - ברמת החישובים והלוחות הזמנים. לאחר מכן, יחידות אלו, כמו תווים, משמשות ליצירת חישובים ולוחות זמנים מורכבים יותר. באותו עיקרון פועלות מערכות תכנון ממוחשבות, שבהן פרויקטים ארכיטקטוניים

והנדסיים מורכבים נבנים ממרכיבים בסיסיים - רכיבים נפרדים ורכיבי ספרייה, מהם נוצרת מודל תלת-ממדי מלא של פרויקט בניין או מבנה מורכב.

המושג של מחזוריות ומבניות, המאפיין את הטבע והמדע, משתקף גם בעולם הנתונים המודרני. כמו בטבע כל היצורים החיים חוזרים לאטומים ולמולקולות, כך גם בעולם הכלים המודרניים לעיבוד נתונים, המידע שואף לעבור לצורתו הפרימיטיבית ביותר.

המרכיבים הקטנים ביותר עם אי-חלוקתם הם אבני הבניין הבסיסיות של תהליכי העסקים. חשוב מההתחלה לתכנן בקפידה כיצד לאסוף, למבנה (לפרק לאטומים) ולאחסן את אבני הבניין הקטנות הללו ממקורות שונים. בעת זאת, ארגון ואחסון הנתונים אינם רק עניין של פירוקם לרכיבים. לא פחות חשוב להבטיח את האינטגרציה והאחסון המובנה שלהם, כך שניתן יהיה בקלות לשלוף, לנתח ולהשתמש בנתונים בכל עת שיידרשו לצורך קבלת החלטות.

לצורך עיבוד מידע יעיל, יש לבחור בקפידה את הפורמט ושיטות האחסון של הנתונים - בדומה לכך שהאדמה צריכה להיות מוכנה לצמיחת עצים. מחסני הנתונים צריכים להיות מאורגנים כך שיבטיחו איכות גבוהה ורלוונטיות של המידע תוך כדי הוצאת נתונים מיותרים או לא רלוונטיים. ככל שה"אדמה המידע" הזו תהיה מובנית יותר, כך המשתמשים יוכלו למצוא את הנתונים הנדרשים מהר יותר ובדיוק רב יותר ולפתור בעיות אנליטיות.

מאגר מידע: קבצים או נתונים

מחסני הנתונים מאפשרים לחברות לאסוף ולשלב מידע ממערכות שונות, וליצור מרכז אחד לניתוח עתידי. הנתונים ההיסטוריים שנאספים מאפשרים לא רק לנתח את התהליכים לעומק, אלא גם לחשוף תבניות שעשויות להשפיע על היעילות של העסק.

נניח שחברה מנהלת מספר פרויקטים במקביל. מהנדס רוצה להבין כמה בטון כבר הוזמן וכמה עוד יש לרכוש. בגישה המסורתית, עליו לחפש ידנית בשרת ולפתוח מספר טבלאות אומדן, להשוות אותן עם דוחות ביצוע ולבדוק את המלאי: וסקריפטים אוטומטיים, המשימה נשארת חצי ידנית ETL הנוכחי. זה לוקח שעות, ולעיתים ימים. אפילו עם תהליכי מהנדס עדיין צריך לציין ידנית את הנתיב לתיקיות או לקבצים ספציפיים בשרת. זה מפחית את האפקט הכללי של האוטומציה, מכיוון שזה ממשיך לגזול זמן עבודה יקר.

כאשר עוברים לניהול נתונים במקום עבודה עם מערכת הקבצים של השרת, המהנדס מקבל גישה למבנה אחסון אחיד מאפשרת לקבל - LLM או אפילו פנייה לסוכן SQL שבו המידע מתעדכן בזמן אמת. בקשה אחת - בצורת קוד, שאילתת מיד נתונים מדויקים על המלאי הנוכחי, כמויות העבודה שבוצעו והאספקות הצפויות, אם הנתונים הוכנו מראש ואוחדו במבנה של מחסן נתונים, שבו אין צורך לנדוד בתיקיות, לפתוח עשרות קבצים ולהשוות ערכים ידנית.

Excel ומאות ואפלי גיליונות RVT מודלים, DWG שרטוטי, PDF במשך זמן רב, חברות בנייה השתמשו במסמכי, ופורמטים מפורזים אחרים, אשר נשמרים בתיקיות מסוימות על השרתים של החברה, מה שהקשה על חיפוש המידע בדיקתו וניתוחו. כתוצאה מכך, הקבצים שנותרו לאחר סיום הפרויקטים, לרוב מועברים חזרה לשרת לתיקיות אחסון ארכיוניות, אשר כמעט ואינן בשימוש בהמשך. אחסון נתונים מסורתי כזה, עם עליית זרם הנתונים, מאבד את הרלוונטיות שלו, בשל פגיעותו לטעויות הנגרמות על ידי גורם אנושי.

קובץ הוא בסך הכל מיכל מבודד שבו נשמרים נתונים. קבצים נוצרו עבור אנשים, ולא עבור מערכות או PDF מסמך, Excel ולכן הם דורשים פתיחה ידנית, קריאה ופרשנות. דוגמה לכך יכולה להיות גיליון שצריך לפתוח בכלי מסוים כדי לגשת למידע הנדרש. ללא חילוץ וניתוח מובנים, המידע, CAD שרטוט שבו נשאר לא מנוצל.

נתונים, לעומת זאת, הם מידע שניתן לקרוא על ידי מכונה, הקשור, מעודכן ומנותח אוטומטית. במחסן המידע מוצג בצורה של טבלאות, רשומות (Data Lake או DWH, כגון מסד נתונים) נתונים אחד וקשרים. זה מאפשר אחסון אחד, ביצוע שאילתות אוטומטיות, ניתוח ערכים ובניית דוחות בזמן אמת.

מאפשר להפסיק את תהליך החיפוש הידני ולמקד את תהליכי העיבוד (8.11 איור) השימוש בנתונים במקום קבצים. חברות שכבר מאמצות גישה כזו מקבלות יתרון תחרותי בזכות מהירות הגישה למידע ואפשרות האינטגרציה המהירה שלו בתהליכי העסקים.

המעבר משימוש בקבצים לנתונים הוא שינוי בלתי נמנע, שיגדיר את עתיד תעשיית הבנייה.

כל חברה בתעשיית הבנייה תתמודד עם בחירה מרכזית: להמשיך לאחסן מידע בקבצים מפוזרים ובסילואים, שצריכים להיקרא על ידי אנשים בעזרת תוכנות מיוחדות, או להמיר אותם בשלבים הראשונים של העיבוד לנתונים מובנים, וליצור בסיס דיגיטלי אינטגרטיבי לניהול אוטומטי של פרויקטים.

Files	Data
Stored in isolated structures	Open and accessible information flows
Stored in folders with limited access for a few employees	Stored in centralized databases and accessible across multiple processes
Require manual search, opening, and analysis	Can be automatically processed and analyzed in real time
Difficult to integrate between systems	Easily integrated and matched for analytics

- אבולוציה של זרימת המידע: מקבצים מבודדים לנתונים אינטגרטיביים 8.11 איור

בתנאים של גידול מהיר בנפחי המידע, שיטות האחסון והעיבוד המסורתיות של קבצים הופכות לפחות ופחות יעילות. בתעשיית הבנייה, כמו בתחומים אחרים, כבר לא מספיק להסתמך על תיקיות מפוזרות עם קבצים בפורמטים שונים או מסדי נתונים שאינם קשורים זה לזה.

חברות השואפות לשמור על תחרותיות בעידן הדיגיטלי, ייאלצו לעבור לפלטפורמות דיגיטליות אינטגרטיביות, להשתמש

בטכנולוגיות נתונים גדולים ובמערכות ניתוח אוטומטיות

המעבר מאחסון קבצים לעבודה עם נתונים ידרוש חשיבה מחדש על הגישות לניהול מידע ובחירה מודעת של פורמטים המתאימים לאינטגרציה עתידית במאגרים מרכזיים. הבחירה הזו משפיעה על כמה ביעילות הנתונים יטופלו, כמה מהר ניתן לגשת אליהם וכמה קל יהיה לשלבם בתהליכים דיגיטליים של החברה.

אחסון נתונים גדולים: ניתוח פורמטים פופולריים ויעילותם

פורמטי האחסון משחקים תפקיד מרכזי בהבטחת יכולת הרחבה, אמינות וביצועים של תשתית האנליטיקה. לצורך ניתוח מבנה פופולרי - **Pandas DataFrame** - ועיבוד נתונים - כמו סינון, קיבוץ ואגרציה - בדוגמאות שלנו השתמשנו בלעבודה עם נתונים בזיכרון.

אין פורמט אחסון משלו, ולכן לאחר סיום העיבוד הנתונים מיוצאים לאחד **Pandas DataFrame** - עם זאת, לפורמטים טבלאיים אלו נוחים להחלפה ותואמים לרוב המערכות. **CSV** או **XLSX** מהפורמטים החיצוניים - לרוב החיצוניות, אך יש להם מספר מגבלות: יעילות אחסון נמוכה, חוסר דחיסה ותמיכה חלשה בגרסאות

■ (ערכים מופרדים בפסיקים): פורמט טקסט פשוט, נתמך באופן רחב על ידי פלטפורמות וכלים שונים **CSV**. הוא פשוט לשימוש, אך אינו תומך בסוגי נתונים מורכבים ובדחיסה.

■ התומך בפונקציות **Microsoft Excel**, פורמט קבצי **XLSX (Excel Open XML Spreadsheet)**: מורכבות כמו נוסחאות, תרשימים ועיצוב. למרות שהוא נוח לניתוח ידני ולוויזואליזציה של נתונים, הוא אינו מותאם לעיבוד נתונים בקנה מידה גדול.

(8.12 איור) הפופולריים, ישנם מספר פורמטים פופולריים לאחסון יעיל של נתונים מובנים **CSV** ו-**XLSX** בנוסף ל- כל אחד מהם בעל יתרונות ייחודיים בהתאם לדרישות הספציפיות לאחסון וניתוח נתונים.

■ פורמט קבצים לאחסון נתונים בעמודות, מותאם לשימוש במערכות ניתוח נתונים. הוא **Apache Parquet**: מציע תכניות דחיסה וקידוד יעילות, מה שהופך אותו לאידיאלי עבור מבני נתונים מורכבים ועיבוד נתונים גדולים.

■ **Parquet** - בדומה ל: (אופטימיזציה של עמודות שורות - **Apache ORC (Optimized Row Columnar)**, מספק דחיסה גבוהה ואחסון נתונים יעיל. הוא מותאם לפעולות קריאה כבדות ומתאים היטב לאחסון אגני **ORC** נתונים.

■ אינו כל כך יעיל מבחינת אחסון נתונים **JSON** - למרות ש: **JSON (JavaScript Object Notation)** הוא מאוד נגיש ופשוט לשימוש, מה שהופך אותו **ORC** או **Parquet** בהשוואה לפורמטים בינאריים כמו לאידיאלי לתרחישים שבהם קריאות ועמידות עם טכנולוגיות אינטרנט חשובות.

■ פורמט אחסון בינארי קל, מהיר ופשוט לשימוש, הממוקד באנליטיקה. הוא נועד להעברת נתונים: **Feather** מה שהופך אותו לבחירה מצוינת עבור פרויקטים שבהם מעורבות **R** (ל- **Python (Pandas)** בצורה יעילה בין הסביבות התכנותיות הללו.

■ (פורמט נתונים היררכי גרסה 5): (מיועד לאחסון וארגון כמויות גדולות של נתונים. הוא תומך במגוון **HDF5** פופולרי במיוחד בחישובים **HDF5**. רחב של סוגי נתונים ומעולה לעבודה עם אוספים מורכבים של נתונים מדעיים בזכות יכולתו לאחסן ולגשת ביעילות לסטים גדולים של נתונים.

		XLSX	CSV	Apache Parquet	HDF5	Pandas DataFrame
	Storage	Tabular	Tabular	Columnar	Hierarchical	Tabular
	Usage	Office tasks, data presentation	Simple data exchange	Big data, analytics	Scientific data, large volumes	Data analysis, manipulation
	Compression	Built-in	None	High	Built-in	None (in-memory)
	Performance	Low	Medium	High	High	High (memory dependent)
	Complexity	High (formatting, styles)	Low	Medium	Medium	Low
	Data Type Support	Limited	Very limited	Extended	Extended	Extended
	Scalability	Low	Low	High	High	Medium (memory limited)

השוואת פורמטי נתונים עם ציון ההבדלים העיקריים בהיבטים של אחסון ועיבוד 8.12 איור.

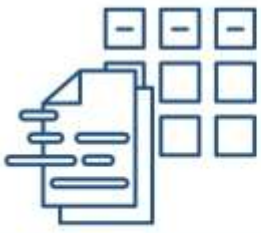
הוכנה טבלה הממחישה את גדלי ETL בתהליך Load-לצורך ביצוע ניתוח השוואתי של הפורמטים המיושמים בשלב הקבצים ואת זמני הקריאה שלהם) איור 8.13. (במסגרת המחקר נעשה שימוש בקבצים עם נתונים זהים: הטבלה כללה שורות ו-10 עמודות, מלאות בערכים אקראיים 10,000.

כמו גם את הגרסאות המכופפות שלהם, XLSX ו-HDF5, CSV, Parquet: במחקר כללו את פורמטי האחסון הבאים Pandas DataFrame והוצגו במבנה של NumPy הנתונים המקוריים נוצרו באמצעות ספריית ZIP. בארכיוני תהליך הבדיקה כלל את השלבים הבאים:

- כל XLSX ו-HDF5, CSV, Parquet: שמירת קבצים: הדאטה פריים נשמר בארבעה פורמטים שונים. פורמט בעל תכונות ייחודיות בדרך האחסון של הנתונים, המשפיעות על גודל הקובץ ומהירות הקריאה שלו.
- ZIP. לצורך ניתוח היעילות של הדחיסה הסטנדרטית, כל קובץ דחוס נוסף בארכיון ZIP-דחיסת קבצים ב-
- זה מאפשר ZIP-זמן הקריאה נמדד עבור כל קובץ לאחר פריקתו מ: (ETL – Load) קריאת קבצים. להעריך את מהירות הגישה לנתונים לאחר הוצאתם מהארכיון.

לא שימש ישירות בניתוח הגדלים או זמני הקריאה, שכן הוא אינו מהווה פורמט Pandas DataFrame חשוב לציין כי אחסון עצמאי. הוא שימש רק כמבנה ביניים ליצירה ושמירה של הנתונים בפורמטים שונים.



		Original Size (MB)	ZIP Size (MB)	Read Time (s)
Structured table 100000 rows x 10 columns 	CSV	18.82	8.73	0.48
	Parquet	9.67	9.22	0.17
	XLSX	12.59	12.59	41.73
	HDF5	8.50	7.62	0.13

איור 8.13 השוואת פורמטי אחסון נתונים לפי גודל ומהירות קריאה.

מדגימים) איור 8.13 (יעילות גבוהה בדחיסה, מפחיתים באופן משמעותי את גודלם כאשר הם CSV ו-HDF5 קבצי לעומת, XLSX מה שיכול להיות מועיל במיוחד בתרחישים הדורשים אופטימיזציה של האחסון. קבצי, ZIP-נארזים ב נשאר דומה למקורי, מה שהופך אותם לפחות משתלמים לשימוש ZIP-זאת, כמעט שאינם ניתנים לדחיסה, וגודלם ב גבוה XLSX בנפחים גדולים של נתונים או בתנאים שבהם מהירות הגישה לנתונים חשובה. בנוסף, זמן הקריאה עבור משמעותית בהשוואה לפורמטים אחרים, מה שהופך אותו לפחות מועדף עבור פעולות קריאה מהירות של נתונים בזכות המבנה הקולונרי שלו, הראה יעילות גבוהה עבור משימות אנליטיות ונפחים גדולים של Apache Parquet, -נתונים.

Apache Parquet אופטימיזציה של אחסון נתונים עם

פורמט זה פותח במיוחד לאחסון Apache Parquet. אחד הפורמטים הפופולריים לאחסון ועיבוד נתונים גדולים הוא מה שמאפשר לצמצם באופן משמעותי את נפח האחסון ולהגביר את מהירות הבקשות, (Pandas-בדומה ל) קולונרי תומך בדחיסה מובנית ומאפיין בעבודה עם CSV, XLSX, Parquet והאנליטיות. בניגוד לפורמטים המסורתיים, כמו ואחסון בענן Spark, Hadoop מערכות עיבוד נתונים גדולים, כולל.

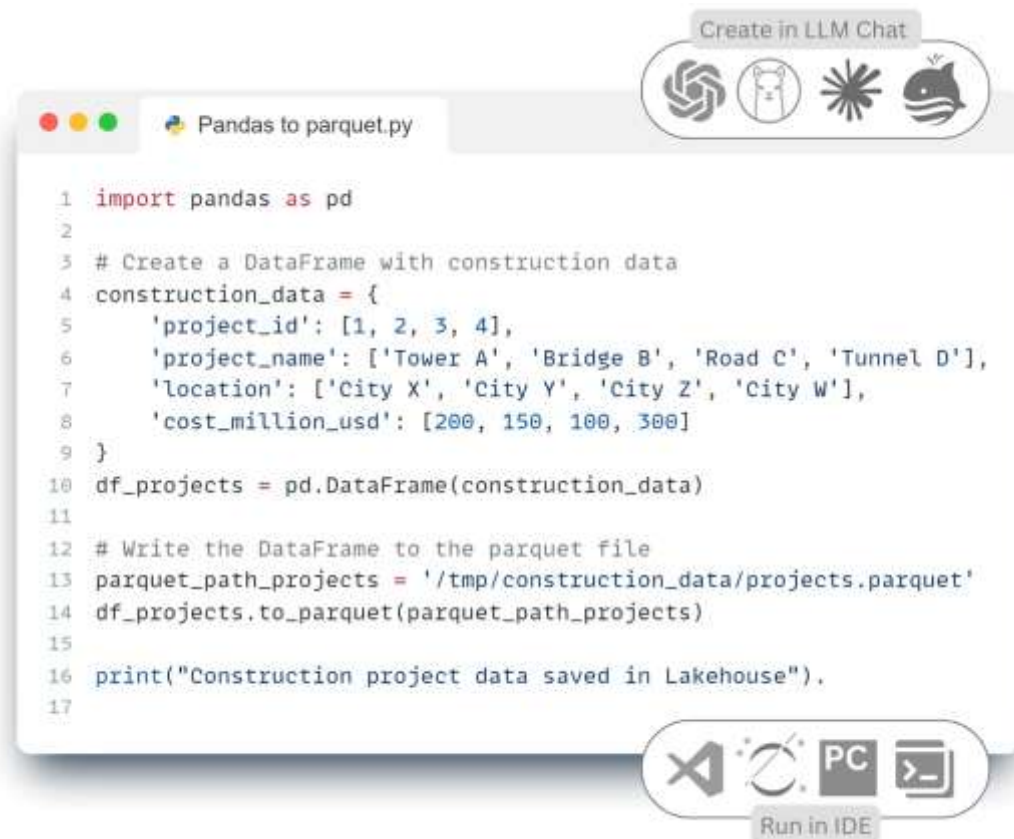
כוללות תמיכה בדחיסה ובקידוד נתונים, מה שמפחית באופן משמעותי את Parquet תכונות מרכזיות של גודל האחסון ומאיץ את פעולות הקריאה של הנתונים על ידי עבודה ישירה עם העמודות הנדרשות, ולא עם כל השורות של הנתונים.

LLM-נשתמש ב, Apache Parquet-לדוגמה מוחשית כיצד ניתן בקלות להשיג את הקוד הנדרש להמרת נתונים ל

LLM (ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN): שלחו בקשה טקסטואלית בצ'אט

Apache Parquet-ל-Pandas DataFrame-כתוב קוד לשמירת נתונים מ

LLM תשובת: ?



```

1 import pandas as pd
2
3 # Create a DataFrame with construction data
4 construction_data = {
5     'project_id': [1, 2, 3, 4],
6     'project_name': ['Tower A', 'Bridge B', 'Road C', 'Tunnel D'],
7     'location': ['City X', 'City Y', 'City Z', 'City W'],
8     'cost_million_usd': [200, 150, 100, 300]
9 }
10 df_projects = pd.DataFrame(construction_data)
11
12 # Write the DataFrame to the parquet file
13 parquet_path_projects = '/tmp/construction_data/projects.parquet'
14 df_projects.to_parquet(parquet_path_projects)
15
16 print("Construction project data saved in Lakehouse").
17
  
```

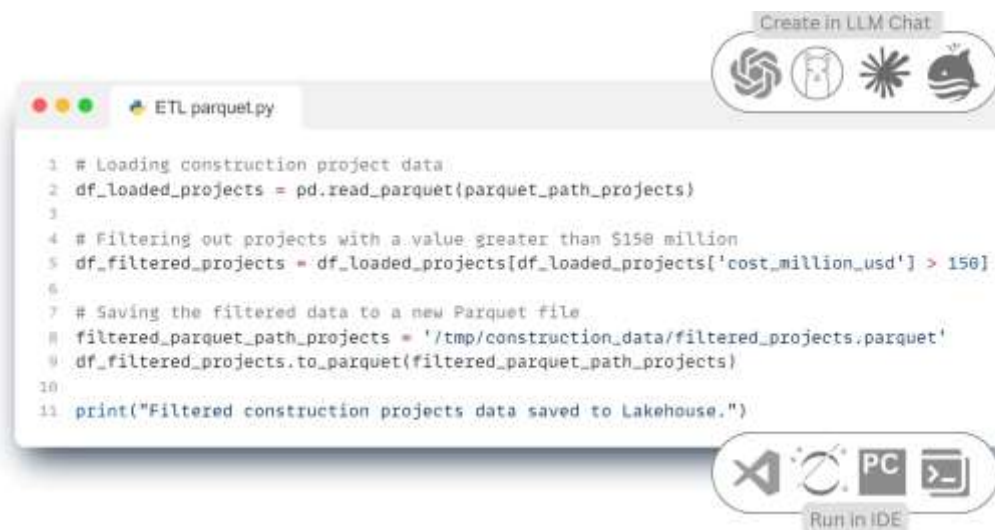
Python. באמצעות מספר שורות **Apache Parquet** איור 8.14 העברת נתונים מדאטה פריים מהזיכרון למבנה אחסון יעיל עבור

כדי לסנן פרויקטים לפי ערך מסוים של Parquet עם נתונים שנשמרו בפורמט ETL-דוגמה הבאה: נמודל את תהליך ה-
 (איור 8.14) "cost_million_usd" אחד מהמאפיינים

LLM-בהמשך השיחה, שלחו בקשה טקסטואלית ל

כתוב קוד שבו אנו רוצים לסנן את הנתונים בטבלה ולשמור רק את הפרויקטים (שורות הטבלה) עולה על 150 (cost_million_usd פרמטר) שהעלות שלהם, Apache Parquet, מנתוני מיליון דולר.

LLM: תשובת



נראה כמו עם פורמטים מובנים אחרים **Apache Parquet** בעבודה עם נתונים בפורמט **ETL-איור 8.15** תהליך ה

מפחית באופן משמעותי את נפח המידע המאוחסן ומאיץ (וכו XLSX, CSV-בהשוואה ל) Parquet השימוש בפורמט משתלב עם מערכות עיבוד Parquet. את פעולות החיפוש. בזכות זה הוא מתאים מאוד גם לאחסון וגם לניתוח נתונים. שונות, ומספק גישה יעילה בארכיטקטורות היברידיות

עם זאת, פורמט אחסון יעיל הוא רק אחד מהמרכיבים של עבודה מלאה עם נתונים. כדי ליצור סביבה עמידה ומדרגת הן מספקות. DWH (Data Warehouse) נדרשת ארכיטקטורה מנוהלת היטב. תפקיד זה ממלאות מערכות מסוג ואלגוריתמים BI אגרגציה של נתונים ממקורות שונים, שקיפות בתהליכי העסקים ואפשרות לניתוח כולל באמצעות כלי של למידת מכונה

מחסן נתונים DWH:

מותאם לאינטגרציה DWH-מותאם לאחסון יעיל של כמויות גדולות של מידע, ה Parquet בדומה לכך, כמו שפורמט ולמבנה של נתונים במטרה לתמוך בניתוח, חיזוי וקבלת החלטות ניהוליות

רישום חשבונאי, CRM, מערכות, ERP, CAFM, CPM: בחברות מודרניות, נתונים מגיעים ממקורות רבים ומגוונים ופתרונות נוספים. כדי לקבל תמונה כוללת, לא מספיק פשוט IoT של מבנים, חיישני CAD ומלאי, מודלים דיגיטליים, מערכת אחסון מרכזית - DWH-לאסוף נתונים - יש לארגן, לסטנדרטיזציה ולרכזם באחסון אחד. תפקיד זה ממלא ה המאפשרת לאגרגט מידע ממקורות שונים, למבנה אותו ולהפוך אותו לזמין לניתוח ולניהול אסטרטגי

זו מערכת אחסון נתונים מרכזית, המאגדת מידע ממקורות רבים - DWH (Data Warehouse) מסדרת אותו ומאפשרת גישה לניתוח ודיווח.

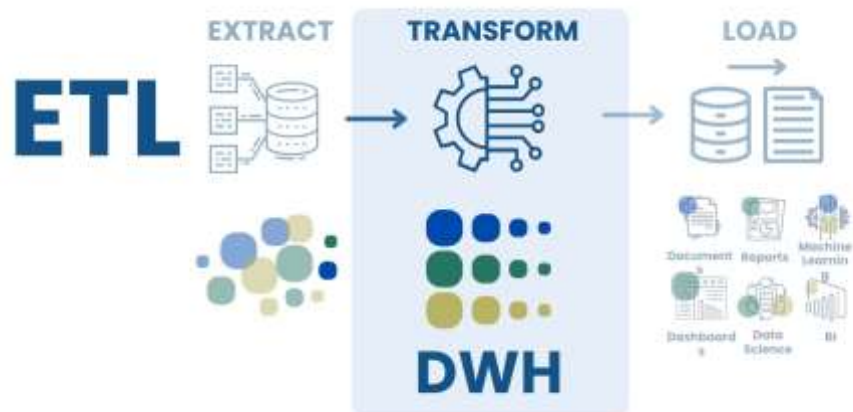
DWH-בחברות רבות, נתונים מפוזרים במערכות שונות, אשר דנו בהן בחלקים הראשונים של הספר (איור 1.24). זהו מסד נתונים מיוחד (מסד DWH מאחד את המקורות הללו, ומספק שקיפות מלאה ואמינות של המידע. מחסן הנתונים DWH-נתונים גדול, (המאגד, מעבד ושומר נתונים ממקורות רבים. תכונות עיקריות של ה

■ חילוף נתונים ממקורות, ניקוי, טרנספורמציה - ETL (Extract, Transform, Load) שימוש בתהליכי טעינה לאחסון ואוטומציה של תהליכים אלו, עליהם דיברנו בחלק השביעי של הספר.

■ יכולים להישמר הן בצורה מצטברת (דוחות מסכמים (והן בצורה DWH-גרנולריות הנתונים – נתונים ב החלו לדבר על נתונים גרנולריים, מה שעשוי CAD מפורטת) נתונים גולמיים. (החל משנת 2024, ספקי להעיד על הכנה של התעשייה למעבר לשימוש במחסני נתונים ייעודיים בענן לעבודה עם נתוני מודלים דיגיטליים של בניינים.

■ ולמידת מכונה Big Data ניתוח, BI תמיכה בניתוח וחיזוי – מחסני נתונים מספקים בסיס לכלי.

משמש כבסיס לניתוח עסקי, מאפשר לנתח מדדי ביצוע מרכזיים, לחזות מכירות, רכישות והוצאות, כמו גם DWH -לאוטומט את הדיווח והוויזואליזציה של הנתונים.



יכול לשמש כמחסן מרכזי, שבו הנתונים, שהופקו ממערכות שונות, עוברים שלבי ETL DWH דיאגרמה 8.16 בתהליך טרנספורמציה וטעינה.

משחק תפקיד מרכזי באינטגרציה, ניקוי ומבנה של מידע, יוצר בסיס חזק לניתוח עסקי ולתהליכי קבלת החלטות DWH עם זאת, בתנאים המודרניים, כאשר כמויות הנתונים הולכות וגדלות ומקורותיהם נעשים מגוונים יותר, הגישה המסורתית ואגם נתונים ELT לעיתים קרובות דורשת הרחבה בצורת גישות DWH-לאחסון מידע ב

מניקוי מסורתי לעיבוד גמיש: ELT ל-ETL אגם נתונים - אבולוציה של

קלאסיים, המיועדים לאחסון נתונים מובנים בפורמט, המותאם לבקשות ניתוח, מתמודדים עם מגבלות DWH מחסני

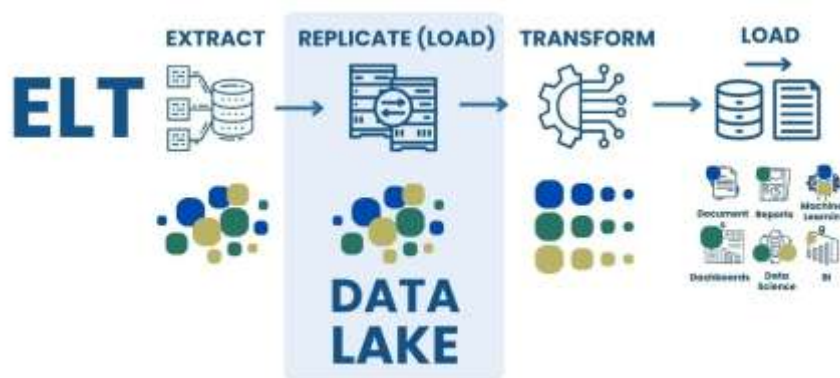
המציעים (Data Lakes), בעבודה עם נתונים לא מובנים וביכולת ההתרחבות. בתגובה לבעיות אלו, הופיעו אגמי נתונים אחסון גמיש של כמויות גדולות של נתונים הטרוגניים.

המאפשרת עבודה עם נתונים לא מובנים, חצי מובנים ונתונים, DWH-אגם נתונים מציע גישה חלופית ל גולמיים ללא תבנית קשיחה מראש. שיטה זו לאחסון נתונים רלוונטית לעיבוד נתונים בזמן אמת, למידת הממין ומאגד נתונים לפני הטעינה, אגם נתונים מאפשר לשמור, DWH-מכונה וניתוח מתקדם. בניגוד ל את המידע במצבו המקורי, ובכך מספק גמישות ויכולת התרבות.

ועניין ב"נתונים גדולים" הביאו להופעת אגמי נתונים, שבהם (RDBMS, DWH) אכזבה ממחסני הנתונים המסורתיים מורכב, הנתונים פשוט נטענים לאחסון חלש-מבנה, ועיבודם מתבצע בשלב הניתוח ETL במקום.

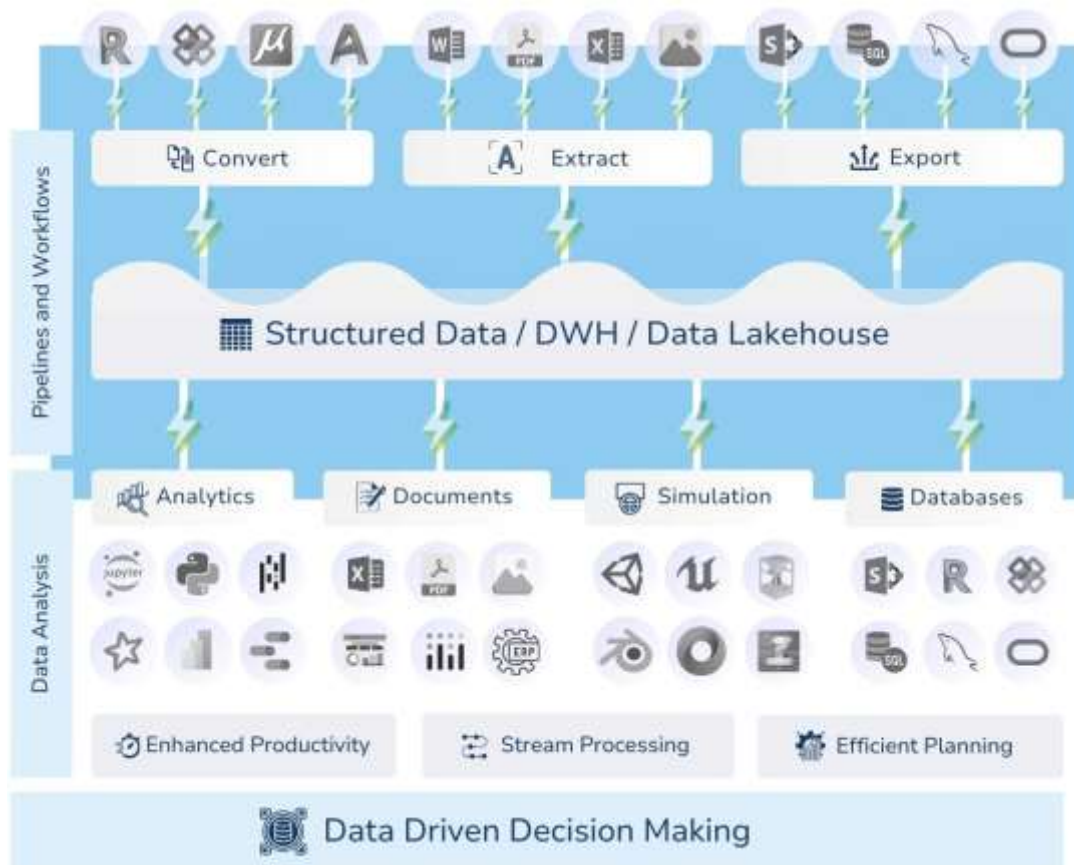
■ (ETL - Extract, Transform, Load) במחסני נתונים מסורתיים, הנתונים בדרך כלל עוברים עיבוד מקדים, טרנספורמציה וניקוי לפני טעינתם לאחסון. זה אומר שהנתונים ממין ומאופיינים כדי לפתור משימות ניתוח (Transform, Load) ודיווח ספציפיות בעתיד. הדגש הוא על שמירה על ביצועי בקשות גבוהים ושלמות הנתונים. עם זאת, גישה זו עשויה להיות יקרה ופחות גמישה מבחינת אינטגרציה של סוגי נתונים חדשים ותבניות נתונים המשתנות במהירות.

■ אגמי נתונים, מצד שני, מיועדים לאחסון כמויות גדולות של נתונים גולמיים בפורמט המקורי שלהם (איור 8.17) (חילוץ, טעינה, המרה), (שבו הנתונים ELT) חילוץ, המרה, טעינה, (נכנס לתמונה ETL במקום תהליך. 8.17) נטענים תחילה לאחסון "כפי שהם" ורק לאחר מכן יכולים להיות מומצים ומנותחים לפי הצורך. זה מספק גמישות רבה ואפשרות לאחסון נתונים הטרוגניים, כולל נתונים לא מובנים, כמו טקסט, תמונות ויומנים.



שבו המידע נטען תחילה "במצב גולמי", ובהמרה מתבצעת בשלב, ELT-באגם נתונים משתמשים ב, ETL-איור 8.17 בניגוד ל היציאה.

מחסני נתונים מסורתיים מתמקדים בעיבוד מקדים של נתונים כדי להבטיח ביצועי שאילתות גבוהים, בעוד שבאגמי נתונים ניתן עדיפות לגמישות: הם מאחסנים נתונים גולמיים וממירים אותם לפי הצורך (איור 8.18).



איור 8.18 הקונספטים המודרניים של מחסנים מכוונים לאחסון ועיבוד כל סוגי הנתונים למטרות קבלת החלטות.

עם זאת, למרות כל היתרונות, אגמי נתונים אינם חסרי חסרונות. חוסר מבנה נוקשה ומורכבות ניהול המידע יכולים להוביל לכאוס שבו הנתונים משוכפלים, סותרים זה את זה או מאבדים את הרלוונטיות שלהם. בנוסף, חיפוש וניתוח נתונים באחסון כזה דורשים מאמצים משמעותיים, במיוחד בעבודה עם מידע הטרונגי. כדי להתגבר על מגבלות אלו Data Lakehouse ולשלב את היתרונות של מחסנים מסורתיים ואגמי נתונים, פותחה ארכיטקטורת

סינרגיה של מחסני נתונים ואגמי נתונים: Data Lakehouse אדריכלות

(מבנה, ניהול, ביצועי ניתוח גבוהים (ואגם נתונים) סקלאביליות DWH כדי לשלב את התכונות הטובות ביותר של ארכיטקטורה זו משלבת את הגמישות של אגמי Data Lakehouse עבודה עם נתונים הטרונגיים, (פותח הגישה של נתונים עם כלים עוצמתיים לעיבוד וניהול, האופייניים למחסנים מסורתיים, ומספקת איזון בין אחסון, ניתוח ולמידת הוא סינתזה של אגמי נתונים ומחסני נתונים, המשלבת את הגמישות והסקלאביליות של Data Lakehouse. מכונה הראשונים עם הניהול והאופטימיזציה של השניים

הוא גישה ארכיטקטונית השואפת לשלב את הגמישות והסקלאביליות של אגמי Data Lakehouse (נתונים עם הניהול וביצועי השאילות במחסני נתונים) איור 8.19

כוללות Data Lakehouse תכונות מפתח של

מבטיח, Apache Parquet פורמט אחסון נתונים פתוח: השימוש בפורמטים פתוחים לאחסון נתונים, כמו יעילות ואופטימיזציה של שאלות.

תומך DWH, Lakehouse-סכימה רק לקריאה: בניגוד לגישה המסורתית של סכימה בלעדית לכתובה ב בסכימה רק לקריאה, מה שמאפשר ניהול גמיש יותר של מבנה הנתונים.

גמישות וסקלאביליות: תומך באחסון וניתוח נתונים מובנים ולא מובנים, ומבטיח ביצועי שאלות גבוהים באמצעות אופטימיזציה ברמת האחסון.

מציע פתרון פשרה, המשלב את היתרונות של שני הגישות, מה שהופך אותו לאידיאלי עבור Data Lakehouse העומסים האנליטיים המודרניים, הדורשים גמישות בעיבוד נתונים.



8.19 **Data Lakehouse** איור - הדור הבא של מערכות אחסון נתונים, שנוצרו כדי לענות על הדרישות המורכבות והמשתנות - תדיר.

רעיון של מחסני נתונים מודרניים נראה פשוט: אם כל הנתונים נמצאים במקום אחד, קל יותר לנתח אותם. עם זאת, במציאות הדברים אינם כל כך חלקים. דמיינו שחברה החליטה להפסיק לחלוטין להשתמש במערכות ניהול וחשבונאות? והחליפה אותן באגם נתונים אחד גדול, אליו יש גישה לכולם. מה יקרה, (או אחרות ERP, PMIS, CAFM) מוכרות סביר להניח שיתחיל כאוס: הנתונים יוכפלו, יתנגדו זה לזה, ומידע קריטי ייאבד או יתעוות. אפילו אם אגם הנתונים משמש רק לניתוח, ללא ניהול נכון, מתעוררות בעיות רציניות.

קשה להבין את הנתונים: במערכות רגילות לנתונים יש מבנה ברור, ואילו באגם יש פשוט אוסף עצום של קבצים וטבלאות. כדי למצוא משהו, על המומחה להבין - מה מייצגת כל שורה ועמודה.

הנתונים עשויים להיות לא מדויקים: אם במקום אחד מאוחסנות גרסאות רבות של אותה מידע, קשה להבין איזו מהן עדכנית. כתוצאה מכך, מתקבלות החלטות על סמך נתונים מיושנים או שגויים.

קשה להכין את הנתונים לעבודה: הנתונים צריכים לא רק להיות מאוחסנים, אלא גם להיות מוצגים בצורה נוחה - בדוחות, גרפים, טבלאות. במערכות מסורתיות זה מתבצע אוטומטית, בעוד שבאגמי נתונים זה דורש עיבוד נוסף.

בסופו של דבר, כל קונספט של אחסון נתונים יש לו את המאפיינים, הגישות לעיבוד והשימושים בעסקים. מסדי נתונים (Data Lake) מספקים מבנה לניתוח, אגמי נתונים (DWH) מסורתיים מכוונים לפעולות טרנזקציה, מחסני נתונים ואגם DWH משלבים את היתרונות של (Data Lakehouse) מאחסנים מידע בצורה גולמית, ואחסונים היברידיים (8.110 איור) נתונים.

	Traditional Approach	Data Warehouse	Data Lake	Data Lakehouse
Data Types	Relational Databases	Structured, ready for analytics	Raw, semi-structured, or unstructured	Mix of structured and unstructured
Use Cases	Transactional Systems	Reporting, dashboards, BI	Big data storage, AI, advanced analytics	Hybrid analytics, AI, real-time data
Processing	OLTP – real-time transactions	ETL – clean and structure before analysis	ELT – store raw data, transform later	ELT with optimized storage and real-time processing
Storage	On-premise servers	Centralized, SQL-based	Decentralized, flexible formats	Combines advantages of DWH and DL
Common Tools	MySQL, PostgreSQL	Snowflake, Redshift, BigQuery	Hadoop, AWS S3, Azure Data Lake	Databricks, Snowflake, Google BigLake

ההבדלים המרכזיים בסוגי נתונים, תרחישי שימוש, שיטות עיבוד: **Data Lakehouse** ו-**DWH, Data Lake** 8.110 איור וגישות אחסון.

הבחירה בארכיטקטורת אחסון נתונים היא תהליך מורכב, התלוי בצרכי העסק, בנפח המידע ובדרישות לניתוח. כל פתרון rarely איזון ביניהם. ארגונים - Lakehouse-גמישות, ו- Data Lake, מספק מבנה DWH: יש לו יתרונות וחסרונות מגבילים את עצמם לארכיטקטורה אחת של נתונים.

ללא קשר לארכיטקטורה הנבחרת, מערכות ניהול נתונים אוטומטיות עולות בהרבה על שיטות ידניות. הק. מאפשרות למזער טעויות אנוש, מזרזות את עיבוד המידע, ומספקות שקיפות ומעקב אחר הנתונים בכל שלב של תהליכי העסק.

ואם מחסני נתונים מרכזיים כבר הפכו לסטנדרט תעשייתי בתחומים רבים בכלכלה, אז בתחום הבנייה המצב עדיין מפוצל. מה שמקשה על יצירת תמונה כוללת ודורש, (ועוד CDE, PMIS, ERP) הנתונים כאן מפוזרים בין פלטפורמות שונות. ארכיטקטורות שיכולות לאחד את המקורות הללו לסביבה דיגיטלית שלמה, המתאימה לניתוח.

ואגם נתונים DWH או CDE, PMIS, ERP

(Common Data Environment, CDE) בחלק מהחברות הפועלות בתחום הבנייה והעיצוב, כבר מיושמת הקונספט של סביבה משותפת לנתונים מבצע את אותן פונקציות כמו מחסן נתונים CDE, למעשה ISO 19650-בהתאם ל (Common Data Environment, CDE) בתעשיות אחרות: מרכז מידע, מספק שליטה על גרסאות, ומעניק גישה למידע מאומת (DWH).

היא מרחב דיגיטלי מרכזי, המשמש לניהול, אחסון, החלפה ושיתוף (CDE) הסביבה הדיגיטלית המרכזית מיושמת לעיתים קרובות באמצעות CDE. פעולה עם מידע פרויקט בכל שלב של מחזור חיי הפרויקט CAD (BIM) טכנולוגיות ענן ומשולבת עם מערכות.

המגזר הפיננסי, הקמעונאות, הלוגיסטיקה והתעשייה משתמשים במשך עשורים במערכות ניהול נתונים מרכזיות מפתחת עקרונות אלו, מתאימה אותם CDE. המאחדות מידע ממקורות שונים, בודקות את עדכניותו ומספקות אנליטיקה למשימות תכנון וניהול מחזור חיי הבניינים.

מסדרת את הנתונים, מתעדת שינויים ומספקת גישה אחידה למידע מאומת. עם המעבר לטכנולוגיות CDE, DWH כמו נתונים מפורטים CDE-ענן ואינטגרציה עם כלים אנליטיים, ההבדלים ביניהם הולכים ומצטמצמים. כאשר מוסיפים ל קלאסיים DWH מאז 2023, ניתן לראות עוד יותר מקבילות עם CAD שהקונספט שלהם נדון על ידי ספקי

תכנון (ERP)-מערכת מידע לניהול פרויקטים (PMIS-בבנייה "כבר דנו ב PMIS ו-ERP קודם לכן בפרק "מערכות משמשת כמחסן נתונים, כולל שרטוטים CDE: פועלים יחד PMIS ו-CDE, משאבים ארגוניים. (בפרויקטים בנייה מנהלת תהליכים כמו בקרה על לוחות זמנים, משימות, משאבים ותקציב PMIS-מודלים ודוקומנטציה פרויקטית, בעוד ש

מה שמספק PMIS, האחראית על ניהול העסק בכללותו) כספים, רכש, כוח אדם, ייצור, (יכולה להשתלב עם ERP, אשר יסייע באיסוף, סדר, DWH-בקרה על עלויות ותקציבים ברמת החברה. עבור אנליטיקה ודיווח ניתן להשתמש ב ולזהות מגמות. בתורו (KPI) מה שיאפשר להעריך מדי ביצוע פיננסיים, ERP ו-PMIS, CDE-ואגירה של נתונים מ על ידי אחסון נתונים גולמיים ולא מובנים) כגון יומנים, נתוני חישנים, DWH יכול להשלים את Data Lake (DL) לצורך ניתוח נוסף DWH-תמונות. (נתונים אלו יכולים להיות מעובדים ומועלים ל

Data ו-DWH מתמקדת בתהליכי עסק, בעוד ש ERP, מתמקדות בניהול פרויקטים PMIS ו-CDE, כך, מתמקדות באנליטיקה ועבודה עם נתונים Lake

ניתן להבחין בהבדלים משמעותיים מבחינת Data Lake ו-DWH עם ERP ו-PMIS, CDE בהשוואת מערכות. עצמאות מספקים, עלות, גמישות אינטגרציה, עצמאות נתונים, מהירות התאמה לשינויים, כמו גם יכולות אנליטיות לעיתים קרובות קשורות לפתרונות וסטנדרטים ספציפיים של ספקים, ERP ו-PMIS, CDE מערכות מסורתיות, כמו מה שהופך אותן לפחות גמישות ומגביר את עלותן בשל רישיונות ותמיכה. בנוסף, נתונים במערכות אלו לעיתים קרובות -כלואים בפורמטים סגורים של ספקים, מה שמגביל את השימוש והניתוח שלהם.

		CDE, PMIS, ERP	DWH, Data Lake
	Vendor Dependency	High (tied to specific solutions and standards of vendors)	Low (flexibility in tool and platform choice)
	Integration Flexibility	Limited (integration depends on vendor solutions)	High (easily integrates with various data sources)
	Cost	High (licensing and support costs)	Relatively lower (use of open technologies and platforms)
	Data Independence	Low (data often locked in proprietary formats)	High (data stored in open and accessible formats)
	Adaptability to Changes	Slow (changes require vendor approval and integration)	Fast (adaptation and data structure modification without intermediaries)
	Analytical Capabilities	Limited (dependent on vendor-provided solutions)	Extensive (support for a wide range of analytical tools)

- ERP ו-PMIS, CDE מציעים גמישות רבה יותר ועצמאות נתונים בהשוואה למערכות כמו Data Lake ו-DWH.

מספקים גמישות רבה יותר באינטגרציה עם מקורות נתונים שונים, והשימוש Data Lake ו-DWH, בניגוד לכך תומכים Data Lake ו-DWH, בטכנולוגיות ופלטפורמות פתוחות מסייע בהפחתת עלות הבעלות הכוללת. יתרה מכך במגוון רחב של כלים אנליטיים, מה שמרחיב את אפשרויות הניתוח והניהול.

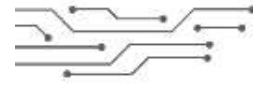
ובזמן שיש גישה לבסיסי נתונים של יישומי CAD עם התפתחות הכלים להנדסה הפוכה עבור פורמטי השאלה נעשית דחופה יותר: עד כמה זה מוצדק להמשיך להשתמש בפלטפורמות סגורות ומבודדות, CAD, כאשר נתוני הפרויקט צריכים להיות זמינים למגוון רחב של מומחים העובדים בעשרות ארגוני קבלנות ועיצוב?

תלות טכנולוגית זו בספק מסוים יכולה להגביל באופן משמעותי את הגמישות בניהול הנתונים, להאט את התגובות לשינויים בפרויקט ולמנוע שיתוף פעולה יעיל בין המשתתפים.

מתמקדות בעיקר באחסון, מבנה - PMIS ו-CDE, DWH, Data Lake הגישות המסורתיות לניהול נתונים - כולל, ועיבוד מידע. עם זאת, עם התפתחות הבינה המלאכותית ולמידת המכונה, עולה הצורך בדרכים חדשות לארגון נתונים המאפשרות לא רק לאגרגט, אלא גם לחשוף קשרים מורכבים, למצוא דפוסים סמויים ולהבטיח גישה מיידית למידע הרלוונטי ביותר.

בכיוון זה, תפקיד מיוחד מתחילות לשחק מסדי נתונים וקטוריים - סוג חדש של מאגרי נתונים, המותאמים לעבודה עם

אמבדינגים בעלי ממד גבוה.



פרק 8.2.

ניהול מחסני נתונים ומניעת כאוס

מאגרי נתונים וקטוריים ותיבת גבול

מסדי נתונים וקטוריים הם סוג חדש של מאגרי נתונים, שאינם רק שומרים נתונים, אלא מאפשרים חיפוש לפי משמעות. השוואת אובייקטים לפי קרבה סמנטית ויצירת מערכות אינטליגנטיות: מהמלצות ועד ניתוח אוטומטי ויצירת הקשר בניגוד למסדי נתונים מסורתיים, המתמקדים בהתאמות מדויקות, מסדי נתונים וקטוריים מוצאים אובייקטים דומים על סמך תכונות - גם בהיעדר התאמה מדויקת.

מסד נתונים וקטורי הוא סוג מיוחד של מסד נתונים, השומר נתונים בצורה של וקטורים רב-ממדיים כאשר כל אחד מהם מייצג תכונות או איכויות מסוימות. הווקטורים יכולים להיות בעלי מספר ממדים - שונה, בהתאם למורכבות הנתונים) במקרה אחד זה יכול להיות מספר ממדים מועט, ובמקרה אחר (אלפים).

היתרון העיקרי של מסדי נתונים וקטוריים הוא החיפוש לפי משמעות סמנטית, ולא לפי התאמה מדויקת של ערכים (k-NN) עם מסננים "שווה" או "מכיל", נעשה שימוש בחיפוש של השכנים הקרובים SQL ו-Pandas במקום שאלות. במרחב התכונות (בחלק הבא של הספר k-NN נרחיב על).

(מודלים גדולים של שפה (ודגמים גנרטיביים, האינטראקציה עם מסדי הנתונים מתחילה להשתנות LLM עם התפתחות עכשיו אפשר לבקש נתונים בשפה טבעית, לקבל חיפוש סמנטי במסמכים, לחלץ אוטומטית מונחים מרכזיים ולבנות או בהבנה של מבנה הטבלאות. נרחיב על כך בפרק SQL-קשרים הקשריים בין אובייקטים - כל זה ללא צורך בידע ב"ותפקידם בעיבוד נתונים ובהליכי עסקים LLM").

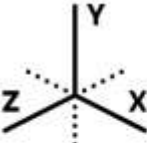
"לא מסדרים את המידע באופן אוטומטי ולא מביאים אותו לסדר. המודל פשוט "שוחה LLM-עם זאת, חשוב להבין ש בתוך מאגר הנתונים ומוצא את הקטע המתאים ביותר, בהתבסס על הקשר הבקשה. אם הנתונים לא נוקו או הוסדרו יזכיר ניסיון למצוא תשובה ב"פסולת" דיגיטלית - זה עשוי לעבוד, אבל איכות (deep search) מראש, החיפוש העמוק -התוצאות תהיה נמוכה יותר. אידיאלי יהיה אם ניתן יהיה לסדר את הנתונים מראש) למשל, להמיר מסמכים ל ולהעלות אותם למסד הנתונים הווקטורי. זה מעלה באופן משמעותי את הדיוק והרלוונטיות של התוצאות (Markdown).

בהתחלה, מסדי נתונים וקטוריים שימשו בלמידת מכונה, אך כיום הם מוצאים שימושים הולכים ומתרבים מעבר לכך - במערכות חיפוש, בהתאמה אישית של תוכן, ובאנליטיקה אינטלקטואלית.

(קופסה מגבילה. (זו מבנה גיאומטרי Bounding Box-אחד מהדוגמאות הבולטות ביותר לגישה וקטורית בבנייה הוא ה מוגדר על ידי הקואורדינטות המינימליות Bounding Box-שמתאר את הגבולות של אובייקט במרחב תלת-ממדי. ה. ויוצר "קופסה" סביב האובייקט. שיטה זו מאפשרת להעריך את המידות והמיקום של X, Y ו-Z, והמקסימליות על צירי האלמנט מבלי הצורך לנתח את כל הגיאומטריה.

כבר 6 ממדים) איור - [רוחב, גובה, עומק, x, y, z], ניתן לייצג כווקטור במרחב רב-ממדי: לדוגמה Bounding Box כל

8.21).-



Bounding Box

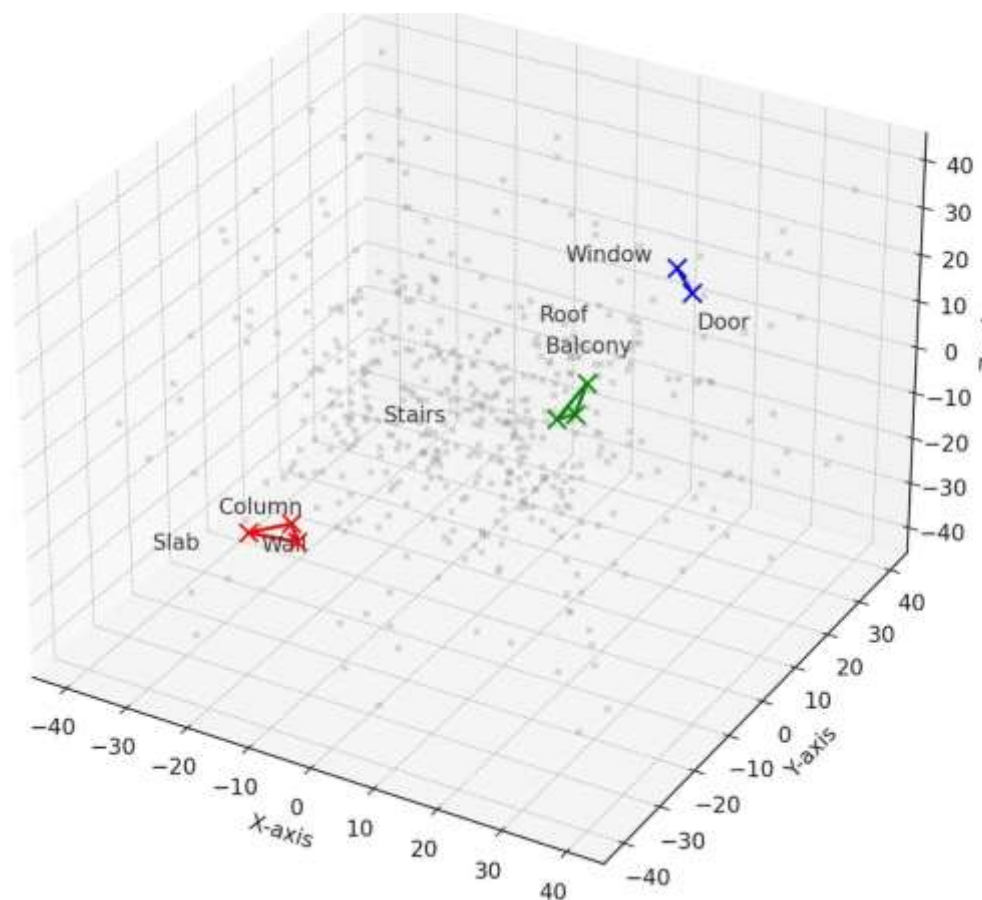
	minX	maxX	minY	maxY	minZ	maxZ	Width	Height	Depth
Column	-15	-5	-25	-15	0	10	10	10	20
Stairs	-5	5	-15	-5	0	10	10	10	10
Door	5	15	5	15	0	10	10	10	10
Window	25	35	-35	-25	10	30	10	20	20
Balcony	15	25	-5	5	20	40	10	20	20

איור 8.21 המידע על קואורדינטות של אלמנטים בתבנית גבול ומיקומם במודל הפרויקט מהווה אנלוגיה לבסיס נתונים וקטורי.

הצגה כזו של נתונים מקלה על מגוון משימות, כולל בדיקות חפיפות בין אובייקטים, תכנון הפצה מרחבית של רכיבי יכול לשמש גשר בין מודלים תלת-ממדיים מורכבים לבין Bounding Box-הבניין וביצוע חישובים אוטומטיים. ה. בסיסי נתונים וקטוריים מסורתיים, ומאפשר לנצל ביעילות את היתרונות של שני הגישות במידול אדריכלי והנדסי.

קופסת גבול היא "וקטורזציה של גיאומטריה", בעוד שהאמבדינג (שיטה להמיר משהו מופשט) היא וקטורזציה של משמעות. "שני הגישות מאפשרות לעבור מחיפוש ידני לחיפוש אינטלקטואלי, בין אם" מדובר באובייקטים תלת-ממדיים במודל פרויקט או במונחים בטקסט.

(k-חיפוש אובייקטים בפרויקט) למשל, מצא את כל החלונות ברוחב >1.5 מ ("דומה לחיפוש השכנים הקרובים ביותר בבסיס נתונים וקטורי, שבו הקריטריונים מגדירים "אזור" במרחב התכונות). נרחיב על חיפוש השכנים הקרובים (k-NN פרמטרים bounding box-אם נוסיף לאטריבוטים של ה. (8.22 איור) (בחלק הבא על למידת מכונה k-NN ביותר נוספים) חומר, משקל, זמן ייצור, (אז הטבלה תהפוך לווקטור בעל ממדים גבוהים, שבו כל אטריבוט הוא מימד חדש. זה כבר קרוב לבסיסי נתונים וקטוריים מודרניים, שבהם הממדים נמדדים במאות או אלפים) למשל, אמבדינג מרשתות).- (עצביות



איור 8.22 חיפוש אובייקטים בפרויקט באמצעות בסיסי נתונים וקטוריים.

רלוונטית לא רק לאובייקטים גיאומטריים, אלא גם לניתוח טקסטים Bounding Box-הגישה שבה נעשה שימוש ב דיוק כפי שניתן לקבץ אובייקטים. (NLP) ושפה. ייצוגים וקטוריים של נתונים כבר בשימוש פעיל בעיבוד שפה טבעית -בפרויקט בנייה לפי קרבה מרחבית) איור 8.22, (ניתן לנתח מילים בטקסט לפי קרבה סמנטית והקשרית

לדוגמה, המילים "אדריכל", "בנייה", "תכנון" במרחב הווקטורי יהיו קרובות זו לזו, מכיוון שיש להן משמעות דומה: במודלים של שפה גדולה, מנגנון זה מאפשר אוטומטית, ללא צורך בקטלוג ידני

■ לקבוע את נושא הטקסט

■ לבצע חיפוש סמנטי על פי תוכן המסמכים

■ ליצור אנוטציות אוטומטיות וסיכומים של טקסט

■ למצוא סינונימים ומונחים קשורים

מסייע לנתח Bounding Box-מסדי נתונים וקטוריים מאפשרים לנתח טקסט ולמצוא בו מונחים קשורים בדיוק כמו ש של רכיבי הפרויקט מסייעת להבין שהייצוג Bounding Box אובייקטים מרחביים במודלים תלת-ממדיים. דוגמת אלא דרך טבעית לארגון נתונים לפתרון בעיות מעשיות, בין אם, ML-הווקטורי אינו מושג "מלאכותי" לחלוטין מעולם ה או בתמונות סמנטיות קרובות במאגר CAD מדובר בחיפוש עמודות בפרויקט

למומחים העובדים עם מסדי נתונים כדאי לשים לב לאחסון וקטורי. התפשטותם מצביעה על שלב חדש, מתחילת להתמזג AI בהתפתחות מסדי הנתונים, שבו מערכות רלציוניות קלאסיות וטכנולוגיות ממוקדות, וליצור פתרונות היברידיים לעתיד.

מורכבים וגדולים ישתמשו במסדי נתונים מיוחדים לחיפוש וקטורי. במקביל, אלו AI משתמשים המפתחים יישומי בודדות לשילוב ביישומים קיימים, ככל הנראה יבחרו את האפשרויות המובנות לחיפוש AI שזקוקים רק לפונקציות (PostgreSQL, Redis) וקטורי במסדי הנתונים שכבר בשימוש.

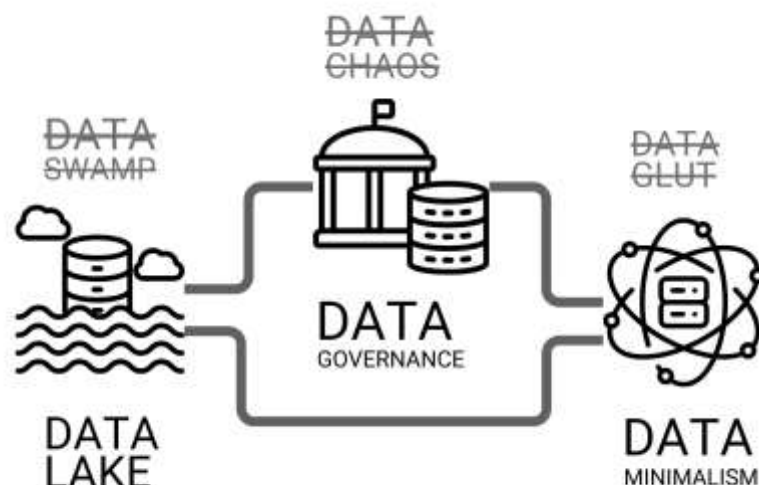
מסדי נתונים וקטוריים ואחרים מציעות גישות שונות, DWH, Data Lake, CDE, PMIS, אף על פי שמערכות כמו לאחסון ולניהול נתונים, היעילות שלהן נקבעת לא רק על ידי הארכיטקטורה, אלא גם על ידי האופן שבו הנתונים עצמם מאורגנים ומנוהלים. אפילו בשימוש בפתרונות מודרניים – בין אם מדובר במסדי נתונים וקטוריים, מסדי נתונים חוסר כללים ברורים לניהול, ארגון ועדכון נתונים עלול – Data Lake רלציוניים קלאסיים או מאגרי נתונים מסוג להוביל לאותן קשיים שבהם נתקלות משתמשים העובדים עם קבצים מבודדים ונתונים בפורמטים שונים.

אפילו הפתרונות החזקים ביותר עלולים להפוך לערימות מידע, (Data Governance) ללא ניהול נתונים מתוכנן כדי להימנע מכך, חברות צריכות לא (Data Swamp) כאוטיות ולא מאורגנות, והופכים אגמים של נתונים לביצות (Data Minimalism), רק לבחור את הארכיטקטורה המתאימה לאחסון, אלא גם ליישם אסטרטגיות מינימליזם נתונים. ניהול גישה ובקרת איכות, מה שמאפשר להפוך את הנתונים לכלי יעיל לקבלת החלטות.

ממשלת נתונים, מינימליזם נתונים וביצת נתונים

ומניעת (Data Minimalism) מינימליזם נתונים, (Data Governance) הבנה ויישום של מושגי ניהול נתונים. הם גורמים מרכזיים בניהול מוצלח של מאגרי נתונים והבטחת ערכם לעסק (ריש (Data Swamp) יצירת ביצת נתונים 8.23).-

מהפרויקטים בתחום הנתונים הגדולים נכשלו, ואחת הסיבות 85%,artner (2017) מחקר של ג המרכזיות לכך היא ניהול לקוי של איכות הנתונים וניהולם.



8.23. Data Governance ו-Data Minimalism אחד מהאספקטים המרכזיים של ניהול נתונים הם 8.23. ריש

הוא מרכיב בסיסי בניהול נתונים, המבטיח שימוש נאות ויעיל בנתונים בכל (Data Governance) ניהול נתונים תהליכי העסקים. מדובר לא רק בהקניית כללים ונהלים, אלא גם בהבטחת זמינות, אמינות וביטחון של הנתונים

■ הגדרה וסיווג נתונים: הגדרה ברורה וסיווג של ישויות מאפשרים לארגונים להבין אילו ישויות נדרשות בחברה ולקבוע כיצד יש להשתמש בהן

■ זכויות גישה וניהול: פיתוח מדיניות ונהלים לגישה לנתונים ולניהולם מבטיחים שרק משתמשים מורשים יוכלו לגשת לנתונים מסוימים

■ הגנה על נתונים מפני איומים חיצוניים: הגנה על נתונים מפני איומים חיצוניים היא אחד ההיבטים המרכזיים בניהול נתונים. זה כולל לא רק אמצעים טכניים, אלא גם הכשרת עובדים בעקרונות אבטחת מידע

זו גישה לצמצום הנתונים לערכים ולמשמעותיים ביותר בעת יצירת - (Data Minimalism) מינימליזם נתונים - (תכונות וישויות) איור 8.24, (מה שמאפשר להפחית הוצאות ולהגביר את היעילות בשימוש בנתונים

■ פישוט תהליך קבלת ההחלטות: צמצום מספר האובייקטים ותכונותיהם לערכים המשמעותיים ביותר מפשט את תהליך קבלת ההחלטות על ידי צמצום הזמן והמשאבים הנדרשים לניתוח ועיבוד נתונים

■ התמקדות בחשוב: בחירת הישויות והתכונות הרלוונטיות ביותר מאפשרת להתמקד במידע שחשוב באמת לעסק תוך הסרת רעש ונתונים מיותרים

■ חלוקת משאבים יעילה: מינימליזציה של נתונים מאפשרת חלוקה יעילה יותר של משאבים, מפחיתה את העלויות של אחסון ועיבוד נתונים, ומעלה את איכותם וביטחונם

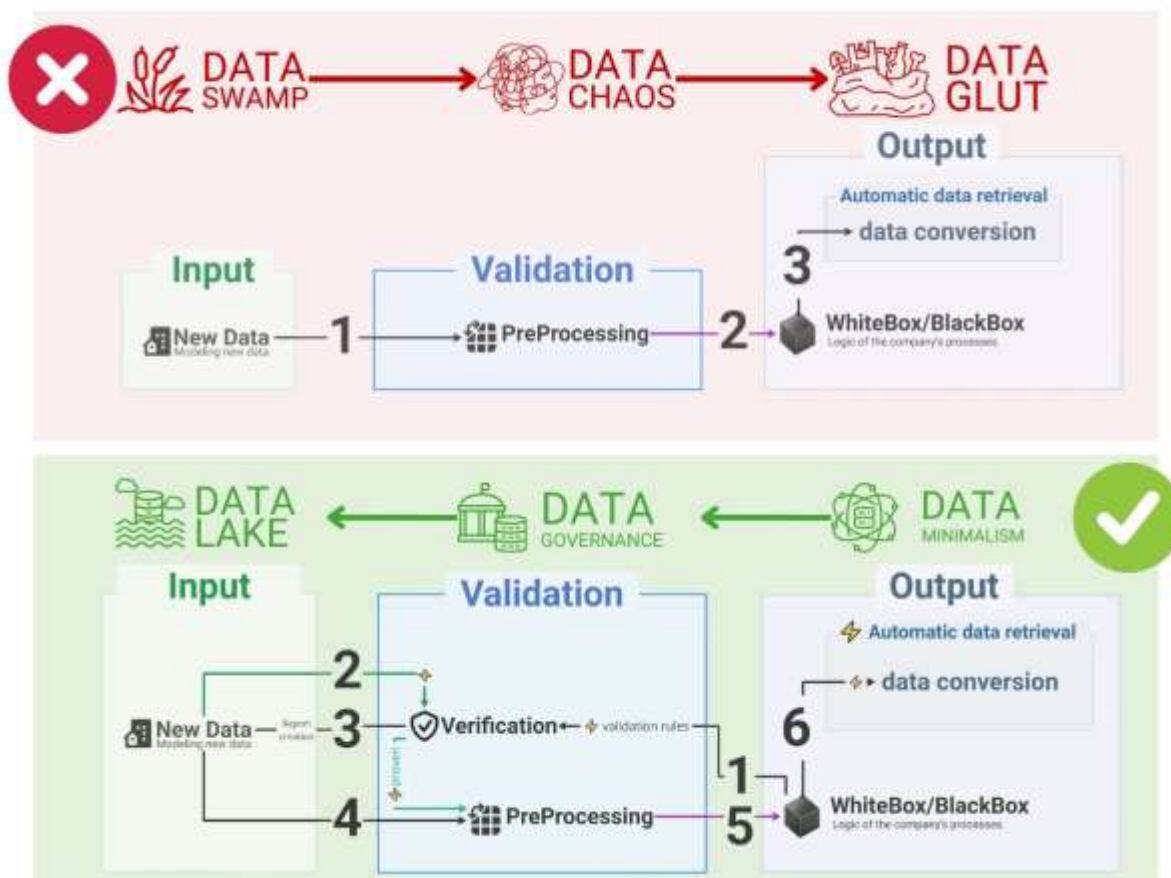
הלוגיקה של עבודה עם נתונים צריכה להתחיל לא מהיצירה שלהם (כשלעצמה) איור 8.24, (אלא מהבנת תרחישי השימוש העתידיים בנתונים הללו עוד לפני תחילת תהליך הגרציה. גישה זו מאפשרת לקבוע מראש את הדרישות המינימליות הנדרשות לתכונות, סוגיהן וערכי הגבול שלהן. דרישות אלו מהוות את הבסיס ליצירת ישויות נכונות ועמידות במודל המידע. חשיבה מוקדמת על מטרות ושיטות השימוש בנתונים תורמת ליצירת מבנה נגיש לניתוח. פרטים נוספים על גישות למידול נתונים ברמות קונספטואליות, לוגיות ופיזיות דיברנו בפרק "מידול נתונים: מודל קונספטואלי, לוגי ופיזי".

בתהליכי עסקים מסורתיים של חברות בנייה, עיבוד הנתונים מזכיר לעיתים קרובות השלכת נתונים בביצה, שבה תחילה נוצרות הנתונים, ולאחר מכן המומחים מנסים לשלב אותם במערכות ובכלים אחרים.

זו תוצאה של איסוף ואחסון נתונים ללא שליטה, ללא ארגון, מבנה וניהול - (Data Swamp) ביצת נתונים נאותים, מה שגורם לכך שהנתונים הופכים לבלתי מובנים, קשים לשימוש וללא ערך.

כיצד למנוע מהזרם של מידע להפוך לביצה

- ניהול מבנה הנתונים: הבטחת סדר וקטלוג של הנתונים מסייעת למנוע את "ביצת הנתונים", מה שהופך אותם למאורגנים ונגישים בקלות.
- הבנה ופירוש של נתונים: תיאור ברור של מקור הנתונים, השינויים שלהם ומשמעותם מבטיח שהנתונים יובנו ויפורשו כראוי.
- שמירה על איכות הנתונים: תחזוקה שוטפת וניקוי של הנתונים מסייעים לשמור על איכותם, רלוונטיותם וערכם לתהליכים אנליטיים ועסקיים.



איור 8.24 כדי למנוע אי סדר באחסון הנתונים, יש להתחיל את תהליך יצירת הנתונים מאיסוף דרישות לתכונות.

על ידי שילוב עקרונות ניהול הנתונים ומינימליזם הנתונים בתהליכי ניהול הנתונים, כמו גם מניעת הפיכת מאגרי הנתונים לביצת נתונים, ארגונים יכולים למקסם את הפוטנציאל של הנתונים שלהם.

השלב הבא בהתפתחות העבודה עם נתונים לאחר פתרון בעיות ניהול ומינימליזם הוא הסטנדרטיזציה של עיבוד אוטומטי הבטחת איכות והטמעת שיטות שהופכות את הנתונים לנוחים לניתוח, טרנספורמציה וקבלת החלטות. זה בדיוק מה שהופכות לכלים חשובים עבור חברות העובדות עם כמויות VectorOps ו-DataOps שמטפלות המתודולוגיות גדולות של מידע ולמידת מכונה.

דאטה אופס ווקטור אופס: סטנדרטים חדשים לעבודה עם נתונים

מאפשרת להבטיח את הדיוק, העקביות והתנועה DataOps אם ניהול נתונים אחראי על שליטה וארגון הנתונים, אז הבלתי פוסקת שלהם בתוך החברה. זה קריטי במיוחד עבור מספר מקרים עסקיים בתחום הבנייה, שם הנתונים נים באופן מתמשך ודורשים עיבוד בזמן. לדוגמה, במצבים שבהם מודלים מידעיים של בניינים, דרישות פרויקט ודוחות יכול להיות מפתח DataOps אנליטיים צריכים להיות מסונכרנים בין מערכות שונות במהלך יום עבודה אחד, תפקיד היא מאפשרת לבנות תהליכי עיבוד נתונים יציבים וחוזרים, מפחיתה את הסיכונים לעיכובים ואובדן עדכניות המידע.

ניהול ברמת ניהול נתונים אינו מספיק בפני עצמו - חשוב שהנתונים לא רק יישמרו, אלא גם ייעשה בהם שימוש פעיל יוצאת למרכז הבמה - מתודולוגיה הממוקדת באוטומציה, אינטגרציה והבטחת זרימה DataOps בפעולות היומיות. כאן מתמשכת של נתונים.

מתמקדת בשיפור שיתוף הפעולה, האינטגרציה ואוטומציה של זרמי נתונים בארגונים DataOps תורמת לדיוק, עקביות וזמינות הנתונים, מה שחשוב מאוד עבור יישומים DataOps הטמעת פרקטיקות הממוקדים בנתונים.

לאורקסטרציה של - (7.44 איור) Apache Airflow הם DataOps הכלים המרכזיים במערכת האקולוגית של לניתוב והמרת זרמי נתונים. יחד, הטכנולוגיות הללו מאפשרות לבנות - (7.45 איור) Apache NiFi-תהליכי עבודה, ו צינורות נתונים גמישים, אמינים ומדריכים, המבטיחים עיבוד אוטומטי, שליטה ואינטגרציה של מידע בין מערכות) פרטים בתהליכי הבנייה, חשוב לקחת בחשבון DataOps אוטומטי. ("כאשר מטמיעים את גישת ETL נוספים בפרק "צינור --ארבעה היבטים בסיסיים

1. אנשים וכלים חשובים יותר מהנתונים: מחסני נתונים מנותקים יכולים להיחשב כבעיה המרכזית, אך במציאות המצב מורכב יותר. מעבר לפיצול הנתונים, תפקיד משמעותי משחקת בידוד הצוותים והנפרדות של הכלים שבהם הם משתמשים. בתחום הבנייה עובדים עם נתונים אנשי מקצוע ממגוון תחומים: מהנדסי נתונים וויזואליזציה, כמו גם מומחים לניהול פרויקטים ואיכות. לכל אחד מהם יש שיטות עבודה BI ואנליסטים, צוותי משלו, ולכן גורם חשוב הוא יצירת מערכת אקולוגית שבה הנתונים מועברים בחופשיות בין המשתתפים ומספקים גרסה אחת, מסונכרנת של המידע.

2. אוטומציה של בדיקות וזיהוי שגיאות: נתוני בנייה תמיד מכילים שגיאות, בין אם מדובר באי דיוקים במודלים שגיאות בחישובים או במפרטים מיושנים. בדיקות נתונים סדירות ותיקון שגיאות חוזרות מאפשרות לשפר

יש ליישם מגננוני בקרה ואימות אוטומטיים, אשר עוקבים DataOps באופן משמעותי את איכותם. במסגרת אחרי נכונות הנתונים, מנתחים שגיאות ומזהים תבניות, כמו גם מתעדים ומתקנים תקלות מערכת בכל מחזור עבודה. ככל שדרגת האוטומציה של הבדיקה גבוהה יותר, כך איכות הנתונים הכוללת גבוהה יותר והסיכוי לשגיאות בשלב הסופי נמוך יותר.

3. נתונים צריכים להיבדק כמו קוד תוכנה: רוב היישומים בתחום הבנייה מבוססים על עיבוד נתונים, אך הבקרה עליהם לעיתים קרובות נשארת בתפקידים משניים. אם מודלים של למידת מכונה מאומנים על נתונים לא הנתונים צריכים לעבור בדיקה, DataOps מדויקים, זה מוביל לחזויות שגויות ולהפסדים כספיים. במסגרת קפדנית כמו קוד תוכנה: בדיקות לוגיות, בדיקות עומס, הערכת התנהגות המודלים כאשר ערכי הקלט משתנים רק נתונים מאומנים ואמינים יכולים לשמש כבסיס לקבלת החלטות ניהוליות.

4. נראות הנתונים ללא פגיעה בביצועים: ניטור נתונים הוא לא רק איסוף מדדים, אלא כלי ניהולי אסטרטגי לאיכות יעבוד ביעילות, הנראות צריכה להיות משולבת בכל שלב בעבודה עם נתונים, מעיצוב ועד DataOps-כדי ש הפעלה. חשוב שהניטור לא יאט את המערכת. בפרויקטי בנייה, קריטי לא רק לאסוף נתונים, אלא גם לעשות זאת כך שעבודת המומחים (כגון המהנדסים) המייצרים את הנתונים לא תופר. האיזון הזה מאפשר לשלוט באיכות הנתונים מבלי לפגוע בביצועים.

חברות בנייה, DataOps הוא לא נטל נוסף על המומחים לנתונים, אלא הבסיס לעבודה שלהם. על ידי יישום DataOps יכולות לעבור מניהול כאוטי לניהול יעיל של מערכת אקולוגית, שבה הנתונים עובדים עבור העסק.

המתמקד בעיבוד, אחסון וניתוח נתונים וקטוריים רב-ממדיים, DataOps הוא השלב הבא בהתפתחות VectorOps (שנדונו בפרק הקודם). (זה במיוחד רלוונטי בתחומים כמו דיגיטליים כפולים, מודלים של רשתות עצביות וחיפוש סמנטי מתבסס על בסיסי נתונים וקטוריים, המאפשרים אחסון, אינדוקס VectorOps. שמתחילים לחדור לתחום הבנייה וחיפוש יעיל של ייצוגים רב-ממדיים של אובייקטים.

המתמקד בעיבוד, ניתוח ושימוש בנתונים וקטוריים, DataOps הוא השלב הבא לאחר VectorOps מתרכז בניהול VectorOps, המתמקד בזרימה, עקביות ואיכות הנתונים, DataOps-בבנייה. בניגוד ל ייצוגים רב-ממדיים של אובייקטים, הנדרשים ללמידת מכונה.

מאפשרת להשיג תיאור מדויק יותר של אובייקטים, מה שקריטי לדיגיטליים VectorOps, בניגוד לגישות המסורתיות שהוסבו לפורמט וקטורי. יישום משולב של CAD כפולים, מערכות עיצוב גרפיטיות וזיהוי אוטומטי של שגיאות בנתוני יוצר בסיס חזק לעבודה אוטומטית ומדרגת עם כמויות גדולות של מידע – מטבלאות VectorOps ו-DataOps קלאסיות ועד מודלים מרחביים עשירים סמנטית.

צעדים נוספים: מאחסון כאוטי לאחסונים מובנים

גישות מסורתיות לאחסון נתוני בנייה לעיתים קרובות מובילות ליצירת "סילונים של מידע" מנותקים, שבהם מידע חשוב אגם, (Data Warehouse) אינו נגיש לניתוח ולקבלת החלטות. מושגים מודרניים לאחסון נתונים, כמו מחסן נתונים וההיברידים שלהם, מאפשרים לאחד מידע מפוזר ולהפוך אותו לנגיש באופן מרכזי לעיבוד נתונים (Data Lake) נתונים בזמן אמת ולניתוח עסקי. חשוב לא רק לבחור את הארכיטקטורה המתאימה לאחסון, אלא גם ליישם עקרונות של ניהול כדי למנוע מהאחסונים להפוך ל"ביצות", (Data Minimalism) ומינימליזם נתוני (Data Governance) נתונים

(Data Swamp) נתונים "בלתי נשלטות".

לסיכום החלק הזה, כדאי להדגיש את הצעדים המעשיים העיקריים שיכולים לעזור ליישם את המושגים שנדונו במשימות היומיומיות שלכם:

■ **בחרו פורמטים יעילים לאחסון נתונים**

☐ לאחסון כמויות גדולות (Apache Parquet, ORC) לפורמטים יעילים יותר XLSX ו-CSV-עברו מ של נתונים

☐ יישמו מערכת גרסאות לנתונים כדי לעקוב אחרי שינויים

☐ השתמשו במטא-נתונים כדי לתאר את המבנה ואת המקור של המידע

■ **צרו ארכיטקטורת נתונים אחידה לחברה**

☐ ואגם נתונים. בחרו את זו שמתאימה RDBMS, DWH: השוו בין ארכיטקטורות שונות לאחסון נתונים ביותר לצרכים שלכם מבחינת יכולת הרחבה, אינטגרציה של מקורות ועיבוד אנליטי

☐ ממקורות שונים עבור המשימות שלכם (ETL) תכננו מפה של תהליכי חילוף, טעינה ושינוי נתונים כדי להציג בצורה ברורה את Draw.io או Miro, Lucidchart השתמשו בכלי ויזואליזציה כמו הצעדים המרכזיים ואת נקודות האינטגרציה

■ **יישמו פרקטיקות של ניהול נתונים ומינימליזם נתוני**

☐ פעלו לפי גישת מינימליזם נתוני - אחסנו ועבדו רק את מה שיש לו ערך אמיתי

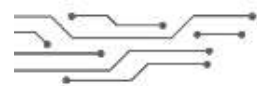
☐ יישמו את עקרונות ניהול הנתונים - קבעו אחריות על הנתונים, ודאו איכות ושקיפות

☐ DataOps, VectorOps למדו יותר על מדיניות ניהול הנתונים ועל מושגי

☐ DataOps קבעו קריטריונים לאיכות הנתונים ולנהלים לבדוק אותם במסגרת

אחסון נתונים מאורגן היטב יוצר בסיס למרכז תהליכי הניתוח של החברה. המעבר מאיסוף קבצים כאוטי לאחסונים מובנים מאפשר להפוך את המידע לנכס אסטרטגי, המסייע בקבלת החלטות מושכלות ובשיפור היעילות של תהליכים עסקיים.

לאחר שההליכים של איסוף, שינוי, ניתוח ואחסון נתונים בצורה מובנית אוטומטיים ומוסדרים, השלב הבא של (Big Data) טרנספורמציה דיגיטלית הוא עבודה מלאה עם נתונים גדולים.





IX חלק

נתונים גדולים, למידת מכונה וחיזויים

החלק התשיעי מוקדש לנתונים גדולים, למידת מכונה ולחיזוי בתעשיית הבנייה. כאן נדונה המעבר מקבלת החלטות אינטואיטיבית לניתוח אובייקטיבי, המבוסס על נתונים היסטוריים בדוגמאות מעשיות מוצג ניתוח נתונים גדולים בבנייה - מהבנת מערך נתוני ההיתרים לבנייה עם מיליוני רכיבים. תשומת לב מיוחדת מוקדשת CAD בסן פרנסיסקו ועד עיבוד פרויקטי לשיטות למידת מכונה לחיזוי עלויות ולוחות זמנים של פרויקטים בבנייה, עם ניתוח מפורט של שכנים הקרובים. מוצג כיצד נתונים מובנים הופכים-k-אלגוריתמים של רגרסיה ליניארית ו לבסיס למודלים חיזוי, המאפשרים להעריך סיכונים, לייעל משאבים ולשפר את היעילות בניהול פרויקטים. החלק כולל גם המלצות לבחירת דגימות נתונים מייצגות ומסביר מדוע לניתוח יעיל לא תמיד נדרשים כמויות עצומות של מידע.

פרק 9.1. נתונים גדולים וניתוחם

נתונים גדולים בבנייה: מהאינטואיציה לחיזוי

המונח "נתונים גדולים" אינו בעל הגדרה מדויקת. במקור, המושג הזה הופיע כאשר כמויות המידע החלו לעלות על יכולות השיטות המסורתיות לעיבודו. כיום, הכמויות והמורכבות של הנתונים בהרבה תחומים, כולל בנייה, גדלו כל כך עד שהם אינם יכולים להתאחסן בזיכרון המקומי של מחשבים ודורשים שימוש בטכנולוגיות חדשות לעיבודם.

מהות העבודה עם נתונים גדולים אינה רק באחסון ובעיבוד, אלא גם ביכולת החיזוי. בתעשיית הבנייה, נתונים גדולים פותחים את הדרך מהחלטות אינטואיטיביות, המבוססות על פרשנות סובייקטיבית של טבלאות וויזואליזציות (כפי שנדון קודם, (לחיזויים מבוססים, הנתמכים על ידי תצפיות אמיתיות וסטטיסטיקה

בניגוד לדעה הרווחת, המטרה של עבודה עם נתונים גדולים אינה לגרום למכונה לחשוב כמו אדם, אלא ליישם מודלים מתמטיים ואלגוריתמים לניתוח מערכי נתונים במטרה לחשוף תבניות, לחזות אירועים ולייעל תהליכים.

אינם עולם קר של אלגוריתמים, חסר השפעה אנושית. להפך, נתונים גדולים (Big Data) נתונים גדולים פועלים בשיתוף פעולה עם האינסטינקטים, השגיאות והיצירתיות שלנו. דווקא חוסר המושלמות של החשיבה האנושית מאפשר למצוא פתרונות לא שגרתיים ולעשות פריצות דרך.

IT-עם התפתחות הטכנולוגיות הדיגיטליות, החלו להיעשות שימושים פעילים בשיטות עיבוד נתונים שהגיעו מעולם הנתונים מובנים ולא מובנים יכולים להתאחד, מה, Apache Parquet ו-Pandas בתעשיית הבנייה. בזכות כלים כמו CAD שמקל על הגישה למידע ומפחית את ההפסדים בניתוח, בעוד שמערכי נתונים גדולים מתוך מסמכים או פרויקטי --מאפשרים לאסוף, לנתח ולחזות נתונים בכל שלב של מחזור חיי הפרויקט (9.212 איור - 9.210 איור)

נתונים גדולים משפיעים באופן מהותי על תעשיית הבנייה, ומשפיעים עליה פוטנציאלית במגוון רחב של היבטים. השימוש בטכנולוגיות נתונים גדולים מביא לתוצאות במספר כיוונים מרכזיים, כולל לדוגמה את הבאים:

- ניתוח פוטנציאל השקעה - חיזוי תשואות וזמני החזר השקעה של פרויקטים על בסיס נתונים של פרויקטים קודמים.
- תחזוקה חיזויית - זיהוי תקלות אפשריות בצידוד לפני שהן מתרחשות בפועל, מה שמפחית את זמן ההשבתה.
- אופטימיזציה של שרשראות אספקה - חיזוי תקלות ושיפור היעילות הלוגיסטית.
- ניתוח יעילות אנרגטית - סיוע בתכנון מבנים עם רמות צריכת אנרגיה נמוכות.
- ניטור בטיחות - שימוש בחיישנים ובמכשירים נישאים למעקב אחר התנאים באתר הבנייה.
- בקרת איכות - פיקוח על עמידה בתקנים טכנולוגיים בזמן אמת.
- ניהול משאבי אנוש - ניתוח ביצועים וחיזוי צרכי כוח אדם.

קשה למצוא תחום בבנייה שבו אנליטיקה של נתונים וחיזויים אינם נדרשים. היתרון המרכזי של אלגוריתמים לחיזוי הוא היכולת שלהם ללמוד באופן עצמאי ולהשתפר באופן מתמיד ככל שמצטברים נתונים.

בעתיד הקרוב, אינטליגנציה מלאכותית לא רק תסייע לבונים, אלא גם תקבל החלטות מפתח - מתהליכי תכנון ועד שאלות של תפעול מבנים.

פרטים נוספים על איך מתבצעים החיזויים וכיצד מודלים לומדים משמשים, יידונו בחלק הבא של הספר: "למידת מכונה וחיזויים".

כדי לעבור לעבודה מלאה עם נתונים גדולים, יש צורך בשינוי הגישה לאנליטיקה. אם במערכות קלאסיות, שהוזכרו עד כה, הדגש היה על קשרים סיבתיים, אז באנליטיקה של נתונים גדולים הדגש עובר לחיפוש אחר תבניות סטטיסטיות וקורלציות, המאפשרות לחשוף קשרים סמויים ולחזות התנהגות של אובייקטים גם ללא הבנה מלאה של כל הגורמים.

שאלת הכדאיות של נתונים גדולים: מתודולוגיה, סטטיסטיקה ודגימת נתונים

באופן מסורתי, הבנייה התבססה על השערות סובייקטיביות וניסיון אישי. מהנדסים הניחו - עם סבירות מסוימת - כיצד יתנהג חומר, אילו עומסים תוכל לעמוד המבנה וכמה זמן יימשך הפרויקט. השערות אלו נבדקו בפועל, לעיתים במחיר של זמן, משאבים וסיכונים עתידיים.

עם הופעת הנתונים הגדולים, הגישה משתנה באופן דרמטי: ההחלטות מתקבלות כעת לא על סמך תחושות אינטואיטיביות, אלא כתוצאה מניתוח של מסדי נתונים רחבים. הבנייה בהדרגה מפסיקה להיות אמנות של אינטואיציה והופכת למדע מדויק של חיזויים.

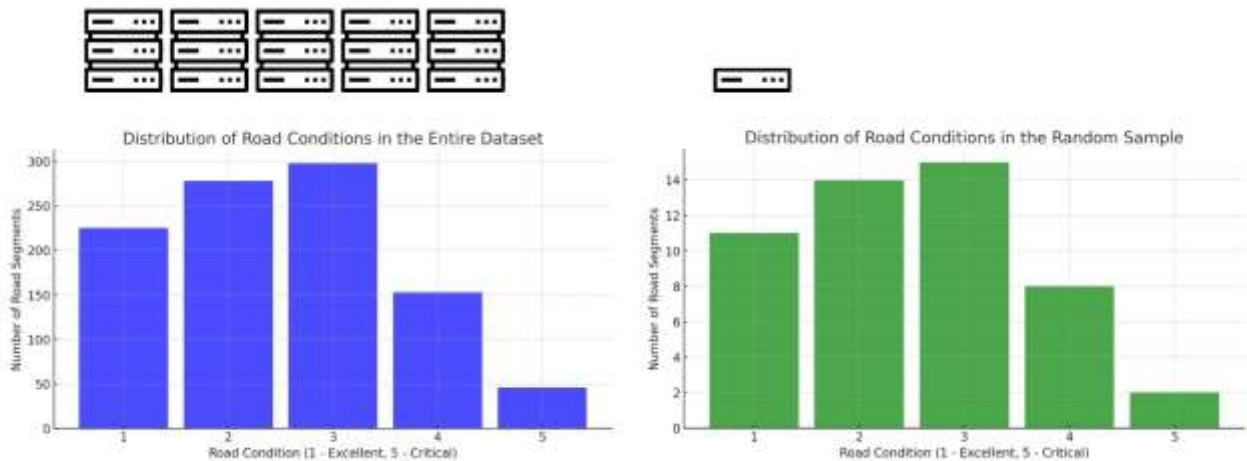
המעבר לרעיון השימוש בנתונים גדולים מעלה באופן בלתי נמנע שאלה חשובה: עד כמה קריטי הוא נפח הנתונים וכמה מידע באמת נדרש עבור אנליטיקה חיזוי אמינה? הדעה הרווחת ש"יותר נתונים" = יותר דיוק "מתבררת לא תמיד כנכונה מבחינה סטטיסטית.

כבר בשנת 1934, הסטטיסטיקאי ירז' ניימן הוכיח כי המפתח לדיוק במסקנות סטטיסטיות אינו תלוי כל כך בנפח הנתונים, אלא בייצוגיות ובאקראיות של הדגימה.

סורקים, IoT, זהו נושא רלוונטי במיוחד בתעשיית הבנייה, שבה כמויות גדולות של נתונים נאספות באמצעות חיישני "בגדלים שונים, מה שמגביר את הסיכון להיווצרות "אזורים עיוורים CAD מצלמות אבטחה, רחפנים ואפילו מודלים. חריגות ועיוותים בנתונים.

ג'יגה-בייט X נבחן דוגמה למעקב אחר מצב הכיסוי הכבישי. סט הנתונים המלא על כל הקטעים הכבישיים עשוי לתפוס ג'יגה- X/50 ולדרוש כעבור יום שלם לעיבוד. יחד עם זאת, דגימה אקראית, הכוללת רק כל קטע חמישים, תתפוס רק

ביט ותעובד בתוך חצי שעה, ובכך תספק דיוק דומה בחלק מהחישובים) איור 9.11).



איור 9.11 היסטוגרמות מצב הכיסי הכבישי: סט הנתונים המלא ודגימה אקראית מציגים תוצאות זהות.

לכן, גורם מפתח להצלחה בניתוח נתונים עשוי להיות לא גודלם, אלא ייצוגיות הדגימה ואיכות השיטות המיועדות לעיבוד המעבר לדגימה אקראית ולגישה יותר סלקטיבית דורש שינוי חשיבה בתעשיית הבנייה. היסטורית, חברות נצמדו ללוגיקה: "ככל שיש יותר נתונים - כך זה טוב יותר", מתוך אמונה שכיסוי כל המדדים האפשריים יבטיח דיוק מקסימלי. גישה זו מזכירה מיתוס פופולרי בניהול פרויקטים: "ככל שאשקול יותר מומחים, כך העבודה תהיה יעילה יותר". עם זאת, כמו במקרה של כוח אדם, איכות וכלים חשובים יותר מכמות. ללא התחשבות בקשרים (קורלציות) בין הנתונים או משתתפי הפרויקט, עלייה בנפח עשויה להוביל רק לרעש, עיוותים, חזרות וחריגות בלתי מוצדקות. בסופו של דבר, לעיתים קרובות מתברר שעדיף להחזיק בסט נתונים קטן אך איכותי, המסוגל לספק תחזיות יציבות ומבוססות, מאשר להסתמך על מידע מסיבי אך כאוטי, המכיל שלל אותות סותרים.

נפח נתונים מופרז לא רק שאינו מבטיח דיוק גבוה יותר, אלא עשוי להוביל למסקנות מעוותות - בשל נוכחות רעשים, תכנים עודפים, קורלציות חבויות ומידע שאינו רלוונטי. בתנאים כאלה, הסיכון להדרכה יתרה של מודלים עולה ואמינות התוצאות האנליטיות פוחתת.

בבנייה, האתגר המרכזי בעבודה עם נתונים גדולים הוא קביעת כמות ואיכות הנתונים האופטימליים. לדוגמה, במעקב אחר מצב מבני בטון, השימוש באלפי חיישנים ואיסוף מידע בכל דקה עשוי להעמיס על מערכת האחסון והניתוח. עם זאת, אם נבצע ניתוח קורלציה ונבחר 10% מהחיישנים המידע ביותר, ניתן להשיג דיוק תחזיות כמעט זהה, תוך שימוש במשאבים הרבה פחותים, לעיתים בעשרות ואף מאות פעמים.

השימוש במיעוט נתונים מפחית הן את נפח האחסון הנדרש והן את זמן העיבוד, מה שמפחית באופן משמעותי את העלויות של אחסון וניתוח נתונים ולעיתים קרובות הופך את הדגימה האקראית לפתרון אידיאלי עבור אנליטיקה תחזיתית, במיוחד בתנאים של פרויקטים תשתיתיים גדולים או בעבודה בזמן אמת. בסופו של דבר, היעילות של תהליכי הבנייה נקבעת לא על ידי נפח הנתונים שנאסף, אלא על ידי איכות הניתוח שלהם. ללא גישה ביקורתית וניתוח קפדני, נתונים עשויים להוביל למסקנות שגויות.

לאחר נפח נתונים מסוים, כל יחידת מידע חדשה מספקת תוצאה מועילה פחות ופחות. במקום איסוף אינסופי של מידע, חשוב להתמקד בייצוגיות שלו ובשיטות הניתוח (איור 9.22).

תופעה זו מתוארת היטב על ידי אלן ווליס, אשר מדגים את השימוש בשיטות סטטיסטיות באמצעות דוגמה של ניסוי שתי קונסטרוקציות חלופיות של פגזים של הצי האמריקאי.

מקבל 1 A תוך כדי ביצוע סדרת יריות זוגיות. בכל סיבוב, $(A-B)$ הצי האמריקאי ניסה שתי קונסטרוקציות חלופיות של פגזים ולהפך. הגישה הסטטיסטית הסטנדרטית מניחה ביצוע מספר, B או 0 בהתאם לכך אם ביצעו טובים יותר או גרועים יותר מ-1, מקבל 1 ביותר מ-53% מהמקרים A קבוע של ניסויים) למשל, 1000 (והגדרת המנצח על סמך התפלגות אחוזים) למשל, אם הוא נחשב לטוב יותר. (כאשר אלן ווליס דן בבעיה כזו עם הקפטן גארט ל. שיילר, הקפטן התנגד לכך שניסוי כזה, כפי שציטט אלן, עשוי להיות חסר תועלת. אם קצין ארטילריה חכם ומנוסה כמו שיילר היה במקום, הוא היה רואה לאחר כמה מאות יריות שהניסוי אינו צריך להסתיים או משום שהשיטה החדשה גרועה באופן ברור, או משום שהיא עולה על מה שציפו. – קבוצת מחקר סטטיסטית ממשלתית בארצות הברית באוניברסיטת קולומביה, בתקופת מלחמת העולם השנייה

עיקרון זה נמצא בשימוש נרחב בתחומים שונים. ברפואה, למשל, ניסויים קליניים של תרופות חדשות נערכים על דגימות אקראיות של מטופלים, מה שמאפשר להשיג תוצאות סטטיסטיות משמעותיות מבלי צורך לבדוק את התרופה על כל האוכלוסייה החיה על פני כדור הארץ. בכלכלה ובסוציולוגיה נערכים סקרים מייצגים, אשר משקפים את דעת הציבור מבלי צורך לשאול כל תושב במדינה.

בדומה לכך, מדינות וארגוני מחקר עורכים סקרים של קבוצות אוכלוסייה קטנות כדי להבין מגמות חברתיות כלליות, חברות בתחום הבנייה יכולות להשתמש בדגימות אקראיות של נתונים כדי לנטר וליצור תחזיות לניהול פרויקטים.

נתונים גדולים יכולים לשנות את הגישה למדעי החברה, אך אינם מחליפים את השכל הישר הסטטיסטי

Significance v. 9(4), 2012, "תומס לנדסל-וולפר", תחזית מצב הרוח של האומה כיום –

מבחינת חיסכון במשאבים, כאשר אוספים נתונים לצורך תחזיות עתידיות וקבלת החלטות, חשוב לשאול את השאלה האם יש טעם להוציא משאבים משמעותיים על איסוף ועיבוד כמויות עצומות של נתונים, כאשר ניתן להשתמש בערכת נתונים קטנה וזולה יותר שניתן להרחיב בהדרגה? היעילות של דגימות אקראיות מראה כי חברות יכולות לצמצם עלויות, בעשרות ואף אלפי פעמים על איסוף והכשרת מודלים, על ידי בחירת שיטות איסוף נתונים שאינן דורשות כיסוי כולל, אך מספקות דיוק וייצוגיות מספקים. גישה כזו מאפשרת גם לחברות קטנות להשיג תוצאות ברמה של תאגידים גדולים, תוך שימוש במשאבים ונתונים הרבה יותר קטנים, דבר שחשוב לחברות השואפות לייעל את ההוצאות ולהאיץ את תהליך קבלת ההחלטות המושכלות, תוך שימוש במשאבים קטנים. בפרקים הבאים נבחן דוגמאות של אנליטיקה ותחזיות על בסיס מערכות נתונים ציבוריות תוך שימוש בכלים של נתונים גדולים.

נתונים גדולים: ניתוח נתוני מערך של מיליון היתרי בנייה בסן פרנסיסקו

עבודה עם מערכות נתונים פתוחות מספקת הזדמנות ייחודית ליישם בפועל את העקרונות שנדונו בפרקים הקודמים, בחירת תכונות רציונליות, דגימה מייצגת, ויזואליזציה וניתוח ביקורתי. בפרק זה נבחן כיצד ניתן לחקור תופעות מורכבות

כגון פעילות בנייה בעיר גדולה, באמצעות נתונים פתוחים - במיוחד, מעל מיליון רשומות על היתרי בנייה בסן פרנסיסקו ממחלקת (CSV רשומות בשני מערכות נתונים בפורמט) נתונים ציבוריים על יותר ממיליון היתרי בנייה (איור 9.12) גולמית לניתוח לא רק של פעילות הבנייה בעיר, אלא CSV הבניינים של סן פרנסיסקו מאפשרים לנו להשתמש בטבלת גם לניתוח ביקורתי של מגמות והיסטוריה של תעשיית הבנייה בסן פרנסיסקו בעשורים האחרונים, משנת 1980 ועד 2019.

דוגמאות קוד שנעשה בהן שימוש ליצירת ויזואליזציות של מערכת הנתונים (איור 9.13- איור 9.18, כמו גם גרפים "ויזואליים עם קוד, הסברים והערות ניתן למצוא בפלטפורמת קגל בחיפוש "סן פרנסיסקו. מגזר הבנייה 1980–2019".

count

1.137695e+06

Public

Building Permits on or after January 1, 2013

June 13, 2020

2,237 Views

Public

Building Permits before January 1, 2013

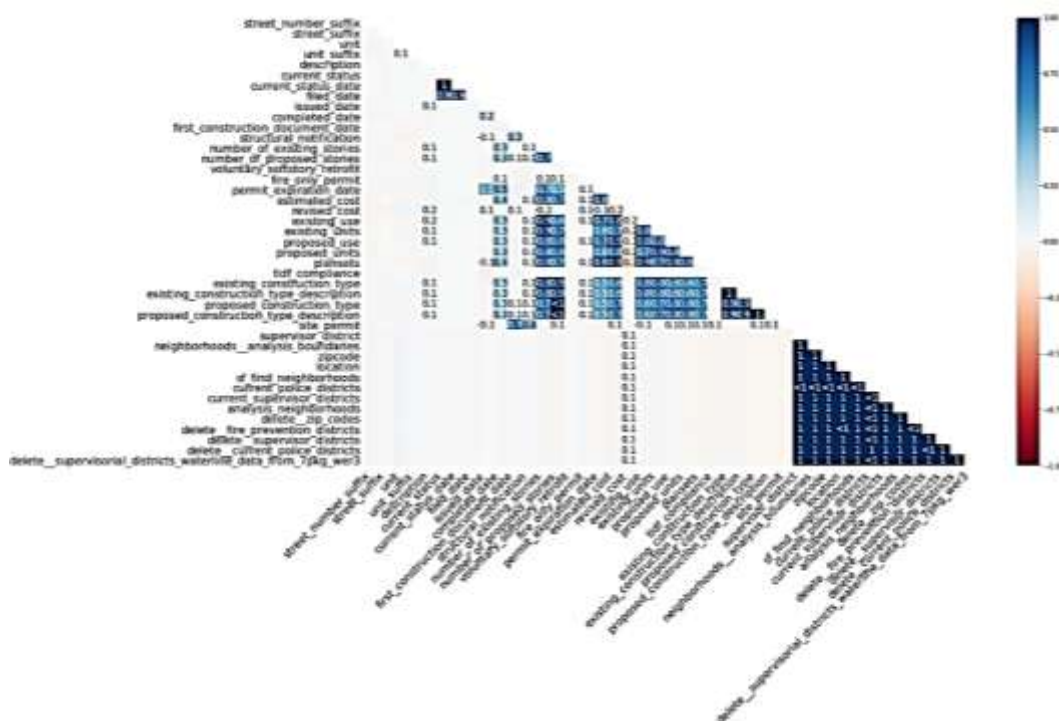
June 13, 2020

880 Views

permit_creation_date	description	current_status	current_status_date	filed_date	issued_date	completed_date
07/01/1998	repair stucco	complete	07/07/1998	07/01/1998	07/01/1998	07/07/1998
12/13/2004	reroofing	expired	01/24/2006	12/13/2004	12/13/2004	NaN
02/18/1992	install auto fire spks	complete	06/29/1992	02/18/1992	03/18/1992	06/29/1992

permit_number	permit_expiration_date	estimated_cost	revised_cost	existing_use	Zipcode	Location
362780	9812394	11/01/1990	780.0	NaN	1 family dwelling	94123.0 (37.7963488760498, -122.4322641443574)
570817	200412131233	06/13/2005	9000.0	9000.0	apartments	94127.0 (37.726258518008386, -122.4644245667462)
198411	9202396	09/18/1992	9000.0	NaN	apartments	94111.0 (37.79506002553974, -122.39593224461805)

איור 9.12 מערכות הנתונים מכילות מידע על היתרי בנייה שהונפקו עם תכונות שונות של האובייקטים.

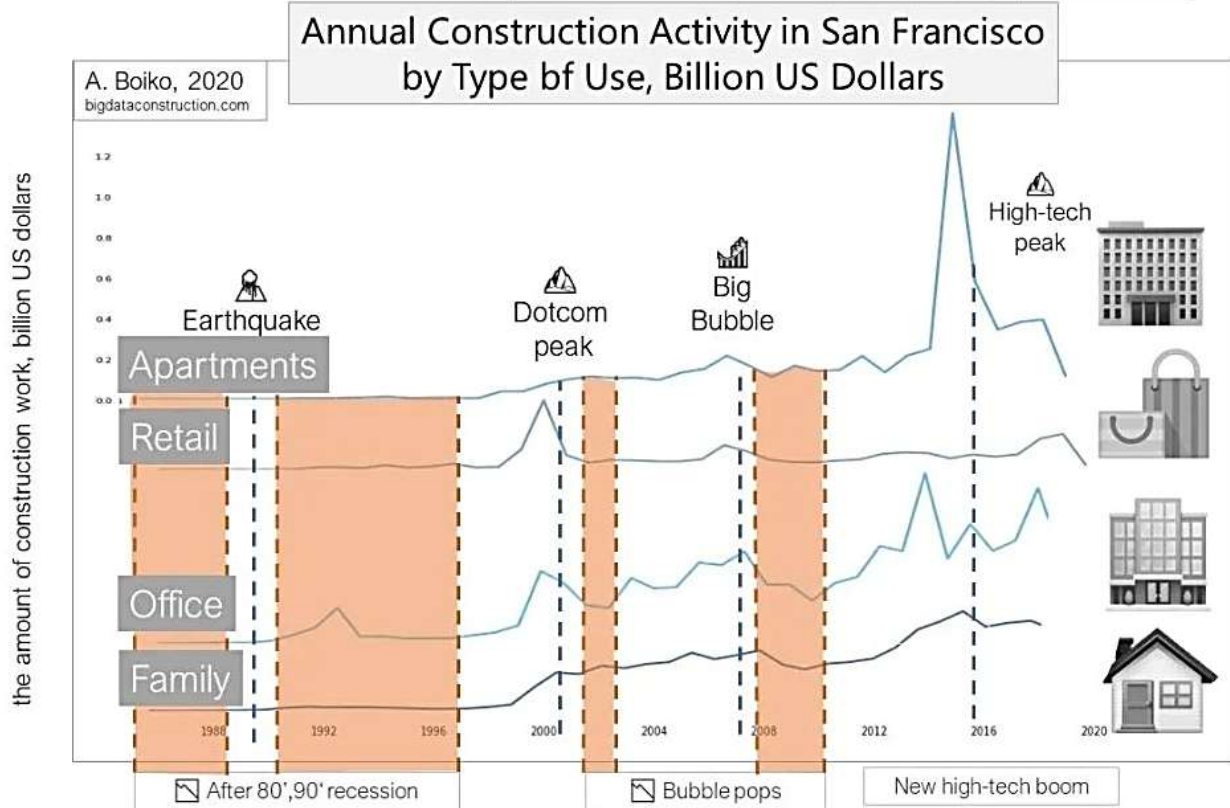
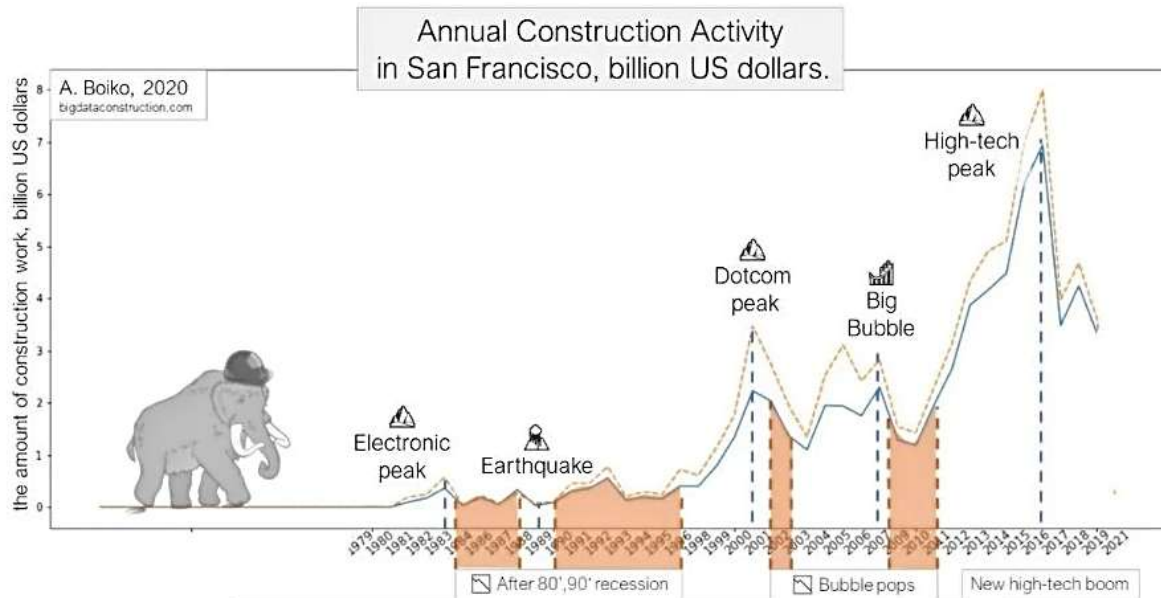


הממחישה את כל תכונות מערכת הנתונים ועוזרת לחשוף קשרים בין זוגות, (Pandas ו-Seaborn) איור 9.13 מפה חמה תכונות.

מהטבלה שסיפקה מחלקת הבניינים של סן פרנסיסקו (איור 9.12), (לא ניתן לראות מגמות או מסקנות. מספרים יבשים בצורה טבלאית אינם בסיס לקבלת החלטות. כדי להפוך את הנתונים לויזואליים וברורים, כפי שנדון בפירוט בפרקים על ויזואליזציה ETL" ויזואליזציה של נתונים, יש להציגם באמצעות ספריות שונות שנדונו בחלק השביעי של הספר בנושא "של תוצאות בצורה של גרפים

על עלות Python, 1,137,695 וספריות ויזואליזציה של Pandas DataFrame לאחר ניתוח הנתונים, באמצעות היתרים, ניתן להסיק כי פעילות הבנייה בסן פרנסיסקו קשורה קשר הדוק למחזורי כלכלה, במיוחד בתעשיית הטכנולוגיה - המתפתחת של עמק הסיליקון (איור 9.14)

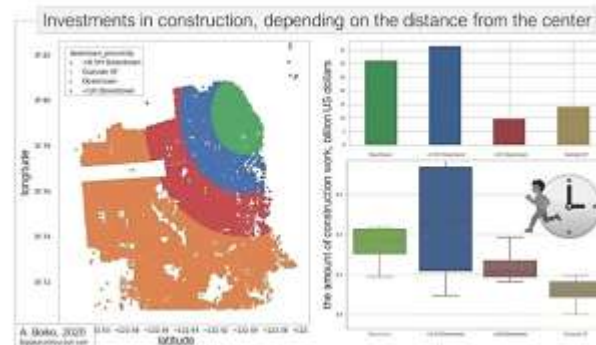
בומים כלכליים ושפלות משפיעים באופן משמעותי על מספר ועלות פרויקטי הבנייה. לדוגמה, השיא הראשון בפעילות והשיאים והשפלות, (Pandas ו-Matplotlib-נעשה שימוש ב) הבנייה חל על בום האלקטרוניקה באמצע שנות ה-80 שלאחר מכן היו קשורים לבועת הדוט-קום ולבום הטכנולוגי של השנים האחרונות.



איור 9.14 בתחום הנדל"ן של סן פרנסיסקו, השקעות מתואמות עם ההתפתחות הטכנולוגית של עמק הסיליקון.

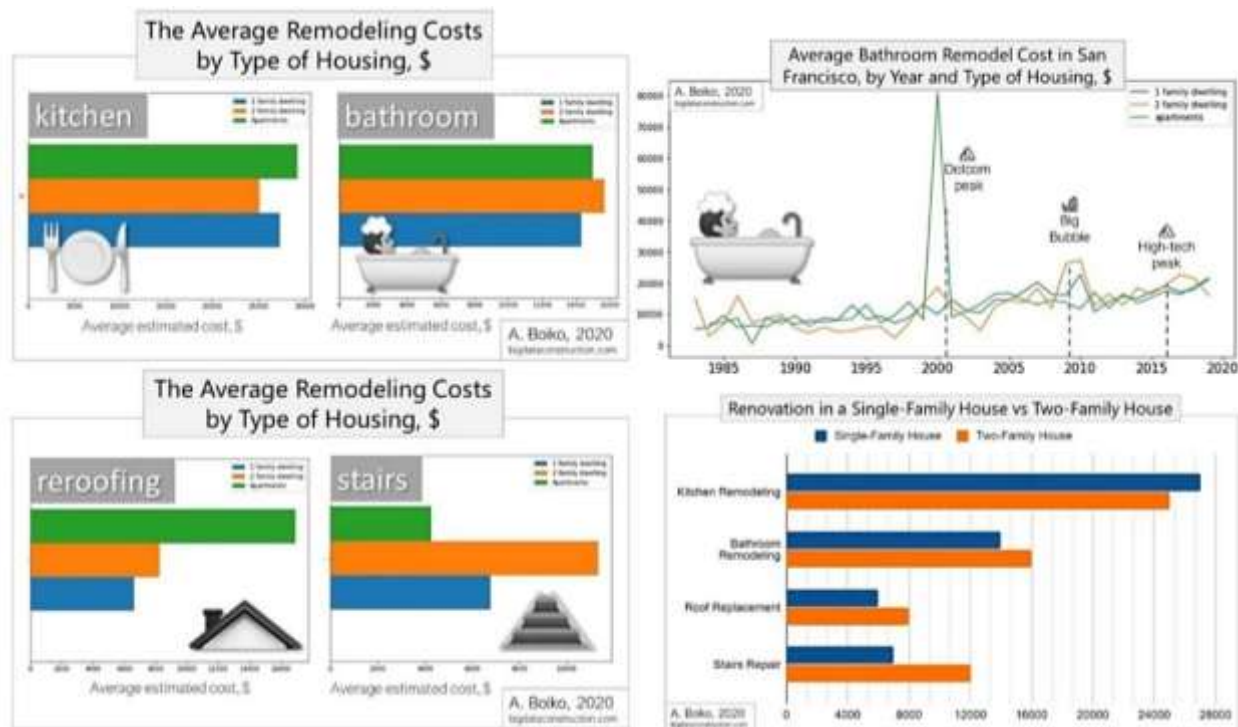
ניתוח נתונים מאפשר להעריך כי בסן פרנסיסקו, חלק גדול מתוך 91.5 מיליארד דולר שהושקעו בבנייה ושיפוץ בעשור (Folium ובספריית הוויזואליזציה Pandas-איור 9.15 - נעשה שימוש ב) האחרון - כמעט 75% - מרוכז במרכז העיר ובמרחק של 2 קילומטרים ממרכז העיר, מה שמשקף את הצפיפות הגבוהה יותר של השקעות באזורים המרכזיים הללו.

עלות ממוצעת של היתרי בנייה משתנה באופן משמעותי בהתאם לאזור, כאשר הבקשות במרכז העיר יקרות פי שלושה מאשר מחוץ לו, בשל עלויות גבוהות יותר של קרקע, כוח עבודה, חומרים ותקנות בנייה מחמירות, המצריכות שימוש



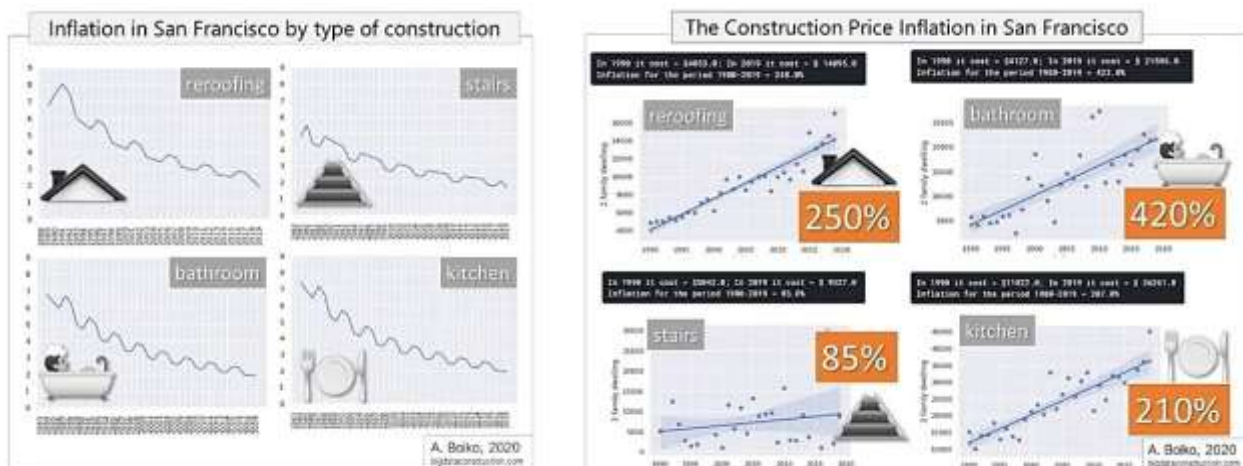
בסך פרנסיסקו 75 אחוזים מהשקעות הבנייה) 91.5 מיליארד דולר (מרוכזות במרכז העיר 9.15 ריש.

מערך הנתונים מאפשר גם לחשב מחירים ממוצעים לשיפוץ לא רק לפי סוגי בתים, אלא גם לפי אזורים בעיר וכתובות ספציפיות (מיקודים). (בסן פרנסיסקו, הדינמיקה של עלות שיפוץ הדירור מציגה מגמות ברורות עבור סוגי שיפוץ שונים: שיפוץ מטבח יקר משמעותית משיפוץ חדר אמבטיה. (Matplotlib ו-Pandas-השתמשו ב- 9.16. ריש) וסוגי דירור שיפוץ ממוצע של מטבח בבית חד משפחתי עולה כ-28,000 דולר לעומת 25,000 דולר בבית דו משפחתי



בסן פרנסיסקו שיפוץ מטבחים עולה כמעט כפול משיפוץ חדרי אמבטיה ובעלי בתים צריכים לחסוך 350 דולר מדי 9.16. ריש חודש במשך 15 שנים כדי לכסות את ההוצאות על שיפוץ בסיסי של הדיור.

ניתן לעקוב אחר האינפלציה בעלות הבנייה בסן פרנסיסקו עם השנים על ידי ניתוח נתונים מקובצים לפי סוגי דיור ושנים המראים עלייה מתמדת בעלות הממוצעת של שיפוץ מאז 1990 (Seaborn ו-Pandas-השתמשו ב- 9.17. ריש) ומזהים מחזורי עלות של שלוש שנים בשיפוץ דירות

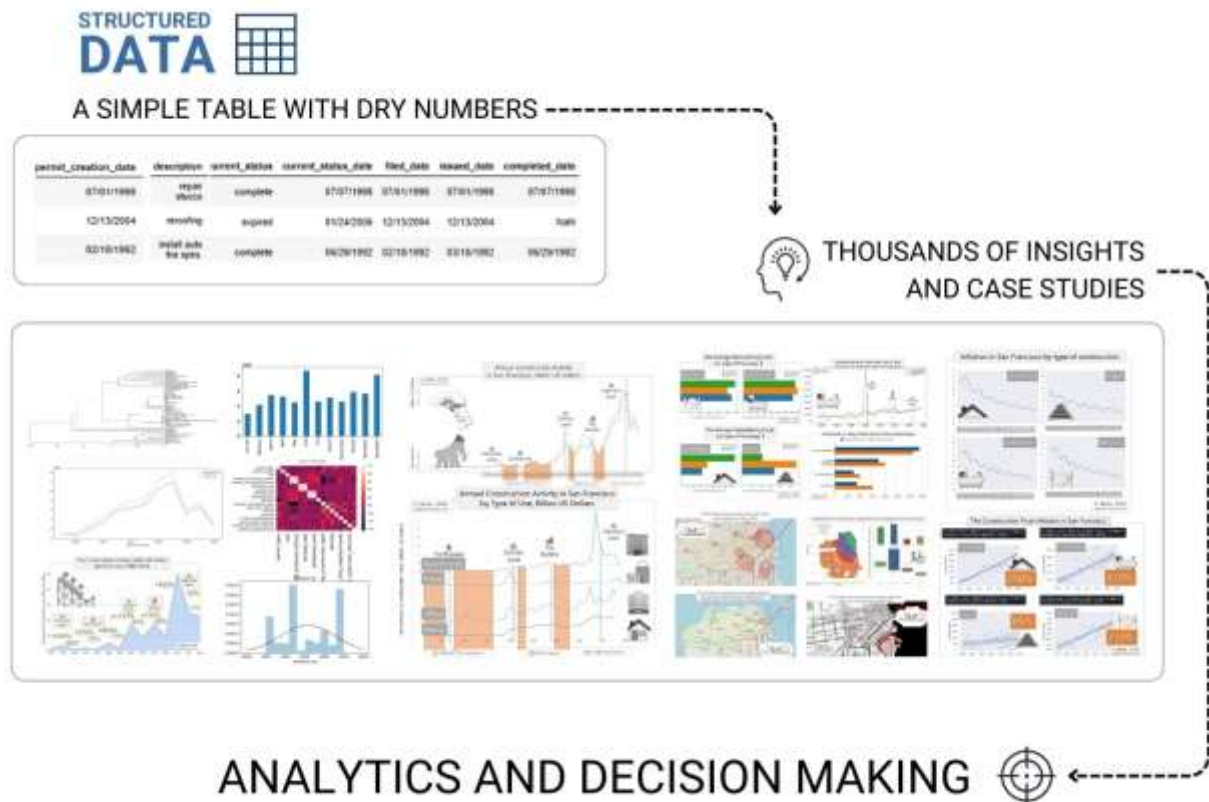


בין השנים 1980 ל-2019 עלות שיפוץ חדרי אמבטיה בסן פרנסיסקו עלתה פי חמישה, בעוד ששיפוץ גגות 9.17. ריש ומטבחים עלה פי שלושה, ושיפוץ מדרגות עלה רק ב-85%.

מאפשר לחשוף כי עלות הבנייה בעיר משתנה (9.13. ריש) מחקר נתונים פתוחים ממחלקת הבנייה של סן פרנסיסקו, מאוד ולעיתים קרובות אינה ניתנת לחיזוי, כאשר היא מושפעת ממגוון גורמים. בין הגורמים הללו נמצאים צמיחה כלכלית

חדשנות טכנולוגית ודרישות ייחודיות של סוגי דיור שונים.

התהליך הפך לנגיש, LLM בעבר, לצורך ניתוח כזה נדרשו ידע מעמיק בתכנות ובאנליטיקה. עם זאת, עם הופעת כלי וברור לקהל רחב של אנשי מקצוע בתעשיית הבנייה, החל מהנדסים במחלקות תכנון וכלה בהנהלה הבכירה של חברות.



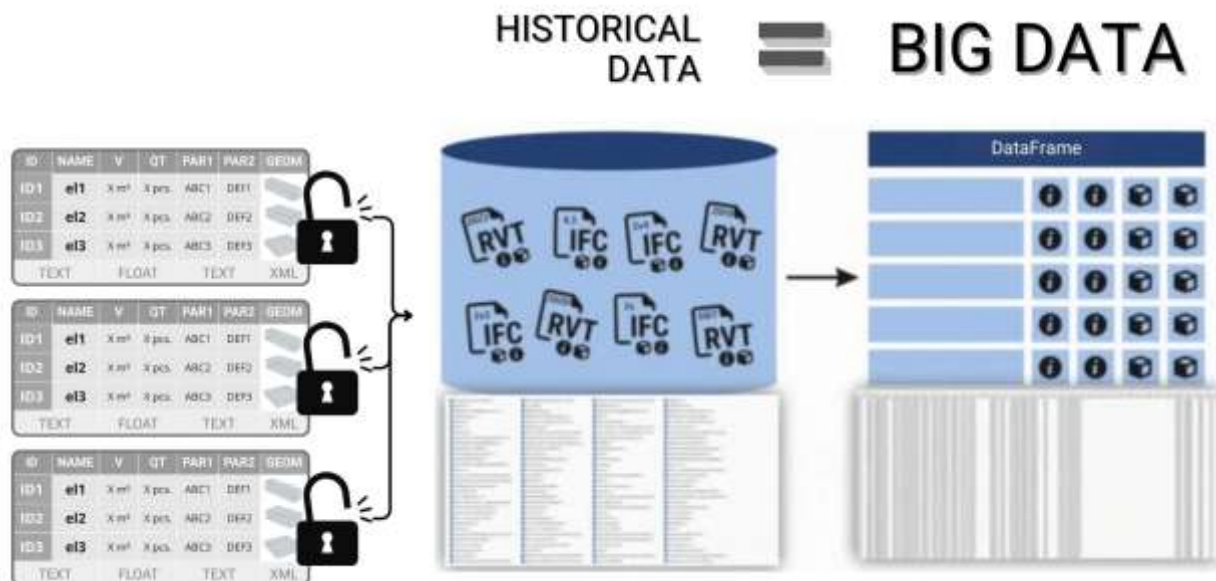
המעבר לנתונים ויזואליים מאפשר אוטומציה של קבלת החלטות באמצעות זיהוי דפוסים סמויים 9.18. ריש.

בדומה לניתוח הנתונים ממערך הנתונים של "מחלקת הבנייה של סן פרנסיסקו", "אנו יכולים להמחיש ולנתח כל מערך CAD או נתונים המתקבלים ממסדי נתונים, IoT, נתונים - החל מתמונות ומסמכים ועד נתוני

CAD (BIM) דוגמה לנתוני גדולים על בסיס נתוני

לצורך איסוף ויצירת CAD (BIM) בדוגמה הבאה ננתח מערך נתונים גדול, תוך שימוש בנתונים מכלים שונים של אינטרנטי אוטומטי (סקריפט), (שהוגדר לחיפוש ואיסוף קבצי פרויקטים crawler-מערך נתונים גדול נעשה שימוש ב הצליח למצוא crawler-במשך מספר ימים, ה. IFC ו-RVT מאתרים המציעים מודלים ארכיטקטוניים בחינם בפורמטים DWG ו-156,024 קבצי RVT ו-6,471 קבצי IFC ולהוריד 4,596 קבצי

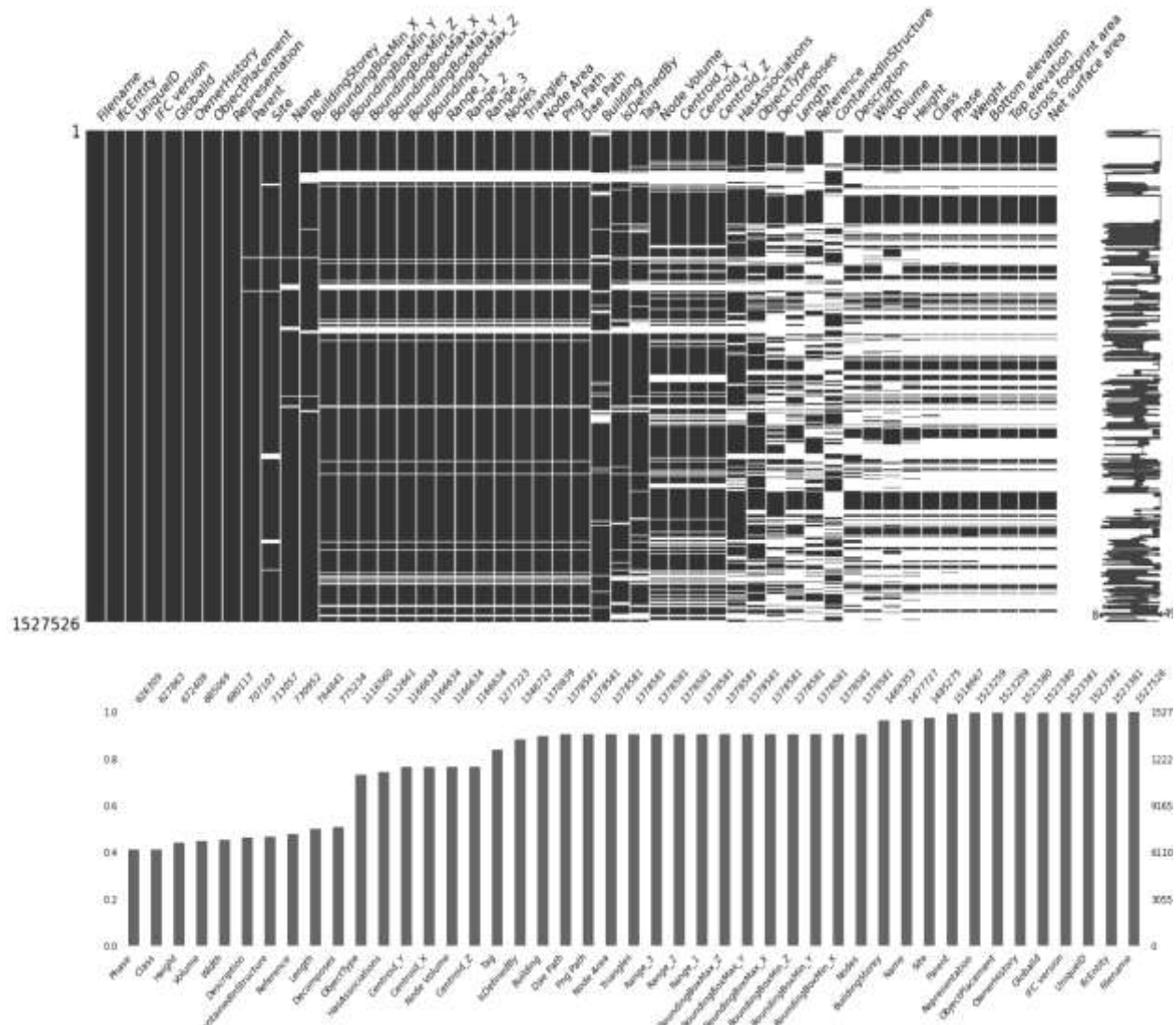
SDKs באמצעות CSV בגרסאות שונות והמרתם לפורמט מובנה IFC ו-RVT לאחר איסוף הפרויקטים בפורמטים Apache נאספו לקובץ טבלה גדול אחד בפורמט IFC ו-RVT חנימיים להנדסה הפוכה, כמעט 10,000 פרויקטים Pandas DataFrame-והועלו לניתוח ב Parquet.



- ריש 9.19 נתוני הפרויקט המובנים מאפשרים לאחד כל כמות של פרויקטים לטבלה דו-ממדית אחת

מכיל כ-4 מיליון ישויות) שורות (ו-24,962 תכונות IFC הנתונים מאוסף זה מכילים את המידע הבא: אוסף קבצי המורכב מכש 6 מיליון ישויות) שורות, (מכיל 27,025 תכונות שונות) עמודות (RVT), (עמודות, (בעוד שאוסף קבצי

- אוספי מידע אלו) ריש 9.110 (מכסים מיליוני רכיבים, עבור כל אחד מהם הושגו והוספו לטבלה הכללית (מלבן המגדיר את גבולות האובייקט בפרויקט) ונוצרו תמונות של כל Bounding Box קואורדינטות הגיאומטריה של - DAE (Collada) - פתוח XML וגיאומטריה בפורמט PNG רכיב בפורמט



של מיליון 100 התכונות הראשונות (missingno ספריית) ריש 9.110. תת-קבוצה של 1.5 מיליון רכיבים והוויזואליזציה בצורה של היסטוגרמה.

כאשר כל התכונות, RVT ו-6,471 פרויקט IFC כך קיבלנו את כל המידע על עשרות מיליוני רכיבים מ-4,596 פרויקט (DataFrame). הומו לצורת טבלה מובנית אחת (Bounding Box) מאפיינים של כל רכיב-ישות וגיאומטרייתם.

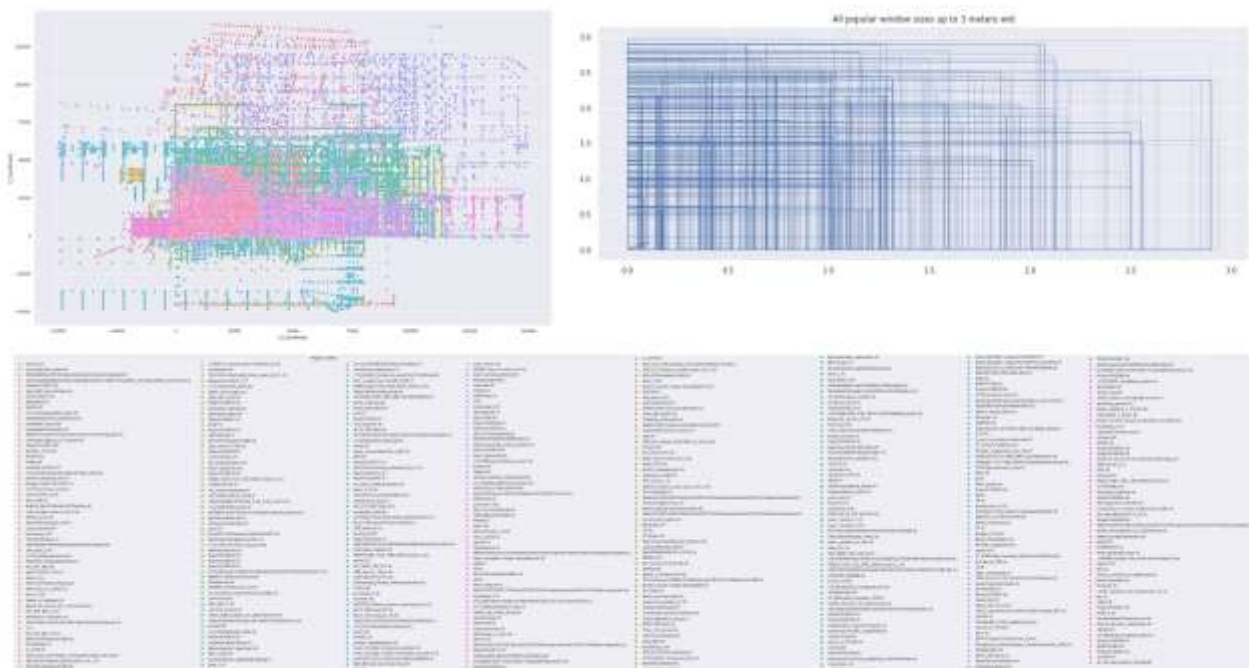
ההיסטוגרמות (ריש 9.110, ריש 9.26, ריש 9.27) שנבנו במהלך הניתוח, מאפשרות להעריך במהירות את צפיפות הנתונים ותדירות הופעת הערכים בעמודות. זה נותן תובנה ראשונית על התפלגות המאפיינים, נוכחות חוסרים ופוטנציאל השימושיות של תכונות מסוימות בניתוח ובבניית מודלים של --- למידת מכונה.

הזמין, "RVT ו-IFC אחד מהשימושים המעשיים של מערך נתונים זה) ריש 9.110. (הוא הפרויקט "5000 פרויקט מלא: מהכנה וניתוח נתונים ועד Pipeline עם פתרון Jupyter Notebook בפרויקט זה מוצג. Kaggle בבלטפורמת וועד Python - pandas, matplotlib, seaborn, folium ויזואליזציה של התוצאות באמצעות ספריות.



- **pandas** וספריית **Python** בעזרת ספריות ויזואליזציה של **CAD (BIM)** ריש **9.111** דוגמאות לניתוח נתונים מפורמטי

על בסיס המידע המטה-דאטי ניתן לקבוע באילו ערים פותחו פרויקטים מסוימים ולהציג זאת על מפה (למשל, בעזרת בנוסף, חותמות הזמן בנתונים מאפשרות לחקור תבניות לפי זמן שמירה או עריכה של קבצים: לפי **folium** ספריית ימי השבוע, שעות היום וחודשים



ויזואליזציה של המיקום הגיאומטרי של כל העמודים ומידות כל החלונות עד **3** מטרים בפרויקטים מהרשימה **9.112** איור בתחתית הגרף.

שהופקו מהמודלים, ניתנים גם לניתוח מצטבר. לדוגמה, באיור **Bounding Box** הפרמטרים הגיאומטריים בצורת מוצגים שני גרפים: הגרף השמאלי מציג את הפיזור של המרחקים בין העמודים בכל הפרויקטים ביחס לנקודת **9.112** האפס, והגרף הימני מציג את מידות כל החלונות בגובה של עד **3** מטרים מתוך מדגם של עשרות אלפי אלמנטים חלוניים ("**OST_Windows**", "**IfcWindows**") לאחר קיבוץ כל מערך הנתונים לפי הפרמטר "קטגוריה" (עם הערך).

תחת השם "5000 פרויקטים Kaggle לדוגמה זו והמערך הנתונים עצמו זמינים באתר Pipeline הקוד האנליטי של המוכן יחד עם מערך הנתונים ניתן Pipeline-את ה. [149] "IFC ו-RVT | DataDrivenConstruction.io". PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse עם תוסף Python, JupyterLab פופולריים עם תוסף PyDev, Thonny, Wing IDE, IntelliJ IDEA, Microsoft Azure Notebooks, Amazon SageMaker. כמו Kaggle.com, Google Collab, Amazon SageMaker.

הנתונים האנליטיים שהושגו כתוצאה מעיבוד ולימוד של כמויות עצומות של נתונים מובנים, ישחקו תפקיד מכריע בתהליכי קבלת ההחלטות בתעשיית הבנייה.

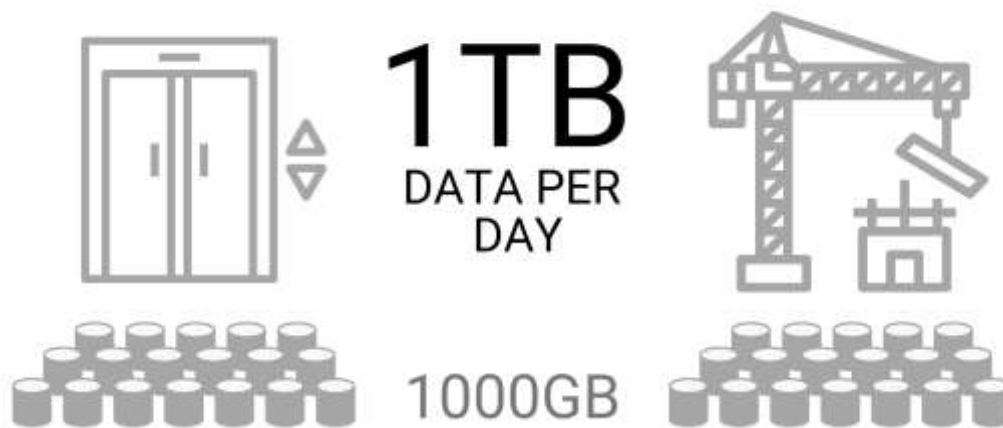
בזכות ניתוח מידע כזה המבוסס על נתונים מפרויקטים קודמים, אנשי מקצוע יכולים לחזות ביעילות, למשל, את הצרכים בחומרים ובכוח אדם ולייעל את ההחלטות התכנוניות עוד לפני תחילת הבנייה.

עם זאת, בעוד שפרטי הפרויקט או ההיתרים לבנייה הם מידע יחסית סטטי, המשתנה באיטיות יחסית, תהליך הבנייה עצמו – מצלמות, מערכות ניטור אוטומטיות, שמעבירות נתונים בזמן אמת: IoT מתמלא במהירות במגוון חיישנים ומכשירי כל אלה הופכים את אתר הבנייה לסביבה דיגיטלית דינמית, שבה יש לנתח נתונים בזמן אמת.

אינטרנט של הדברים וחוזים חכמים

משלו והופך IP) האינטרנט של הדברים (מייצג גל חדש של טרנספורמציה דיגיטלית, שבו כל מכשיר מקבל כתובת IoT הוא מושג המניח חיבור של אובייקטים פיזיים לאינטרנט לצורך איסוף, עיבוד והעברת IoT. לחלק מהרשת הגלובלית נתונים. בבנייה, זה אומר אפשרות לשלוט בתהליכי הבנייה בזמן אמת, למזער אובדן חומרים, לחזות בלאי של ציוד ולייעל את קבלת ההחלטות.

הכנה לעתיד באמצעות בנייה מחוברת] "150, [בעשור הקרוב תעבור תעשיית הבנייה CFMA לפי מאמר אתר בנייה – Connected Construction טרנספורמציה דיגיטלית רחבת היקף, שיא שלה יהיה המושג משולב ומאוטומטי לחלוטין.



או מכשירים להעברת נתונים באתר הבנייה יכולים לייצר ולהעביר טרה-בייטים של נתונים ביום IoT מכשירי 9.113 אור.

אתר הבנייה הדיגיטלי שכל מרכיבי הבנייה – מתכנון ולוגיסטיקה ועד ביצוע עבודות ובקרת איכות באתר הבנייה יתאגדו במערכת אקולוגית דינמית דיגיטלית אחת. קודם לכן, בפרק השביעי של – *Arone* במאפשר לארגן עיבוד נתונים בזמן אמת, *Apache NiFi* הספר, כבר בחנו את האפשרויות של הכלי החינמי והפתוח מאיסוף ממקורות שונים ועד העברה לאחסון או לפלטפורמות אנליטיות –.

נתונים על התקדמות הבנייה, צריכת חומרים, מצב הציוד ובטיחות יועברו בזמן אמת למערכות אנליטיות. זה מאפשר לחזות סיכונים פוטנציאליים, להגיב במהירות להפרות ולייעל תהליכים באתר. מרכיבי המפתח של אתר הבנייה הדיגיטלי – כוללים:

- מעקב אחרי פרמטרים של הסביבה, ניטור ציוד הבנייה ובקרת תנאי העבודה – *IoT* חיישני
 - דיגיטליים כפולים – מודלים וירטואליים של בניינים ותשתיות, המאפשרים לחזות סטיות אפשריות ולמנוע טעויות.
 - מערכות לוגיסטיות אוטומטיות – ניהול שרשראות אספקה בזמן אמת כדי לצמצם השבתות והוצאות
 - מערכות בנייה רובוטיות – שימוש במכונות אוטונומיות לביצוע משימות שגרתיות ומסוכנות
- לא רק יגבירו "*Connected Site (Construction)*" ומושג אתר הבנייה הדיגיטלי *IoT*-רובוטיזציה, שימוש נרחב באתר היעילות ויקטינו עלויות, אלא גם יפתחו עידן חדש של בטיחות, בנייה בת קיימא וניהול פרויקטים צפוי

(זיהוי בתדר רדיו. (הן משמשות לזיהוי *RFID* הם גם תוויות *IoT* אחד המרכיבים החשובים של רכיבי ומעקב אחרי חומרים, ציוד ואפילו עובדים באתר הבנייה, מה שמגביר את שקיפות וניהול המשאבים של הפרויקט.

משמשת לזיהוי אוטומטי של אובייקטים באמצעות אותות רדיו. היא מורכבת משלושה מרכיבים *RFID* טכנולוגיית מרכזיים:

- (פסיביות או אקטיביות –) מכילות מזהה ייחודי ומודבקות על חומרים, כלים או ציוד *RFID* תוויות
- סורקים – מכשירים שקוראים מידע מהתוויות ומעבירים אותו למערכת
- מסד נתונים מרכזי – מאחסן מידע על מיקום, מצב ותנועת אובייקטים

בבנייה-*RFID* השימוש ב

- 'ניהול אוטומטי של חומרים – תוויות על מוצרים מבטון מוכנים, ברזל או חבילות עם פאנלים סנדוויץ' מאפשרות לשלוט במלאי ולמנוע גניבות
- של עובדים רושמות את זמן תחילת וסיום המשמרת, מה שמבטיח *RFID* בקרת עבודה של עובדים – תעודות ניהול זמן עבודה
- עוקבת אחרי תנועת ציוד, מונעת השבתות ומגבירה את היעילות של הלוגיסטיקה *RFID* ניטור ציוד – מערכת את ערכת הטכנולוגיה הזו משלימים חוזים חכמים מבוססי טכנולוגיות בלוקצ'יין, המאפשרים לאוטומט תשלומים, בקרת אספקות ואכיפת תנאי חוזים ללא צורך במתווכים, מה שמפחית את הסיכון להונאה ועיכובים

כיום, בהיעדר מודל נתונים אחיד, חוזים חכמים מהווים פשוט קוד שמסכימים עליו המשתתפים. עם זאת, בגישה מונחת נתונים ניתן ליצור מודל משותף של פרמטרי החוזה, לקודד אותו בבלוקצ'יין ולמכן את ביצוע התנאים.

RFID ותוויות IoT לדוגמה, במערכת ניהול אספקה, חוזה חכם יכול לעקוב אחרי משלוח הסחורה באמצעות חיישני – ולבצע אוטומטית את התשלום עם הגעתה. באופן דומה, באתר הבנייה, חוזה חכם יכול לתעד את סיום שלב העבודה לדוגמה, התקנת ברזל או יציקת יסודות – בהתבסס על נתונים מחלל או חיישני בנייה וליזום אוטומטית את התשלום. הבא לקבלן מבלי צורך בבדיקות ידניות ובמסמכים פיזיים.

אך למרות הטכנולוגיות החדשות והמאמצים של ארגונים בינלאומיים לסטנדרטיזציה, מספר סטנדרטים מתחרים מסבכים IoT. את הנוף של

(IoT) לפי מחקר של סיסקו שפורסם בשנת 2017, כמעט 60% מהיוזמות בתחום האינטרנט של הדברים שלהן כהצלחה מלאה IoT-נעצרות בשלב הוכחת הקונספט, ורק 26% מהחברות רואות את פרויקטי היתרה מכך, שליש מהפרויקטים שהושלמו לא משיגים את המטרות המוצהרות ולא נחשבים להצלחה גם לאחר היישום.

אחת הסיבות המרכזיות היא חוסר תאימות בין הפלטפורמות המעבדות נתונים מחיישנים שונים. כתוצאה מכך, הנתונים נשארים מבודדים במסגרת פתרונות נפרדים. חלופה לגישה זו, כמו במקרים דומים (שעליהם דנו בספר זה, היא ארכיטקטורה הבנויה סביב הנתונים עצמם כנכס המרכזי.

משחקים תפקיד מרכזי לא רק במעקב אחרי מצב הטכנולוגי של הציוד, אלא גם בניתוח חיזוי IoT חיישני המאפשר להפחית סיכונים באתר הבנייה ולהגביר את הפרודוקטיביות הכללית של התהליכים, על ידי חיזוי תקלות וסטיות.

יכולים להתעבד בזמן אמת על ידי אלגוריתמים של למידת RFID ותוויות IoT הנתונים שנאספים באמצעות חיישני מכונה, המסוגלים לזהות אנומליות ולהתריע מראש למהנדסים על תקלות פוטנציאליות. זה יכול להיות כמו הופעת סדקים. מיקרוסקופיים במבנים מבטון, או הפסקות לא אופייניות בעבודת מנוף, המעידות על תקלות טכניות או הפרות של נהלים יתרה מכך, אלגוריתמים מתקדמים לניתוח התנהגות מאפשרים לתעד דפוסי התנהגות, שיכולים להעיד, למשל, על עייפות פיזית של העובדים, מה שמגביר את רמת הניהול הפרואקטיבי של הבטיחות ורווחת העובדים באתר.

קורות בפתאומיות. בדרך כלל, ישנם סטיות rarely – בתעשיית הבנייה, תאונות ותקלות – בין אם טכניות או אנושיות קטנות שמקדימות אותן, אשר נשארות לא מורגשות. ניתוח חיזוי ולמידת מכונה מאפשרים לזהות את האותות הללו בשלב מוקדם, עוד לפני שההשלכות הקריטיות מתרחשות.

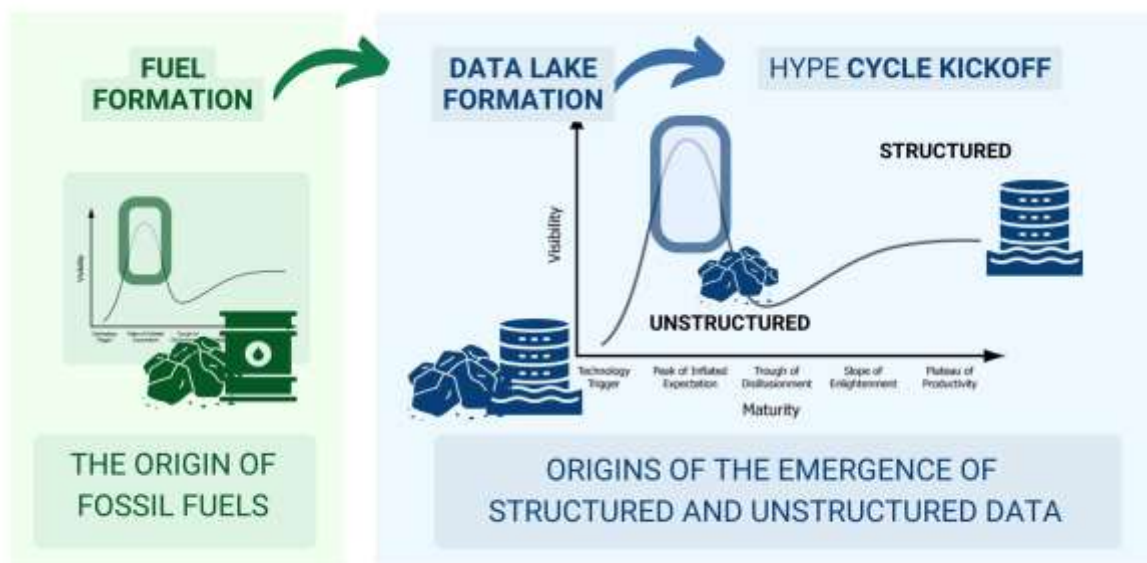
יוצרים עקבות דיגיטליות של פרויקטי הבנייה, אז RFID ותוויות IoT אם המסמכים, קבצי הפרויקט והנתונים מחיישני למידת מכונה מסייעת להפיק מהם ידע שימושי. עם עליית נפח הנתונים ודמוקרטיזציה של הגישה לנתונים, תעשיית הבנייה מקבלת הזדמנויות חדשות בתחום הניתוח, החיזוי והשימוש בבינה מלאכותית.



פרק 9.2. למידת מכונה וחיזויים

למידת מכונה ובינה מלאכותית ישנו את האופן שבו אנו בונים

בסיסי הנתונים של מערכות שונות בעסקי הבנייה - עם התשתית המתקבת והולכת ומסובכת שלהן - הופכים להיות קרקע פורייה לפתרונות עתידיים. השרתים של החברה, בדומה ליער, עשירים במידע חשוב, לעיתים קרובות מוסתר מתחת לפני השטח, בעמקי התיקיות והשרתים. כמויות הנתונים ממערכות שונות, הנוצרות כיום - לאחר השימוש, צניחה לתחתית השרת ולאחר התאבנות של שנים רבות - יהפכו לדלק עבור מודלים של למידת מכונה ומודלים לשוניים בעתיד על בסיס המודלים הפנימיים הללו, תוך שימוש באחסונים מרכזיים, ייבנו צ'אט פנימיים של חברות) למשל, גרסה נפרדת המאפשרים לקבל מידע במהירות ובנוחות וליצור, (LlaMa, Mistral, DeepSeek, מותאם מקומית ChatGPT של את הגרפים, הלוחות והמסמכים הנדרשים



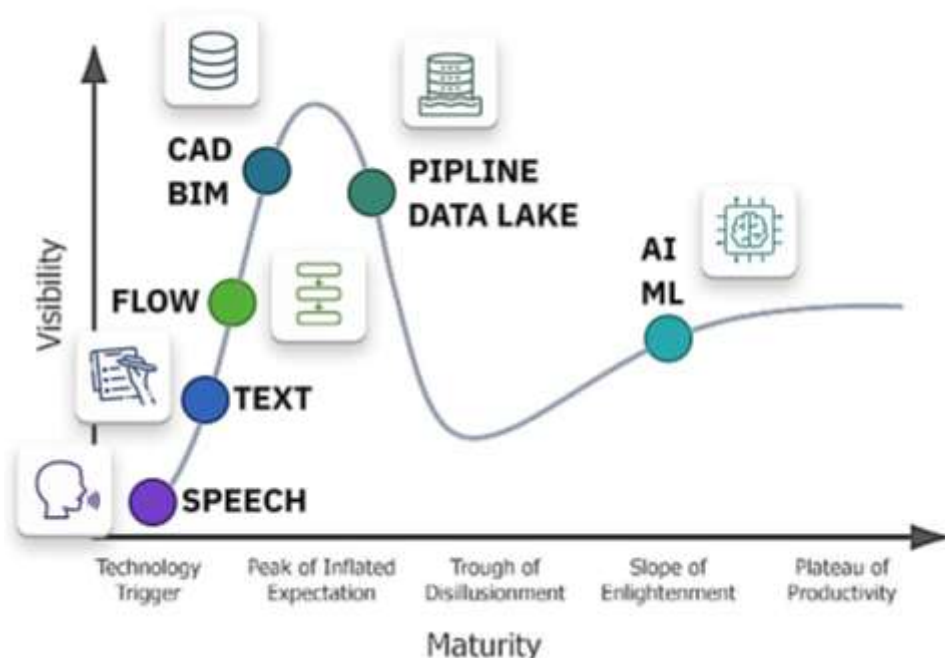
כמו שעצים הופכים לפחם, כך גם מידע עם הזמן, תחת לחץ הזמן והאנליסטים, הופך למקורות אנרגיה יקרי ערך עבור 9.21. ריס - העסק.

התאבנות של חומר צמחי בשילוב עם לחץ וטמפרטורה יוצרת מסה הומוגנית ומיוחדת של עצים ממינים שונים, שחיו בזמנים שונים - פחם] 152. [כך גם המידע, שנרשם על דיסקים קשיחים בפורמטים שונים ובזמנים שונים תחת לחץ של מחלקות האנליזה וטמפרטורת ניהול האיכות, בסופו של דבר יוצר מסה הומוגנית ומסודרת של מידע יקר ערך) ריס - 9.21).

השכבות הללו) או לעיתים קרובות, פיסות מידע מבודדות (נוצרות בעזרת עבודה קפדנית של ארגון נתונים על ידי אנליסטים מנוסים, המתחילים בהדרגה לחלק מידע יקר ערך מנתונים, שלכאורה, כבר אינם רלוונטיים

ברגע שהשכבות המבשילות הללו מפסיקות פשוט "לבעור" בדו"ח, ומתחילות לזרום בתהליכי העסק, מעשירות את ההחלטות ומשפרות את התהליכים, החברה מוכנה לשלב הבא - המעבר ללמידת מכונה ואינטליגנציה מלאכותית) ריס - 9.22).

היא קבוצה של שיטות לפתרון בעיות של אינטליגנציה (ML - Machine learning) למידת מכונה מלאכותית. אלגוריתמים של למידת מכונה מזהים תבניות בכמויות גדולות של נתונים ומשתמשים בהן כדי ללמוד באופן עצמאי. כל סט נתונים חדש מאפשר לאלגוריתמים מתמטיים להשתפר ולהתאים את עצמם בהתאם למידע שהתקבל, מה שמאפשר להעלות את דיוק ההמלצות והתחזיות באופן מתמיד.



- דעיכת הטכנולוגיות ליצירת נתונים ושימוש בכלי אנליזה פותחים את הדרך לנושא של למידת מכונה 9.22. ריס

כפי שאמר בראיון בשנת 2023 מנכ"ל משפיע של קרן השקעות הגדולה בעולם (שבבעלותה חבילות מניות מרכזיות כמעט בכל החברות הגדולות המייצרות תוכנה לבנייה, כמו גם חברות המחזיקות בכמות הגדולה ביותר של נדל"ן בעולם למידת מכונה תשנה את עולם הבנייה - [55]).

ל.אינטליגנציה מלאכותית יש פוטנציאל עצום. היא תשנה את האופן שבו אנו עובדים, את האופן שבו אנו חיים. אינטליגנציה מלאכותית ורובוטיקה יישנו את האופן שבו אנו עובדים וכיצד אנו בונים, ונוכל להשתמש באינטליגנציה מלאכותית ורובוטיקה כאמצעי ליצירת פרודוקטיביות רבה יותר [153]. - מנכ"ל הקרן הגדולה בעולם, ראיון, ספטמבר 2023.

פועלת על ידי עיבוד כמויות גדולות של נתונים, תוך שימוש בשיטות סטטיסטיות כדי לחקות היבטים (ML) למידת מכונה של חשיבה אנושית. עם זאת, רוב החברות אינן מחזיקות במערכי נתונים כאלה, ואם ישנם, הם לעיתים קרובות אינם להיות יעילה יותר ML-מסומנים כראוי. כאן יכולות לעזור טכנולוגיות סמנטיות ולמידה מעבירה - שיטה המאפשרת לבעבודה עם כמויות קטנות של נתונים, היתכנות אשר נדונה בפרקים הקודמים של חלק זה.

מהות הלמידה המעבירה היא שבמקום לעבד כל משימה מאפס, ניתן להשתמש בידע שהושג בתחומים קרובים. יש להבין כי תבניות וגילויים מתחומים אחרים של הכלכלה יכולים להיות מותאמים ויישומים בתחום הבנייה. לדוגמה, שיטות אופטימיזציה של תהליכי לוגיסטיקה, שפותחו בתחום הקמעונאות, מסייעות לשפר את היעילות של ניהול שרשראות

האספקה בבנייה. ניתוח נתונים גדולים, המיושם באופן פעיל בתחום הפיננסים, יכול לשמש לחיזוי עלויות ולניהול סיכונים בפרויקטים של בנייה. טכנולוגיות ראייה ממוחשבת ורובוטיקה, המתפתחות בתעשייה, כבר מוצאות יישום בבקרת איכות אוטומטית, ניטור בטיחות וניהול אתרים באתר הבנייה.

הלמידה המעבירה מאפשרת לא רק להאיץ את הכנסת החדשנות, אלא גם להפחית את העלויות של פיתוחן, תוך שימוש בניסיון שכבר צברו תחומים אחרים.

$$\text{labor productivity in construction} = f(\text{AI})$$

טכנולוגיות אינטליגנציה מלאכותית ורובוטיקה יהפכו לכוח המניע המרכזי של העתיד לשיפור הפרודוקטיביות בתחום 9.23. ריש - הבנייה

ריש, 4.419. החשיבה האנושית בנויה על עיקרון דומה: אנו נשענים על ידע קודם כדי לפתור בעיות חדשות (ריש 4.420, ריש 4.421). גישה זו גם פועלת - על ידי הפשטת מודל הנתונים והפיכתו ליותר. ML אלגנטי, ניתן להפחית את מורכבות המשימה עבור האלגוריתמים של --של נתונים ומצמצם את העלויות החישוביות

מע סובייקטיבית לחזוי סטטיסטי

הולך ונעלם. בתנאים של (ריש 9.24) העידן שבו החלטות אסטרטגיות היו תלויות באינטואיציה של מנהלים בודדים תחרות גוברת ומצבים כלכליים מורכבים, גישה סובייקטיבית הופכת להיות מסוכנת ולא יעילה מדי. חברות שממשיכות להסתמך על דעות אישיות במקום על ניתוח נתונים אובייקטיבי, מפסידות את היכולת להגיב במהירות לשינויים.

הסביבה התחרותית דורשת דיוק וחזרתיות, המבוססים על נתונים, תבניות סטטיסטיות וסיכויים ניתנים לחישוב. החלטות כבר לא יכולות להסתמך על תחושות, הן חייבות להסתמך על קורלציות, מגמות ומודלים: חיזוי שהושגו באמצעות אנליטיקה ולמידת מכונה. זו לא רק החלפת כלים - זו החלפת לוגיקת חשיבה. מהנחות - לעובדות, מסיכויים סובייקטיביים - להפרשים מחושבים סטטיסטית, מהרגשה - לעובדות



(הדעה של העובד המרוויח ביותר (עם הגעת הנתונים הגדולים ולמידת HiPPO עידן ההחלטות המתקבלות על ידי 9.24. ריש המכונה ייעלם.

מנהיגים המסתמכים אך ורק על תחושותיהם האישיות יתקלו במציאות חדשה: סמכות כבר לא קובעת את הבחירה. במרכז הניהול נמצאות כעת מערכות המנתחות מיליוני פרמטרים וכיוונים, חושפות דפוסים סמויים ומציעות אסטרטגיות אופטימליות.

הסיבה העיקרית לכך שחברות עדיין נמנעות מהטמעת למידת מכונה היא חוסר השקיפות שלה. רוב המודלים עבור מנהלים פועלים כמו "קופסאות שחורות", מבלי להסביר כיצד הם מגיעים למסקנותיהם. זה מוביל לבעיות: אלגוריתמים יכולים לחזק סטריאוטיפים ואפילו ליצור מצבים אבסורדיים, כמו במקרה של הצ'אט-בוט של מיקרוסופט, שהפך במהרה לכלי תקשורת רעיל.

IBM גארי קספרוב, אלוף העולם לשעבר בשחמט, משקף על תבוסתו למחשב "Deep Thinking" בספר Big Blue. הוא טוען שהערך האמיתי של אינטליגנציה מלאכותית אינו בהעתקת אינטליגנציה אנושית, אלא בהשלמת יכולותינו. אינטליגנציה מלאכותית צריכה לבצע משימות שבהן בני אדם חלשים, בעוד שבני אדם מביאים יצירתיות. מחשבים שינו את הגישה המסורתית לניתוח שחמט. במקום ליצור סיפורים מרתקים על משחקים, תוכניות שחמט ממוחשבות מעריכות כל מהלך באובייקטיביות, בהתבסס רק על כוחו או חולשתו, הממשית. קספרוב מציין שהנטייה האנושית לתפוס אירועים כסיפורים מחוברים, ולא כפעולות נפרדות. לעיתים קרובות מובילה למסקנות שגויות - לא רק בשחמט, אלא גם בחיים בכלל.

לכן, אם אתם מתכננים להשתמש בלמידת מכונה לצורך חיזוי וניתוח, חשוב להבין את העקרונות הבסיסיים שלה - כיצד פועלים האלגוריתמים וכיצד מעובדים הנתונים, לפני שתתחילו להשתמש בכלים של למידת מכונה ואינטליגנציה מלאכותית בעבודתכם. הדרך הטובה ביותר להתחיל היא באמצעות ניסיון מעשי.

וסט הנתונים Jupyter Notebook אחד הכלים הנוחים ביותר להיכרות ראשונית עם נושא למידת מכונה וחיזויים הוא שיאפשר לכם להבין בצורה מוחשית את השיטות המרכזיות לניתוח נתונים ובניית מודלים, Titanic, הקלאסי הפופולרי

של למידת מכונה

דאטהסט טיטניק: שלום עולם בעולם ניתוח נתונים וגדולים

אחד מהדוגמאות המפורסמות ביותר לשימוש בלמידת מכונה בניתוח נתונים הוא ניתוח סט הנתונים "טיטניק", אשר בלימוד "Hello World" משמש לעיתים קרובות ללימוד הסיכוי להישרדות נוסעים. חקר טבלה זו הוא אנלוגיה לתוכנית שפות תכנות.

בשנת 1912 הביאה למותם של 1502 מתוך 2224 אנשים. סט הנתונים של טיטניק RMS Titanic הטביעה של מכיל לא רק מידע על כך אם נוסע שרד, אלא גם תכונות כמו: גיל, מין, מעמד כרטיס ועוד פרמטרים. סט הנתונים הזה זמין בחינם, וניתן לפתוח ולנתח אותו בפלטפורמות שונות באף-ליין ובאונליין.

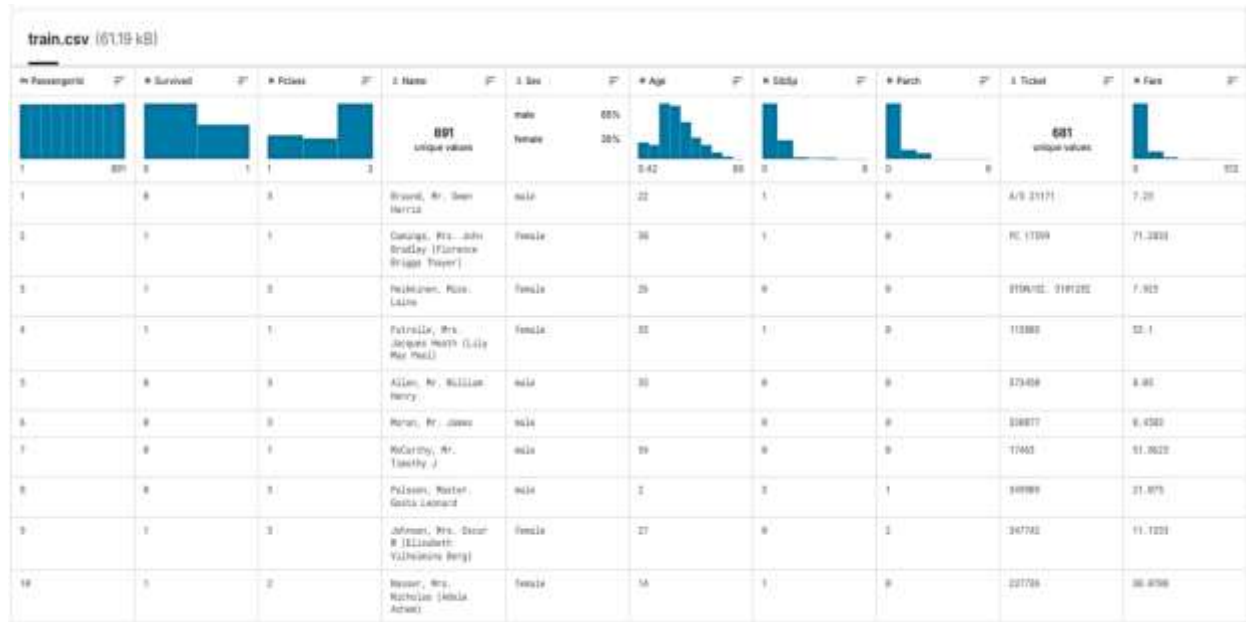
קישור לסט הנתונים של טיטניק:

<https://raw.githubusercontent.com/datasciencedojo/datasets/master/titanic.csv>

אחת הסביבות - Jupyter Notebook ושינויים עתידיים בתכנות "כבר נדונה LLM-עם תמיכה ב IDE" בפרק Jupyter Notebook-הפופולריות ביותר לפיתוח ניתוח נתונים וללמידת מכונה. חלופות חינוכיות מבוססות ענן ל ללא התקנת תוכנה ומספקות גישה Python המאפשרות להריץ קוד Kaggle ו-Google Collab הן הפלטפורמות החינמיות למשאבים חינוכיים.

קגל - הפלטפורמה הגדולה ביותר לניתוח נתונים, תחרויות בלמידת מכונה עם סביבה מובנית להרצת קוד. נכון לאוקטובר 2023, לקגל יש יותר מ-15 מיליון משתמשים מ-194 מדינות.

הורד והשתמש בנתוני הטיטניק בפלטפורמת קגל (איור 9.25), (כדי לאחסן את הנתונים) עותק שלהם (ולהריץ קוד -מיוחד IDE פייתון עם ספריות מותקנות מראש ישירות בדפדפן, מבלי צורך להתקין



איור 9.25 סטטיסטיקות של טבלת הטיטניק – הנתונים הפופולריים ביותר ללימוד ניתוח נתונים ולמידת מכונה

במהלך טביעתה בשנת 1912. הנתונים RMS Titanic נתוני הטיטניק כוללים מידע על 2224 נוסעים שהיו על הספינה מה שמאפשר להשתמש בהם, (test.csv) וקבוצת בדיקה (train.csv) מוצגים בשתי טבלאות נפרדות – קבוצת אימון הן לאימון מודלים והן להערכת דיוקם על נתונים חדשים

קבוצת האימון מכילה הן את תכונות הנוסעים (גיל, מין, מעמד כרטיס ועוד) (והן מידע על מי ששרד) (עמודה עם ערכים משמשת לאימון המודל. קבוצת הבדיקה) (איור 9.26 train.csv) (קובץ) בינאריים "שרד". ("קבוצת האימון כוללת רק את תכונות הנוסעים ללא מידע על שורדים) (ללא העמודה "שרד". ("קבוצת הבדיקה test.csv) (קובץ – 9.27 -מיועדת לבדוק את ביצועי המודל על נתונים חדשים ולהעריך את דיוקן

כך, יש לנו כמעט תכונות זהות של נוסעים בקבוצות האימון והבדיקה. ההבדל המרכזי הוא שבקבוצת הבדיקה יש לנו רשימה של נוסעים שבה חסרה העמודה "שרד" – המשתנה המטרה שאותו אנו רוצים ללמוד לחזות באמצעות "אלגוריתמים מתמטיים שונים. לאחר בניית המודל נוכל להשוות את הפלט של המודל שלנו עם הפרמטר האמיתי "שרד" מקבוצת הבדיקה, אשר נחשב להערכת התוצאות

העמודות העיקריות בטבלה, פרמטרי הנוסעים בקבוצות האימון והבדיקה:

■ מזהה ייחודי של נוסע – PassengerId

■ אם הנוסע שרד, 0 אם מת) (קבוצת הבדיקה) 1 – Survived

■ מעמד כרטיס) 1, 2 או 3 – Pclass

■ שם הנוסע – Name

■ מין הנוסע (זכר/נקבה) – Sex

■ גיל – Age

■ מספר אחים/אחיות או בני זוג על הספינה – SibSp

■ מספר הורים או ילדים על הספינה – Parch

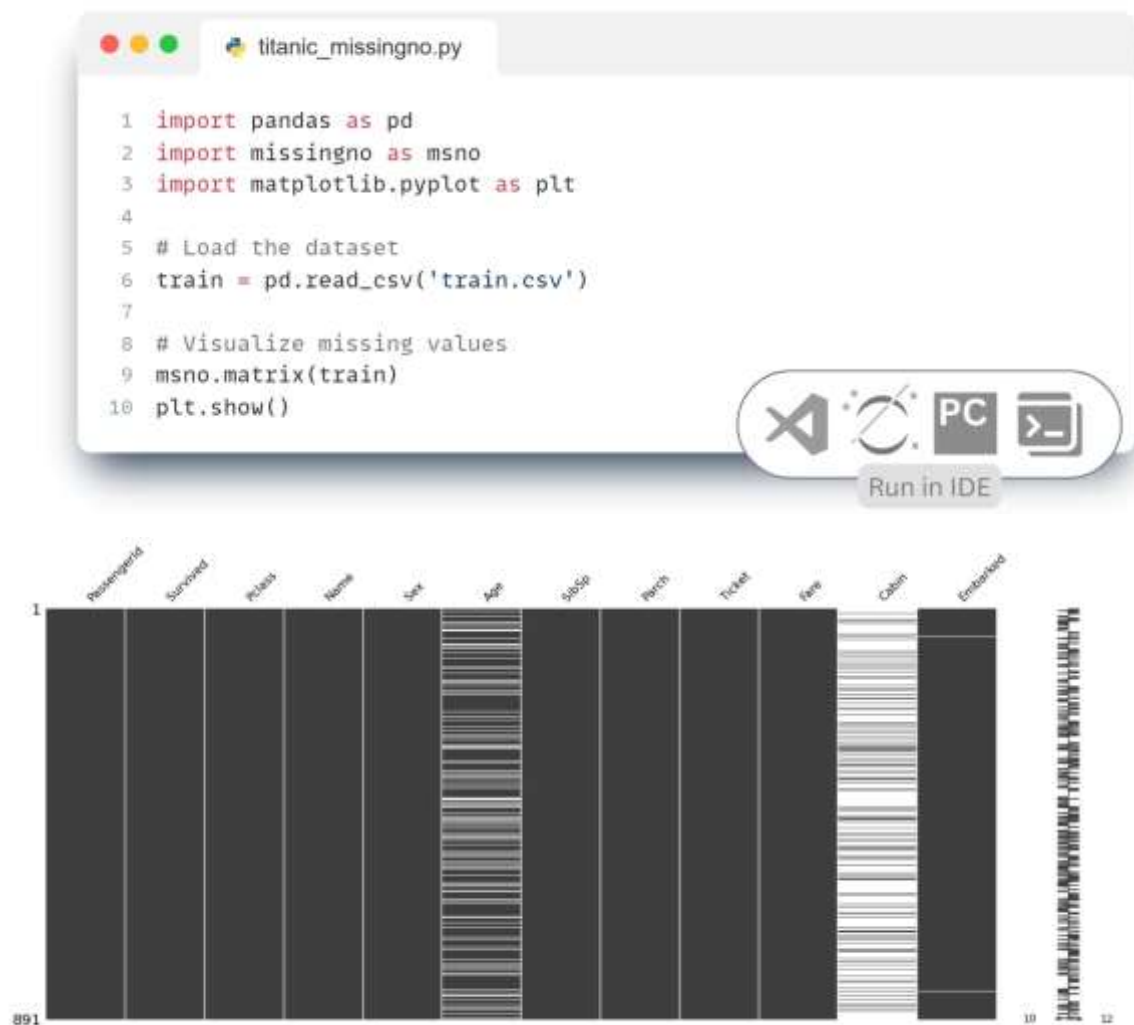
■ מספר כרטיס – Ticket

■ מחיר כרטיס – Fare

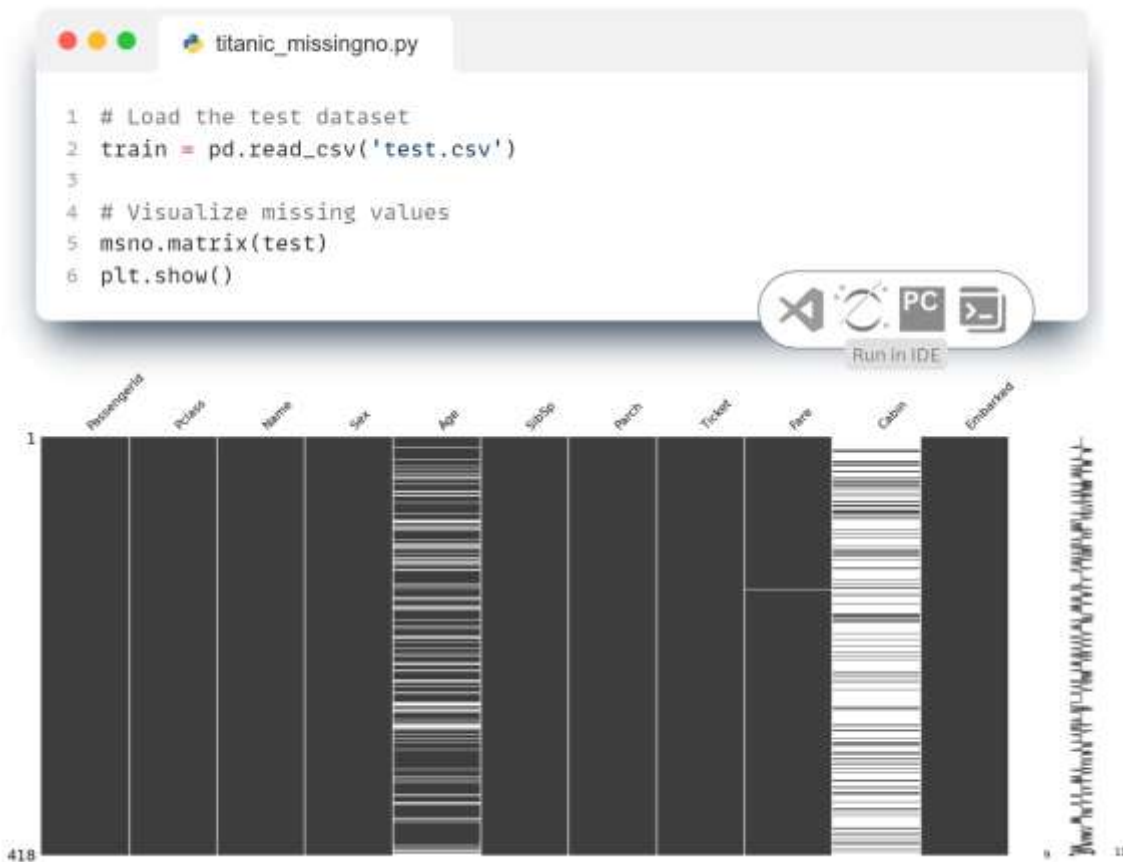
■ מספר קבינה (רבים מהנתונים חסרים) – Cabin

■ (סאות'המפטון = S, קווינסטאון = Q, שרבור = C) נמל העלייה – Embarked

(איור 9.26, איור 9.27) missingno כדי להמחיש נתונים חסרים בשתי הטבלאות ניתן להשתמש בספריית הממחשה ערכים חסרים בצורה של היסטוגרמה, שבה השדות הלבנים מראים נתונים חסרים. המחשה כזו מאפשרת להעריך במהירות את איכות הנתונים לפני עיבודם.



איור 9.26 באמצעות מספר שורות קוד ניתן להמחיש את הנתונים החסרים בקבוצת האימון של הטיטניק, כאשר הפרמטר המרכזי "לאימון הוא הפרמטר "שרד".



9.27 ויזואליזציה של נתונים חסרים בסט הנתונים של טיטאניק, אשר מכיל רק את מאפייני הנוסעים ללא מידע נוסף

השערות ומבצע תחזיות על בסיס סט הנתונים, ניתוח חזותי מסייע לחשוף דפוסים formulates לפני שנ מרכזיים בנתונים, להעריך את איכותם ולקבוע תלות אפשרית. ישנם מגוון שיטות ויזואליזציה המסייעות להבין, טוב יותר את סט הנתונים של טיטאניק. ניתן להשתמש בגרפים של התפלגות לניתוח קבוצות גיל של נוסעים דיאגרמות הישרדות בהתאם למגדר ולמחלקה, כמו גם מטריצות של נתונים חסרים כדי להעריך את איכות המידע ולהבין את הנתונים.

לסייע לנו בוויזואליזציה של הנתונים מסט הנתונים של טיטאניק, לשם כך נשלח את LLM-נבקש מ (ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN אחר מודל):

הראה בבקשה כמה גרפים פשוטים לסט הנתונים של טיטאניק. הורד את סט הנתונים בעצמך והצג ↩ את הגרפים

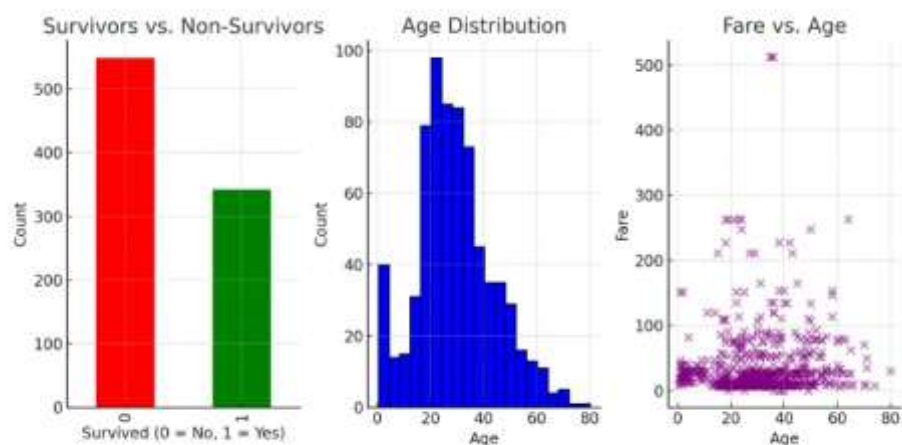
? בצורה של קוד מוכן וגרפים, הממחישים את פרמטרי סט הנתונים LLM תשובת

Create in LLM Chat

titanic_plot.py

```

1 import pandas as pd
2 import matplotlib.pyplot as plt
3
4 # Load the dataset
5 file_path = "/mnt/data/titanic.csv"
6 df = pd.read_csv(file_path)
7
8 # Convert column names to lowercase for consistency
9 df.columns = df.columns.str.lower()
10
11 # Basic Plots
12 plt.figure(figsize=(10, 5))
13
14 # Bar Chart: Survival Count
15 plt.subplot(1, 3, 1)
16 df['survived'].value_counts().plot(kind='bar', color=['red', 'green'])
17 plt.title("Survivors vs. Non-Survivors")
18 plt.xlabel("Survived (0 = No, 1 = Yes)")
19 plt.ylabel("Count")
20
21 # Histogram: Age Distribution
22 plt.subplot(1, 3, 2)
23 df['age'].dropna().plot(kind='hist', bins=20, color='blue', edgecolor='black')
24 plt.title("Age Distribution")
25 plt.xlabel("Age")
26 plt.ylabel("Count")
27
28 # Scatter Plot: Fare vs. Age
29 plt.subplot(1, 3, 3)
30 plt.scatter(df['age'], df['fare'], alpha=0.5, color='purple')
31 plt.title("Fare vs. Age")
32 plt.xlabel("Age")
33 plt.ylabel("Fare")
34
35 # Show the plots
36 plt.tight_layout()
37 plt.show()
    
```



מסייע להשיג מידית ויזואליזציה של נתוני הסט **9.28 LLM** איור.

ויזואליזציה של נתונים היא שלב חשוב, המאפשר להכין את סט הנתונים לבניית מודל למידת מכונה, אליו ניתן לעבור רק לאחר הבנת הנתונים.

למידת מכונה בפעולה: מנוסעים בטיטאניק לניהול פרויקטים

ההשערה המרכזית, המיועדת לחקר יסודות למידת מכונה על בסיס סט הנתונים של טיטאניק, היא כי קבוצות מסוימות של נוסעים היו בעלות סיכוי גבוה יותר לשרוד.

טבלה קטנה של נוסעי טיטאניק הפכה לפופולרית ברחבי העולם, ומיליוני אנשים משתמשים בה ללמידה ניסויים ובדיקת מודלים כדי לגלות אילו אלגוריתמים והשערות יאפשרו לבנות את המודל המדויק ביותר לחיזוי הישרדות על בסיס סט הנתונים האימוני של נוסעי טיטאניק.

האטרקטיביות של סט הנתונים של טיטאניק מוסברת על ידי הקומפקטיות שלו: עם כמה מאות שורות ושנים עשר עמודים הוא מספק הזדמנויות רחבות לניתוח. סט הנתונים מהווה, יחסית, דוגמה קלאסית פשוטה לפתרון סיווג (9.26 איור) בינארי, שבו מטרת המשימה – הישרדות – מבטאת בפורמט נוח של 0 או 1.

טוען כי בבסיס היקום קיימות החלטות בינאריות. באופן דומה, גם עסק [7] "It from Bit" - ג'ון וילר ב המנוהל על ידי אנשים, המורכבים ממולקולות, למעשה בנוי על סדרה של בחירות בינאריות.

בנוסף, הנתונים מבוססים על אירוע היסטורי אמיתי, מה שהופך אותם לערך למחקר, בניגוד לדוגמאות שנוצרו באופן בפתרון ETL ו-Data Pipeline אחת מהפלטפורמות הגדולות ביותר לעבודה עם, Kaggle מלאכותי. רק בפלטפורמת Data Pipeline בעיות על בסיס סט הנתונים של טיטאניק השתתפו 1,355,998 אנשים, שפיתחו 53,963 פתרונות. (9.29 איור) [157] ייחודיים.

זה נראה מדהים, אבל רק 1000 שורות נתונים על נוסעי "טיטאניק" עם 12 פרמטרים הפכו לזירה עבור ייחודיים. מסט נתונים קטן נולדים תובנות Data-Pipelines-מיליוני השערות, שרשראות לוגיות ו אינסופיות, השערות ופירושים – מדגמים פשוטים של הישרדות ועד אנסמבלים מורכבים, המתחשבים בדפוסים חבויים ובמבוכים מורכבים של מחשבה.

Machine Learning from Disaster

Submit Prediction

Data
Code
Models
Discussion
Leaderboard
Rules

Titanic Tutorial Updated 3y ago <a>29858 comments · Titanic - Machine Learning from Disaster	16916 Gold ***
Titanic competition w/ TensorFlow Decision Forests Updated 2y ago Score: 0.80143 · <a>318 comments · Titanic - Machine Learning from Disaster	1098 Gold ***
Titanic Data Science Solutions Updated 6y ago <a>2590 comments · Titanic - Machine Learning from Disaster	10723 Gold ***
Exploring Survival on the Titanic Updated 7y ago Score: 0.80382 · <a>1072 comments · Titanic - Machine Learning from Disaster	3968 Gold ***

איור 9.29 חמשת הפתרונות הראשונים מתוך 53,963 צינורות מוכנים ופתוחים. כמעט 1.5 מיליון אנשים כבר ניסו לפתור את [157]. המשימה הזו רק בקגל

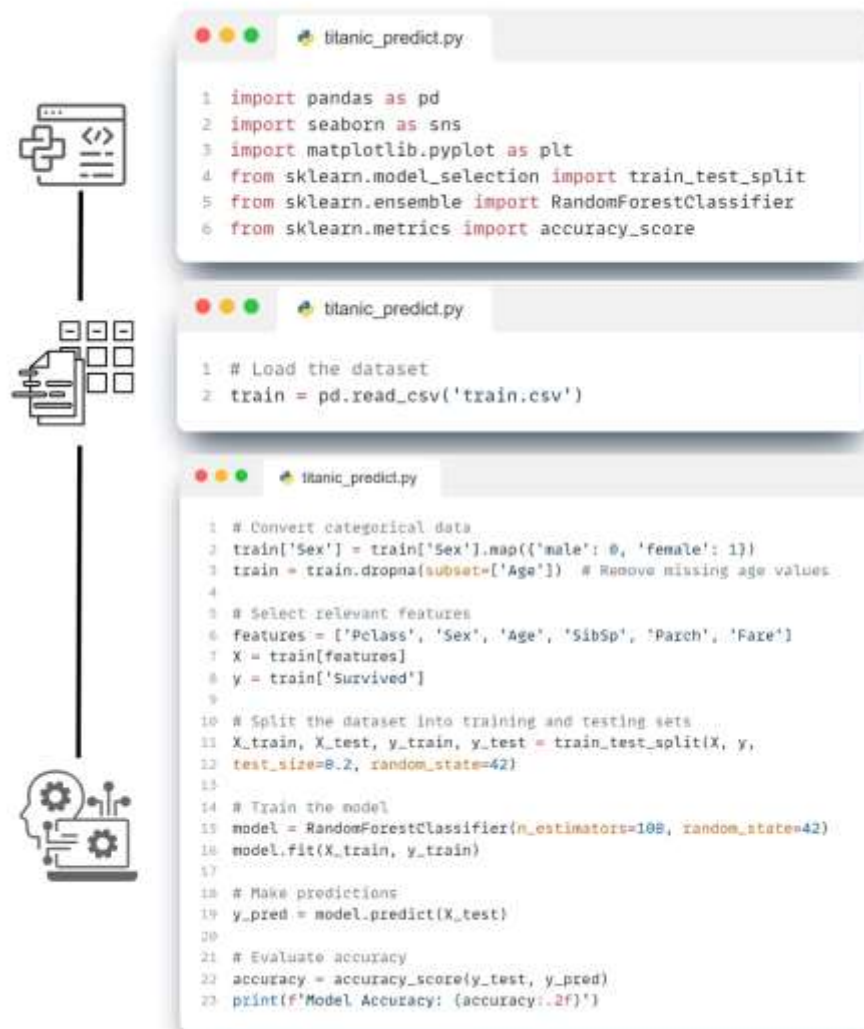
אם אפילו טבלה קטנה כזו יכולה לייצר מיליוני פתרונות ייחודיים (איור 9.29), (מה אפשר לומר על מערכות נתונים - בתעשייה האמיתית, שבהן הפרמטרים נמדדים בעשרות אלפים

סטנדרטי של בניין קטן מכיל עשרות אלפי ישויות עם אלפי פרמטרים - מהמאפיינים CAD פרויקט הגיאומטריים ועד לאטריבוטים של עלות וזמן. דמיינו כמה תובנות פוטנציאליות, קשרים, תחזיות והיפותזות ניהוליות טמונות בנתונים של כל הפרויקטים של החברה שלכם, שנאספו במהלך השנים האחרונות. נתוני פרויקטים היסטוריים אינם רק ארכיון - הם זיכרון חי של הארגון, עקבות דיגיטליים שניתן לנתח כדי לבנות כמות גדולה של היפותזות ייחודיות

הדבר החשוב ביותר הוא שאין צורך לחכות שהקהילה בקגל תתעניין בחברה שלכם או בנתונים שלכם. כבר היום אפשר להתחיל לעבוד עם מה שיש לכם: להריץ אנליטיקה על הנתונים שלכם, לאמן מודלים על הנתונים שלכם, לזהות חזרות גישה פתוחה לנתונים, LLM, אנומליות ודפוסים. במקום שדרושים שנים של ניסויים וייעוץ יקר, היום מספיק יוזמה ורצון ללמוד

כדי לבנות אלגוריתם למידת מכונה שיחזיק תחזיות על הישרדות נוסעים בהתבסס על מערך הנתונים לפתור את המשימה הזו עבורנו LLM-נבקש מ, train.csv האימוני

להתבסס על מערך הנתונים האימוני של נוסעי הטיטניק, בנה מודל למידת מכונה לחיזוי הישרדות



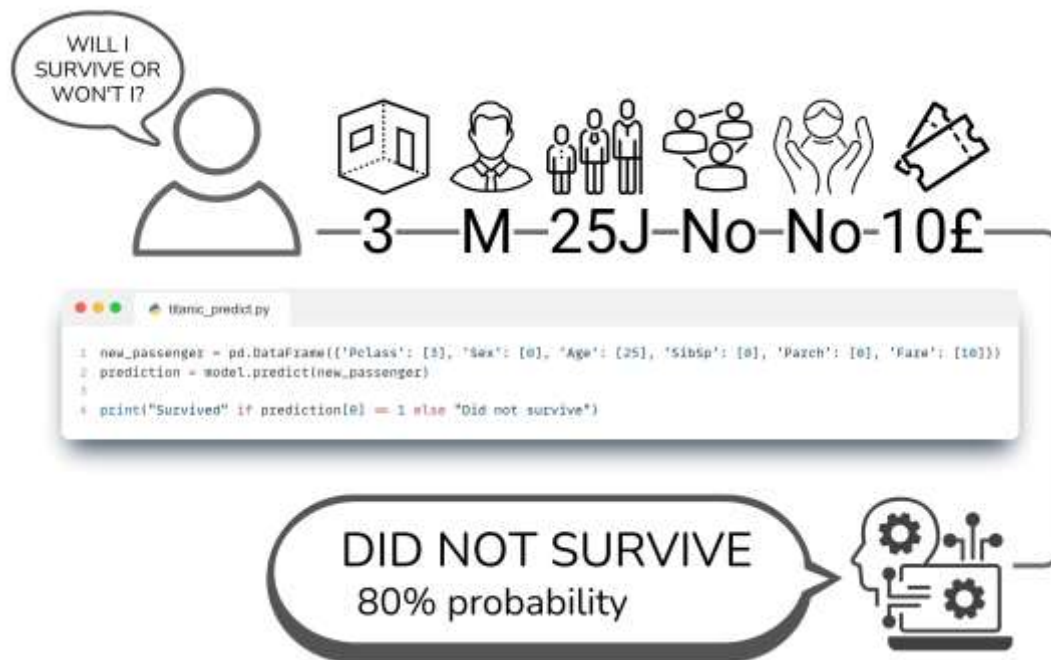
9.210 Random Forest. בנה תחזיות על הישרדות בטיטניק באמצעות אלגוריתם למידת מכונה LLM איור 9.210

(איור 9.210) טוען את נתוני הנוסעים של "טיטניק", מנקה אותם, ממיר משתנים קטגוריאליים ל-LLM-הקוד שהתקבל מ לחזות אם נוסע שרד RandomForestClassifier) למשל, מגדר לפורמט מספרי (ומאמן מודל באמצעות אלגוריתם -או לא) נרחיב על אלגוריתמים פופולריים בפרקים הבאים)

test.csv בעזרת הקוד, הנתונים במהלך האימון מחולקים לסט אימון וסט בדיקה) באתר קגל כבר נוצרו קבצים מוכנים (איור 9.26) (לצורך האימון, (ולאחר מכן המודל מאומן על הנתונים האימוניים ובדק על train.csv)-איור 9.27) (עם נתונים test.csv-הנתונים הבודקים, כדי להבין עד כמה המודל טוב בתחזיותיו. לאחר האימון, הנתונים הבודקים מ אמיתיים על מי ששרד או לא (מוזנים למודל, והוא מנבא מי שרד ומי לא. במקרה שלנו, הדיוק של מודל למידת המכונה --שהתקבל הוא כ-80%, מה שמראה שהוא מצליח לזהות דפוסים בצורה טובה

ניתן להשוות למידת מכונה לילד שמנסה להכניס בלוק מלבני לחור עגול. בשלביו הראשונים, האלגוריתם מנסה גישות רבות, נתקל בשגיאות ובאי-התאמות. תהליך זה עשוי להיראות לא יעיל, אך הוא מספק למידה חשובה: על ידי ניתוח כל שגיאה, המודל משפר את תחזיותיו ומקבל החלטות מדויקות יותר ויותר.

עכשיו ניתן להשתמש במודל הזה (איור 9.210) כדי לחזות את סיכויי ההישרדות של נוסעים חדשים. לדוגמה, אם נזין עם הפרמטרים: גבר, "כיתה שלישית", "בן 25", "ללא model.predict מידע על נוסע באמצעות הפונקציה קרובים על הסיפון, "המודל יניב תחזית – שהנוסע עם סיכוי של 80% לא ישרוד במקרה של אסון, אם הוא היה על -הספינה טיטאניק בשנת 1912) איור 9.211.



איור 9.211 המודל שיצרנו, עכשיו יכול לחזות עם סיכוי של 80% אם נוסע חדש בטיטאניק ישרוד או לא.

מודל החיזוי של ההישרדות נוסעי "טיטאניק" מדגים מושג רחב הרבה יותר: מדי יום אלפי מומחים בתעשיית הבנייה מקבלים החלטות "דו-ערכיות" דומות – החלטות של חיים או מוות, פרויקט, תקציב, כלי, רווח או הפסד, בטיחות או סיכון. כמו בדוגמה של "טיטאניק", "שבה התוצאה תלויה בגורמים (מין, גיל, כיתה, גם בתעשיית הבנייה כל אחד מהאספקטים של ההחלטה מושפע ממגוון גורמים ומשתנים) (עלויות חומרים, הכשרה של עובדים, לוחות זמנים, מזג האוויר, לוגיסטיקה, סיכונים טכניים, הערות ומאות אלפי פרמטרים נוספים).

בתעשיית הבנייה, למידת מכונה מיושמת באותם עקרונות כמו בתחומים אחרים: המודלים מאומנים על נתונים היסטוריים – מפרויקטים, חוזים, תקציבים – כדי לבדוק השערות שונות ולמצוא את הפתרונות היעילים ביותר. תהליך זה מזכיר במידה רבה את חינוכו של ילד באמצעות שיטת ניסוי וטעייה: עם כל מחזור המודלים מתאימים את עצמם ומדויקים יותר.

השימוש בנתונים שנצברו פותח אופקים חדשים לבנייה. במקום חישובים ידניים מעייפים, ניתן לאמן מודלים המסוגלים

לחזות עם דיוק גבוה את המאפיינים המרכזיים של פרויקטים עתידיים. כך, אנליטיקה חזויית הופכת את תעשיית הבנייה למקום שבו ניתן לא רק לתכנן, אלא גם לחזות בביטחון את התפתחות האירועים.

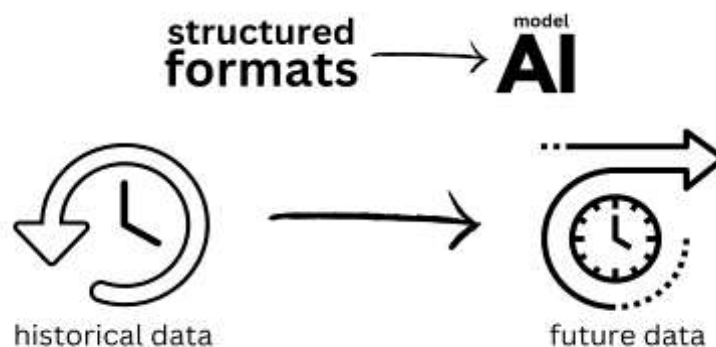
תחזיות וניבויים על בסיס נתונים היסטוריים

הנתונים שנאספו על פרויקטי החברה פותחים אפשרות לבניית מודלים המסוגלים לחזות את העלות ואת המאפיינים הזמניים של אובייקטים עתידיים, שעדיין לא מומשו – ללא חישובים ידניים מעייפים והשוואות. זה מאפשר להאיץ ולפשט את תהליכי ההערכה, בהתבסס לא על השערות סובייקטיביות, אלא על תחזיות מתודולוגיות מבוססות.

קודם לכן, בחלק הרביעי של הספר, דנו בשיטות המסורתיות לחישוב עלות פרויקטים, כולל השיטה המשאבית, וכן הזכרנו את הגישות הפרמטריות והאקספרטיות. שיטות אלו עדיין רלוונטיות, אך בפרקטיקה המודרנית הן מתחילות להתעשר בכלים של ניתוח סטטיסטי ולמידת מכונה, מה שמאפשר להגדיל משמעותית את הדיוק והחזרתיות של ההערכות.

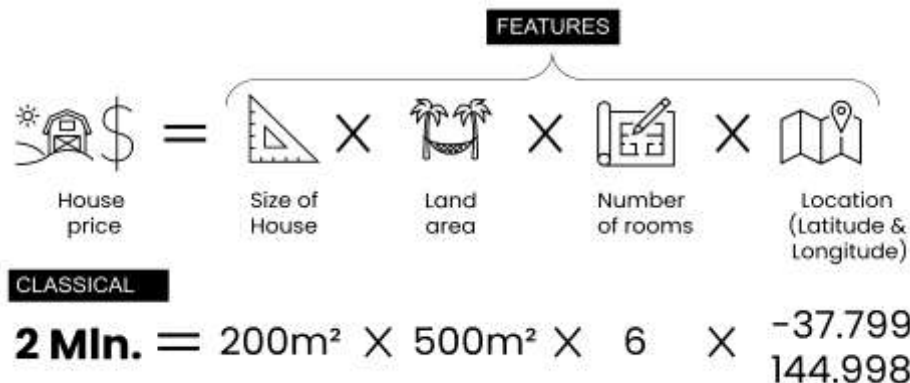
תהליכי חישוב ידני וחצי-אוטומטי של מחירים ומאפיינים זמניים יושלמו בעתיד על ידי דעות ותחזיות של מודלים של למידת מכונה, המסוגלים לנתח נתונים היסטוריים, למצוא תבניות חבויות ולהציע פתרונות מבוססים. נתונים חדשים יוצרים טקסטים, תמונות (LLM) ותסריטים ייווצרו אוטומטית מתוך המידע הקיים – בדומה לאופן שבו מודלים לשוניים וקוד על בסיס נתונים שנאספו במשך שנים ממקורות פתוחים.

כמו שהאדם היום מתבסס על ניסיון, אינטואיציה וסטטיסטיקה פנימית בהערכת אירועים עתידיים, בשנים הקרובות העתיד של פרויקטים בנייה ייקבע יותר ויותר על ידי שילוב של ידע מצטבר ודגמים מתודולוגיים של למידת מכונה.



נתונים איכותיים ומסודרים מההיסטוריה של החברה - חומר שעל בסיסו נבנים דגמי למידת מכונה וחזויים 9.212 איור

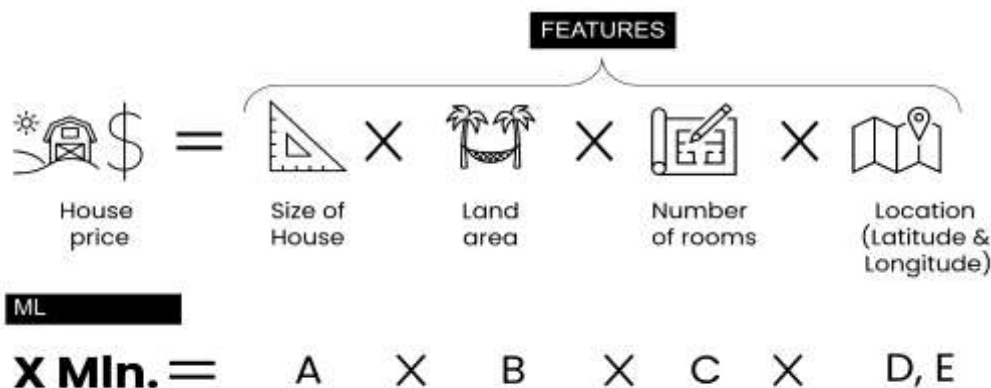
נבחן דוגמה פשוטה: חיזוי מחיר בית על בסיס שטחו, גודל המגרש, מספר החדרים והמיקום הגיאוגרפי. אחד הגישות היא גישה כזו דורשת נוסחה (9.213 איור) בניית מודל קלאסי, שמנתח את הפרמטרים הללו ומחשב את המחיר המשוער מדויקת ומוכרת מראש, דבר שלמעשה כמעט בלתי אפשרי במציאות.



כדי להעריך את עלות הבית ניתן להשתמש באלגוריתם קלאסי עם נוסחה קבועה, שיש למצוא אותה **9.213** איור.

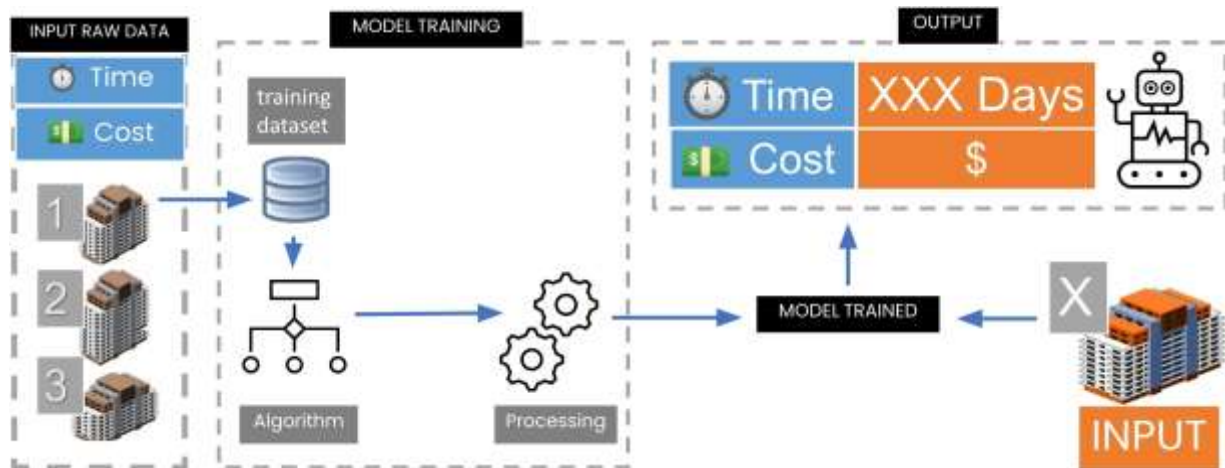
למידת מכונה מאפשרת להימנע מחיפוש ידני של נוסחאות ולהחליף אותן באלגוריתמים נלמדים, שמזהים באופן עצמאי תלותים, שמעלים בהרבה על דיוק כל משוואה שנקבעה מראש. כחלופה, ניצור אלגוריתם למידת מכונה, שיגנר את המודל **9.214** איור) על בסיס הבנה מוקדמת של הבעיה ונתונים היסטוריים, שעשויים להיות לא שלמים

בדוגמת בעיית קביעת המחיר, למידת מכונה מאפשרת ליצור סוגים שונים של דגמים מתודולוגיים, שאינם דורשים ידע על המנגנון המדויק של קביעת העלות. המודל "לומד" מנתונים על פרויקטים קודמים, מתאמן על תבניות אמיתיות בין פרמטרי הבניינים, עלותם וזמני הביצוע



בניגוד להערכה קלאסית לפי נוסחאות, האלגוריתם של למידת מכונה מתאמן על נתונים היסטוריים **9.214** איור

כל פרויקט במערך הנתונים האימוני, (supervised machine learning) בהקשר של למידת מכונה תחת פיקוח מכיל גם את התכונות הנכנסות (למשל, נתוני עלות וזמן בנייה של בניינים דומים, וגם את הערכים המיועדים) למשל ככל שמערך הנתונים **9.215** איור) עלות או זמן. (מערך נתונים כזה משמש ליצירה והתאמה של מודל למידת מכונה -גדול יותר ואיכות הנתונים בו גבוהה יותר, כך המודל יהיה מדויק יותר והתוצאות יהיו מדויקות יותר



שהוכשר על נתונים על עלות ולוח זמנים של פרויקטים קודמים, יקבע עם סבירות מסוימת את העלות, **מודל ML 9.215** איור ולוח הזמנים של פרויקט חדש.

לאחר יצירת והכשרת המודל להערכת בניית פרויקט חדש, מספיק לספק למודל תכונות חדשות עבור הפרויקט החדש, והמודל עם סבירות מסוימת יספק תוצאות חישוביות, המבוססות על תכונות שנלמדו בעבר.

מונחים מרכזיים בלמידת מכונה

למידת מכונה אינה קסם, אלא מתמטיקה, נתונים וחיפוש תכונות. היא אינה מחזיקה באינטליגנציה אמיתית, אלא מהווה תוכנה, שהוכשרה על נתונים כדי לזהות תכונות ולקבל החלטות ללא מעורבות מתמדת של אדם.

:- (9.216 איור) כדי לתאר את המבנה שלה, למידת מכונה משתמשת במגוון מונחים מרכזיים

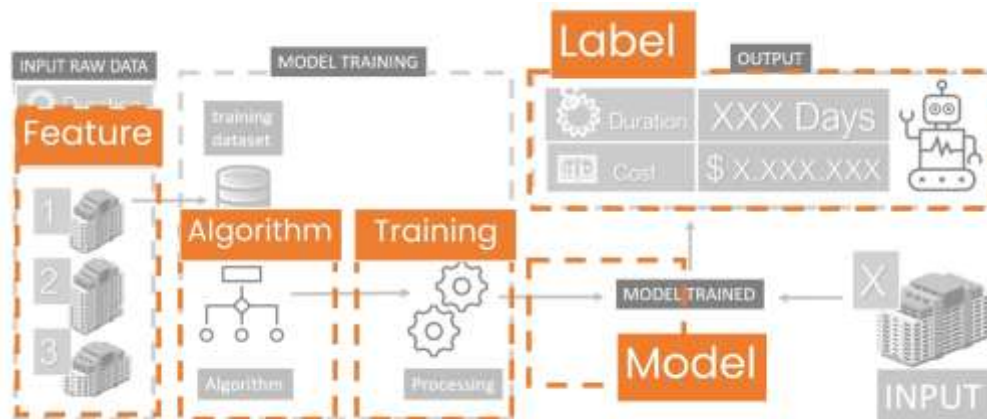
■ אלו משתנים מטרה או תכונות (הפרמטר "הישרדות" בנתוני הטיטניק, (אשר המודל צריך - (Labels) תוויות לחזות. דוגמה: עלות הבנייה) למשל, בדולרים, (משך זמן הבנייה) למשל, בחודשים)

■ אלו משתנים בלתי תלויים או תכונות, שממששות כקלט למודל. במודל חיזוי, הן - (Features) תכונות משמשות לחיזוי התוויות. דוגמאות: שטח המגרש) במטרים רבועים, (מספר הקומות בבניין, שטח הבניין הכולל) במטרים רבועים, (מיקום גיאוגרפי) קו רוחב וקו אורך, (סוג החומרים שנעשה בהם שימוש בבנייה מספר התכונות גם קובע את ממדי הנתונים

■ זהו אוסף של השערות שונות, אחת מהן מתקרבת לפונקציית המטרה, אותה יש לחזות או - (Model) מודל לאפיין. דוגמה: מודל למידת מכונה, המשתמש בשיטות ניתוח רגרסיה לחיזוי עלות וזמן הבנייה

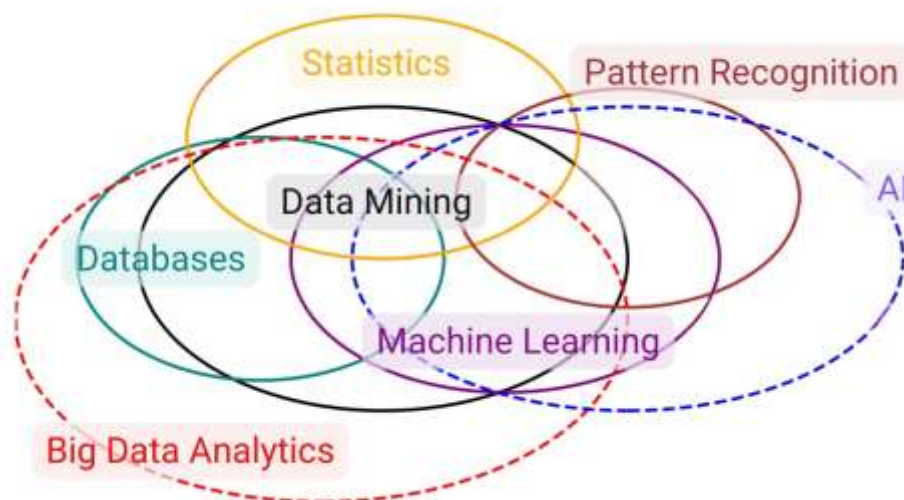
■ זהו תהליך חיפוש ההשערה הטובה ביותר במודל, שמתאימה - (Learning Algorithm) אלגוריתם למידה או יער KNN, בדיוק לפונקציית המטרה, תוך שימוש באוסף נתוני אימון. דוגמה: אלגוריתם רגרסיה ליניארית אקראי, המנתח נתונים על עלות וזמן הבנייה כדי לחשוף קשרים ודפוסים

■ במהלך תהליך האימון, האלגוריתם מנתח את נתוני האימון, מוצא דפוסים שמתאימים - (Training) אימון לקשר בין התכונות הקלטיות לתוויות המטרה. תוצאת תהליך זה היא מודל למידת מכונה מאומן, מוכן לחזות דוגמה: תהליך שבו האלגוריתם מנתח נתונים היסטוריים על בנייה) עלות, זמן, תכונות הנכס (כדי ליצור מודל חיזוי



למידת מכונה משתמשת בתוויות ובתכונות כדי ליצור מודלים, אשר מאומנים על נתונים באמצעות אלגוריתמים 9.216. ריש לחיזוי תוצאות.

למידת מכונה אינה קיימת בבידוד, אלא היא חלק מאקוסיסטמה רחבה יותר של דיסציפלינות אנליטיות, הכוללות סטטיסטיקה, מסדי נתונים, ניתוח נתונים אינטלקטואלי, זיהוי תבניות, אנליטיקה של נתונים גדולים ובינה מלאכותית מדגים כיצד תחומים אלו חופפים ומשלימים זה את זה, יוצרים בסיס מקיף למערכות קבלת החלטות 9.217. ריש - ואוטומציה מודרניות



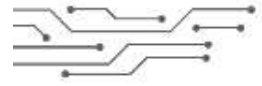
הקשרים בין תחומים שונים של ניתוח נתונים: סטטיסטיקה, למידת מכונה, בינה מלאכותית, נתונים גדולים, זיהוי 9.217. ריש תבניות וניתוח נתונים אינטלקטואלי.

המטרה העיקרית של למידת מכונה היא להעניק למחשבים את היכולת לרכוש ידע באופן אוטומטי ללא התערבות או עזרה של בני אדם ולתקן את פעולותיהם בהתאם

כך, בעתיד, תפקיד האדם יהיה רק לספק למכונה יכולות קוגניטיביות - הוא יקבע את התנאים, המשקלים והפרמטרים ומודל למידת המכונה יעשה את כל השאר

בפרק הבא נבחן דוגמאות ספציפיות ליישום אלגוריתמים. בטבלאות אמיתיות ובמודלים פשוטים יראו כיצד נבנה חיזוי

שלב אחר שלב.



פרק 9.3.

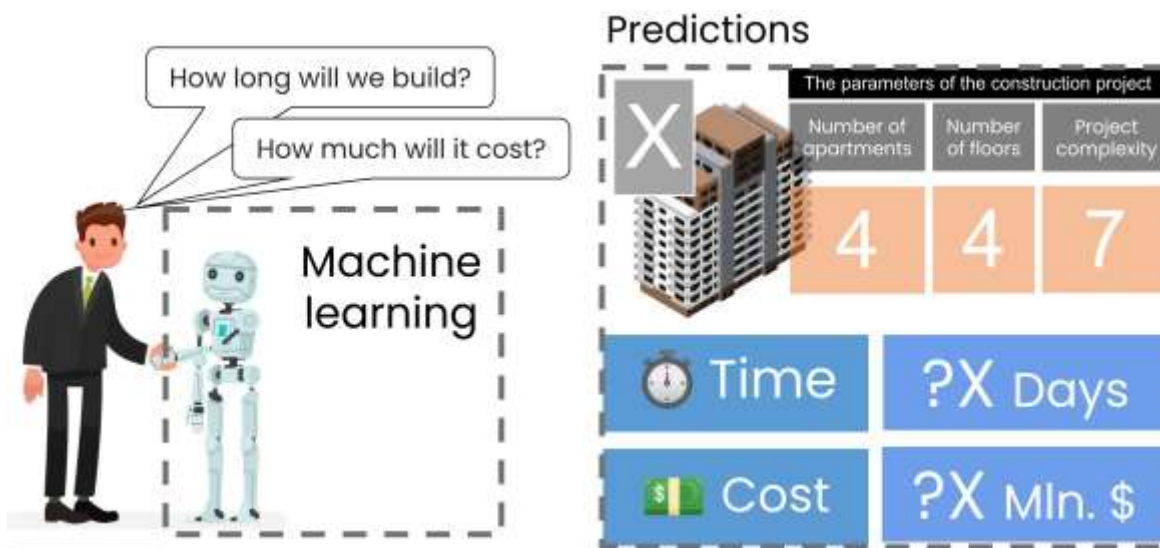
חיזוי עלות וזמן באמצעות למידת מכונה

דוגמה לשימוש בלמידת מכונה לקביעת עלויות וזמנים של פרויקט

הערכת זמני ועלויות הבנייה היא אחד התהליכים המרכזיים בפעילות של חברת בנייה. באופן מסורתי, הערכות אלו מתבצעות על ידי מומחים בהתבסס על ניסיון, מדריכים ומסדי נתונים נורמטיביים. עם זאת, בתנאים של טרנספורמציה כדי לשפר את הדיוק (ML) דיגיטלית ועלייה בזמינות הנתונים, מתאפשרת השימוש במודלים של למידת מכונה ואוטומציה של הערכות אלו.

יישום למידת מכונה בתהליך חישוב העלויות והזמנים של הבנייה מאפשר לא רק לשפר את היעילות של התכנון, אלא גם מהווה נקודת מוצא לשילוב מודלים אינטליגנטיים בתהליכים עסקיים אחרים - מניהול סיכונים ועד אופטימיזציה של לוגיסטיקה ורכישות.

חשוב לדעת במהירות כמה זמן ידרוש הבנייה של פרויקט ומה תהיה העלות הכוללת שלו. שאלות אלו לגבי זמני ועלויות הפרויקט תופסות באופן מסורתי מקום מרכזי במחשבה של לקוחות וחברות בנייה מאז ומעולם.



בפרויקטי בנייה, גורמי הצלחה מרכזיים הם מהירות ואיכות הערכת זמני ועלויות הבנייה 9.31. ריש.

בדוגמה הבאה יופקו נתונים מרכזיים מפרויקטים קודמים, ועל בסיסם תפותח מודל למידת מכונה, אשר יאפשר לנו בעזרת (9.31. ריש) מודל זה להעריך את העלות והזמנים של פרויקטי בנייה חדשים עם פרמטרים חדשים

נבחן שלושה פרויקטים עם שלושה מאפיינים מרכזיים: מספר דירות) כאשר 100 דירות שוות ל-10 לצורך הפשטה) מספר קומות ומדד מורכבות בנייה בסולם של 1 עד 10, כאשר 10 הוא המדד הגבוה ביותר של מורכבות. בלמידת "מכונה, תהליך ההמרה והפשטה של ערכים כמו 100 ל-10 או 50 ל-5 נקרא "נורמליזציה

נורמליזציה בלמידת מכונה היא תהליך של הבאת נתונים מספריים שונים לאותו סולם כדי להקל על עיבודם וניתוחם. תהליך זה חשוב במיוחד כאשר הנתונים נמצאים בסולמות וביחידות מידה שונות.

היו 50 דירות (לאחר נורמליזציה - 5, 7) קומות והערכת מורכבות 2, מה שציין (9.32. ריש) נניח שבפרויקט הראשון בנייה יחסית פשוטה. בפרויקט השני היו כבר 80 דירות, 9 קומות ופרויקט יחסית מורכב. בתנאים אלו, הבנייה של הבית המשותף הראשון והשני ארכה 270 ו-330 ימים, והעלות הכוללת של הפרויקט הייתה 4.5 ו-5.8 מיליון דולר בהתאמה.

The parameters of the construction project			The key parameters of the project	
Construction project	Number of apartment	Number of floors	Project complexity	
	100 = 100 apt		Multi-story	
1	5	7	2	270
2	8	9	6	330
3	3	5	3	230
X	4	4	7	?X
				\$ 4.502.000
				\$ 5.750.000
				\$ 3.262.000
				\$?X. XXX.XXX

X דוגמת סט פרויקטים קודמים שישמשו להערכת הזמן והעלות של פרויקט עתידי 9.32. ריש.

בבניית מודל למידת מכונה עבור נתונים כאלה, המשימה העיקרית היא לקבוע את המאפיינים הקריטיים (או התוויות) לחיזוי, במקרה זה - זמן ועלות הבנייה. עם סט נתונים קטן, נשתמש במידע על פרויקטי בנייה קודמים כדי לתכנן חדשים בהתבסס על X באמצעות אלגוריתמים של למידת מכונה, עלינו לחזות את העלות והמשך הבנייה של פרויקט חדש מאפיינים שניתנים של הפרויקט החדש, כמו 40 דירות, 4 קומות ומורכבות יחסית גבוהה של הפרויקט - 7. בתנאים אמיתיים, מספר הפרמטרים הנכנסים יכול להיות הרבה יותר גדול - מעשרות ועד מאות גורמים. אלה עשויים לכלול: סוג, חומרי הבניין, אזור אקלימי, רמת הכשרה של קבלנים, זמינות של רשתות הנדסיות, סוג היסוד, עונה להתחלת העבודה - 'הערות מנהלי העבודה וכו'.

כדי ליצור מודל חיזוי של למידת מכונה, עלינו לבחור אלגוריתם לבנייתו. אלגוריתם בלמידת מכונה הוא כמו מתכון מתמטי, המלמד את המחשב כיצד לחזות (לערבב בסדר הנכון את הפרמטרים) (או לקבל) החלטות בהתבסס על נתונים.

כדי לנתח נתונים על פרויקטי בנייה קודמים ולחזות את הזמן והעלות של פרויקטים עתידיים, ניתן להשתמש באחד מהאלגוריתמים הפופולריים של למידת מכונה -

■ רגרסיה ליניארית: אלגוריתם זה מנסה למצוא תלות ישירה בין המאפיינים, לדוגמה, בין מספר הקומות לעלות, הבנייה. המטרה של האלגוריתם היא למצוא משוואה ליניארית שמתארת בצורה הטובה ביותר את הקשר הזה.

מה שמאפשר לבצע חיזויים

שכנים הקרובים: אלגוריתם זה משווה את הפרויקט החדש עם פרויקטים קודמים שהיו דומים- k אלגוריתם מספר (דוגמאות לימודיות k מסווג את הנתונים בהתבסס על אילו מתוך k -NN. בגודלם או במורכבותם השכנים הקרובים k נמצאות הכי קרוב אליהם. בהקשר של רגרסיה, התוצאה היא ממוצע או חציון מתוך

עצי החלטה: זהו מודל חיזוי המפריד את הנתונים לתת-קבוצות, בהתבסס על תנאים שונים, תוך שימוש במבנה עץ. כל צומת בעץ מייצג תנאי או שאלה, המובילה להפרדה נוספת של הנתונים, וכל עלה - חיזוי או תוצאה סופית. האלגוריתם מחלק את הנתונים לקבוצות קטנות יותר בהתבסס על תכונות שונות, לדוגמה, קודם כל לפי מספר הקומות, ואז לפי מורכבות וכו', כדי לבצע חיזוי

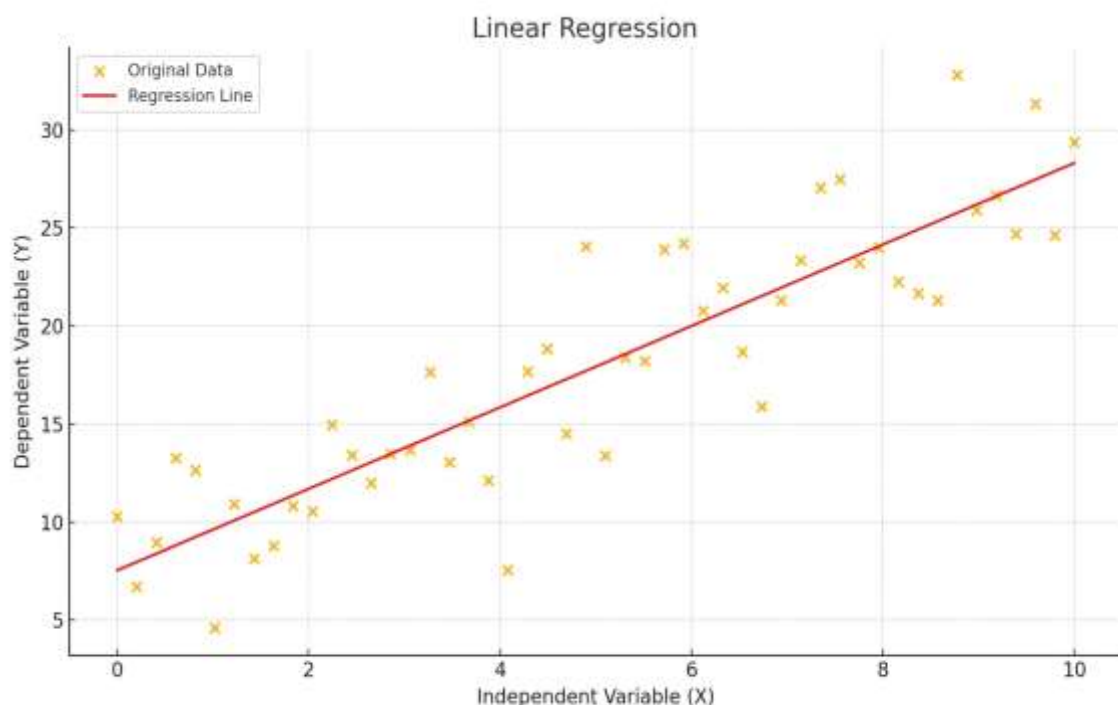
בואו נבחן את אלגוריתמי למידת המכונה להערכת עלות פרויקט חדש על בסיס שני אלגוריתמים פופולריים: רגרסיה שכנים הקרובים- k ליניארית ואלגוריתם

תחזית עלות זמן הפרויקט באמצעות רגרסיה ליניארית

רגרסיה ליניארית היא אלגוריתם בסיסי לניתוח נתונים, המאפשר לחזות את ערך המשתנה בהתבסס על קשר ליניארי עם משתנים אחרים. מודל זה מניח שיש קשר ליניארי ישיר בין המשתנה התלוי לבין אחד או יותר משתנים בלתי תלויים. והמטרה של האלגוריתם היא למצוא קשר זה

הפשטות והבהירות של רגרסיה ליניארית הפכו אותה לכלי פופולרי בתחומים שונים. כאשר עובדים עם משתנה אחד, רגרסיה ליניארית כרוכה במציאת הקו המתאים ביותר שעובר דרך נקודות הנתונים

למשתנה היוצא X רגרסיה ליניארית מוצאת את הקו הטוב ביותר (הקו האדום) שמאפר את התלות בין המשתנה הנכנס - בהתבסס על התלות הליניארית שנחשפה) איור 9.33 X עבור ערכים חדשים של Y קו זה מאפשר לחזות ערכים של Y .



איור 9.33 עקרון הפעולה של רגרסיה ליניארית במציאת הקו הטוב ביותר שיעבור דרך ערכי האימון.

מניבה את הערך החזוי של המשתנה התלוי (X) קו זה מוצג על ידי משוואה, שבה הכנסת ערך של המשתנה הבלתי תלוי תוך שימוש בתלות הליניארית ביניהם, X בהתבסס על ערכים ידועים של Y תהליך זה מאפשר לחזות ביעילות את (Y). דוגמה למציאת קו ממוצע סטטיסטי כזה ניתן לראות בדוגמת הערכת נתונים על היתרי בנייה בסן פרנסיסקו (איור 9.17) שבו חישבה הרגרסיה הליניארית את האינפלציה עבור סוגים שונים של נכסים.

ונבקש ממנו לבנות עבורנו LLM-בואו נטען את טבלת הנתונים של הפרויקט (איור 9.32 מהפרק הקודם) (ישירות ל מודל פשוט של למידת מכונה

של LLM (ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN: שלחו בקשה טקסטואלית בצ'אט של

X יש להציג את בניית מודל פשוט של למידת מכונה לחיזוי עלות זמן ביצוע של פרויקט חדש - (איור 9.32 כהדמיה מצורפת)

LLM: תשובת

Create in LLM Chat



```

1 from sklearn.linear_model import LinearRegression
2 import numpy as np
3
4 # Data from the image
5 # Features: Number of apartments, number of floors, project complexity
6 X = np.array([
7     [5, 7, 2],
8     [8, 9, 6],
9     [3, 5, 3]
10 ])
11
12 # Targets: Time (days), Cost ($)
13 y_time = np.array([270, 330, 230])
14 y_cost = np.array([4502000, 5750000, 3262000])
15
16 # Creating linear regression models for time and cost
17 model_time = LinearRegression()
18 model_cost = LinearRegression()
19
20 # Training the models
21 model_time.fit(X, y_time)
22 model_cost.fit(X, y_cost)
23
24 # The new project's features
25 X_new = np.array([[4, 4, 7]])
26
27 # Predicting the time and cost for the new project
28 predicted_time = model_time.predict(X_new)
29 predicted_cost = model_cost.predict(X_new)
30
31 predicted_time, predicted_cost

```

Result

(array([238.4444444]), array([3042337.77777778]))

Run in IDE



9.34 בחר ברגרסיה ליניארית ליצירת מודל למידת מכונה לחיזוי עלות זמן הפרויקט LLM איור

זיהה אוטומטית את הטבלה מהתמונה המצורפת והמיר את הנתונים מהפורמט הוויזואלי למערך טבלה) איור 9.34 LLM שורה 6. (מערך זה שימש כבסיס ליצירת תכונות ותוויות, שעל בסיסן נוצר מודל למידת מכונה) איור 9.34 – שורות – --שבו נעשה שימוש ברגרסיה ליניארית, (22-17

בעזרת מודל רגרסיה ליניארית בסיסי, שהוכשר על "קבוצת נתונים קטנה מאוד, "נעשו חיזויים עבור פרויקט בנייה במשימה שלנו, פרויקט זה מתאפיין בנוכחות של 40 דירות, 4 קומות ורמת קושי Project X-היפותטי חדש, המיוצג כ

(איור 9.32) 7.

Project בהתאם לחיזויים שנעשו בעזרת מודל הרגרסיה הליניארית, בהתבסס על קבוצת נתונים מוגבלת וקטנה עבור
- החדש (איור 9.34 - שורות 24-29) X

■ (238.4444444) משך הבנייה יהיה כ-238 ימים

■ (3042337.777) הסכום הכולל של ההוצאות יהיה כ-\$3,042,338

כדי להמשיך לחקור את ההשערה לגבי עלות הפרויקט, יהיה מועיל להתנסות באלגוריתמים ובשיטות שונות של למידת
החדש בהתבסס על קבוצת נתונים היסטורית קטנה Project X מכונה. לכן, נחזות באותם ערכים של עלות זמן עבור
K-Nearest Neighbours (k-NN) בעזרת אלגוריתם

K-nearest neighbor (k-NN) תחזיות עלות זמן הפרויקט בעזרת אלגוריתם

K-Nearest Neighbours (k-NN). כהערכה נוספת להערכת עלות ומשך הפרויקט החדש, נשתמש באלגוריתם
הוא שיטה של למידת מכונה מפוקחת, המיועדת הן לסיווג והן לרגרסיה K-Nearest Neighbors (k-NN) אלגוריתם
כבר נדון על ידינו קודם לכן בהקשר של חיפוש בבסיסי נתונים וקטוריים, שם הוא משמש k-NN כמו כן, אלגוריתם
למציאת הווקטורים הקרובים ביותר (למשל, טקסטים, תמונות או תיאורים טכניים). (בגישה זו, כל פרויקט מיוצג כנקודה
- במרחב רב-ממדי, כאשר כל מימד תואם לאטריבוט מסוים של הפרויקט

במקרה שלנו, בהתחשב בשלושה אטריבוטים של כל פרויקט, נציג אותם כנקודות במרחב תלת-ממדי. כך, הפרויקט הקרוב
יש לציין כי בתנאים אמיתיים, מספר הנקודות וממדי. $(x=4, y=4, z=7)$ יתמקם במרחב הזה עם קואורדינטות X שלנו
המרחב עשויים להיות גבוהים בהרבה.

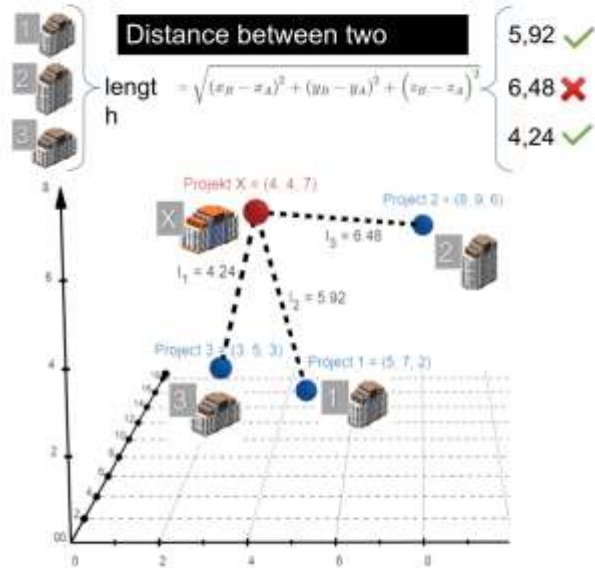
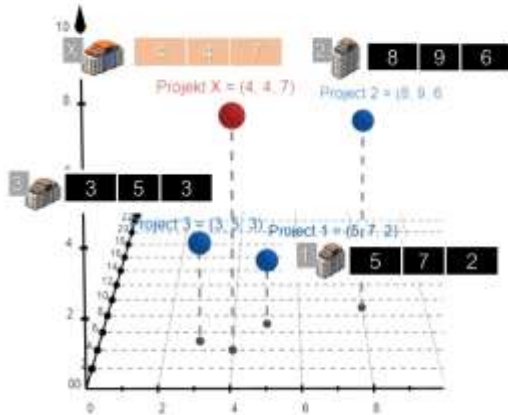
לבין X פועל על ידי מדידת המרחק בין הפרויקט הרצוי K-NN (k-nearest neighbors) אלגוריתם
הפרויקטים בבסיס הנתונים האימוני. בהשוואת המרחקים הללו, האלגוריתם קובע אילו פרויקטים נמצאים
X. הכי קרוב לנקודת הפרויקט החדש

מאשר פרויקטים, X-מתוך מאגר הנתונים שלנו ממוקם רחוק משמעותית מ $(x=8, y=9, z=6)$ לדוגמה, אם הפרויקט השני
הפרויקטים הקרובים ביותר, על $(k=2)$ אחרים, ניתן להוציא אותו מהניתוח הבא. כתוצאה מכך, ניתן להשתמש רק בשני
בסיסם ייקבע הערך הממוצע.

שיטה דומה, דרך חיפוש שכנים, מאפשרת להעריך את הדמיון בין הפרויקטים, מה שעוזר להסיק מסקנות לגבי העלות
האפשרית והזמנים להשלמת הפרויקט החדש בהתבסס על פרויקטים דומים שבוצעו בעבר

k-nearest neighbors algorithm

The KNN algorithm assumes that similar things exist in close proximity. In other words, similar things are near to each other.



הפרויקטים מיוצגים כנקודות במרחב רב-ממדי, ולצורך הערכת הדמיון וחיזוי, נבחרים הפרויקטים הקרובים ב-K-NN, במערכת ביותר בהתבסס על המרחקים.

כוללת מספר שלבים מרכזיים k-NN עבודת:

- הכנת הנתונים: תחילה טוענים את מערכי הנתונים האימוניים והבדיקות. הנתונים האימוניים משמשים לאימון "של האלגוריתם, והנתונים הבדיקים - לבדוק את היעילות שלו".
- המצביע על כמה שכנים קרובים (נקודות נתונים) יש לקחת בחשבון K, נבחר מספר K: בחירת פרמטר הוא מאוד חשוב, מכיוון שהוא משפיע על התוצאה "K" באלגוריתם. ערך.
- תהליך הסיווג והרגרסיה עבור נתוני הבדיקה:
 - חישוב מרחקים: עבור כל אלמנט מנתוני הבדיקה, מחושב המרחק לכל אלמנט מנתוני האימון. לשם כך ניתן להשתמש בשיטות שונות למדידת מרחקים, כגון מרחק אוקלידי (השיטה הנפוצה ביותר, (מרחק מנהטן או מרחק המינג.
 - הנקודות הקרובות K השכנים הקרובים ביותר: לאחר חישוב המרחקים, הם ממין ונבחרים K מיון ובחירת ביותר לנקודת הבדיקה.
 - הגדרת הקטגוריה או הערך של נקודת הבדיקה: אם זו משימת סיווג, הקטגוריה של נקודת הבדיקה מוגדרת השכנים הנבחרים. אם זו משימת רגרסיה, אז מחושבת הממוצע K על סמך הקטגוריה הנפוצה ביותר בין (השכנים K) או מדד אחר של נטייה מרכזית (של הערכים של
- סיום התהליך: ברגע שכל נתוני הבדיקה סווגו או נעשו עבורם תחזיות, התהליך יושלם.

יעיל בהרבה יישומים מעשיים ונחשב לאחד הכלים הבסיסיים (k-NN) השכנים הקרובים K אלגוריתם בארסנל של מומחי למידת מכונה. אלגוריתם זה פופולרי בזכות הפשטות והיעילות שלו, במיוחד במשימות שבהן הקשרים בין הנתונים קלים לפרשנות.

השכנים הקרובים, (זוהו שני פרויקטים) ממדגם הקטן שלנו (עם המרחק הקטן K בדוגמה שלנו, לאחר יישום אלגוריתם, על סמך פרויקטים אלה, האלגוריתם קובע את הממוצע של מחירים וזמן הבנייה שלהם. לאחר ניתוח X ביותר מהפרויקט יעלה כ-3,800,000 דולר ויימשך X האלגוריתם, באמצעות חישוב ממוצע השכנים הקרובים, מגיע למסקנה כי הפרויקט כ-250 ימים.

k-nearest neighbors algorithm



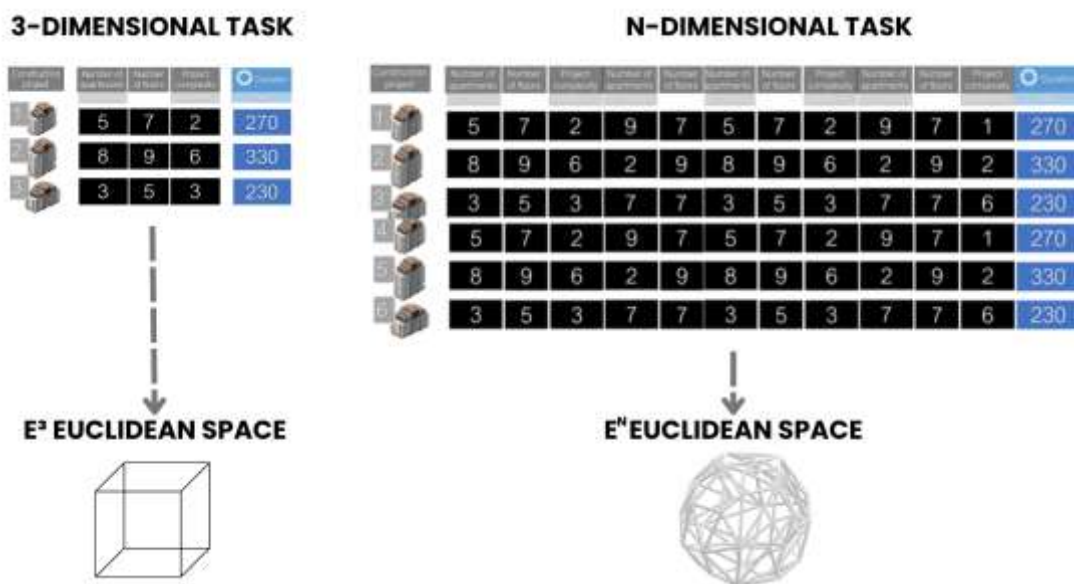
תוך ניתוח שני הפרויקטים הקרובים ביותר X, השכנים הקרובים קובע את העלות ואת לוח הזמנים של הפרויקט-K אלגוריתם במדגם.

פופולרי במיוחד במשימות סיווג ורגרסיה, לדוגמה, במערכות המלצה, שבהן (k-NN) השכנים הקרובים K אלגוריתם נמצא k-NN, הוא משמש להמלצה על מוצרים או תוכן על סמך העדפות הדומות לאינטרסים של משתמש ספציפי. בנוסף בשימוש נרחב באבחון רפואי לסיווג סוגי מחלות על סמך תסמינים של המטופל, בזיהוי דפוסים ובסקטור הפיננסי להערכת יכולת ההחזר של לקוחות.

גם כאשר יש נתונים מוגבלים, מודלים של למידת מכונה יכולים לספק תחזיות מועילות ולחזק באופן משמעותי את המרכיב האנליטי בניהול פרויקטי בנייה. עם הרחבה וניקוי של נתונים היסטוריים, ניתן לעבור למודלים מורכבים יותר - לדוגמה, תוך התחשבות בסוגי מבנים, מיקום, עונה של תחילת הבנייה וגורמים נוספים.

במשימה הפשוטה שלנו, לצורך ויזואליזציה במרחב תלת-ממדי, נעשה שימוש בשלושה מאפיינים, אך פרויקטים אמיתיים

כוללים בממוצע מאות או אלפי מאפיינים, מה שמגדיל משמעותית את ממדי המרחב ואת מורכבות הצגת הפרויקטים vectors.-



בדוגמה הפשוטה, לצורך ויזואליזציה תלת-ממדית נעשה שימוש בשלושה מאפיינים, בעוד שפרויקטים אמיתיים כוללים מספר רב יותר.

שבו 40 דירות, 4 קומות ורמת מורכבות 7, הניב X יישום אלגוריתמים שונים על אותו מערך נתונים עבור הפרויקט ערכי תחזית שונים. אלגוריתם הרגרסיה הליניארית חזה זמן סיום של 238 ימים ועלות של 3,042,338 דולר, בעוד --חזה 250 ימים ו-3,882,000 דולר k-NN שאלגוריתם

דיוק התחזיות המתקבלות באמצעות מודלים של למידת מכונה תלוי ישירות בנפח ובאיכות הנתונים המקוריים. ככל שיותר פרויקטים משתתפים בהכשרה, וככל שהמאפיינים (תכונות) (והתוצאות) תיוגים (שלהם מוצגים בצורה מלאה ומדויקת יותר, כך עולה הסבירות לקבל תחזיות מהימנות עם ערכי שגיאה מינימליים

תפקיד חשוב בתהליך זה ממלאים שיטות עיבוד מקדים של נתונים, כולל

- נורמליזציה המאפשרת להביא את המאפיינים לאותו סולם
- זיהוי והסרת חריגות, המונעות עיוות של המודל
- קידוד תכונות קטגוריאליות, המאפשר עבודה עם נתוני טקסט
- מילוי ערכים חסרים, המגביר את עמידות המודל

בנוסף, להערכת יכולת הכללה של המודל ועמידותו בפני מערכי נתונים חדשים נעשה שימוש בשיטות של תיקוף צולב. המאפשרות לחשוף התאמה יתרה ולשפר את מהימנות התחזיות

אפילו אם נראה לכם שהכאוס של המשימות שלכם בלתי ניתן לתיאור פורמלי, ידעו - כל אירוע בעולם ובמיוחד תהליכי הבנייה כפופים לחוקי מתמטיקה, שעשויים לדרוש תמיכה בחישוב ערכים לא דרך נוסחאות מחמירות אלא בעזרת סטטיסטיקה ונתונים היסטוריים.

כפי שחישובי עלויות מסורתיים, המבוצעים על ידי מחלקות חישוב, מתמודדים עם חוסר ודאות ומקורות פוטנציאליים לטעויות, כך גם מודלים של למידת מכונה. עם זאת, כאשר ישנם נתונים איכותיים בכמות מספקת, מודלים של למידת מכונה יכולים להציג דיוק דומה, ולעיתים אף גבוה יותר, בהשוואה להערכות של מומחים.

למידת מכונה, ככל הנראה, תהפוך לכלי עזר אמין לניתוח, המאפשר: לחדד חישובים, להציע תרחישים, חלופיים ולגלות תלות נסתרת בין פרמטרים של פרויקט. מודלים כאלה לא יתיימרו להיות אוניברסליים אך כבר בקרוב יוכלו לתפוס מקום חשוב בחישובים ובתהליכי קבלת החלטות פרויקטיות. טכנולוגיות למידת מכונה לא יחריגו את מעורבותם של מהנדסים, אומדנים ואנליסטים, אלא, להיפך, ירחיבו את יכולותיהם, תוך הצעת נקודת מבט נוספת, המבוססת על נתונים היסטוריים.

ביישום נכון של למידת מכונה בתהליכי העבודה של חברות הבנייה, יש לה פוטנציאל להפוך לרכיב חשוב במערכת תמיכת ההחלטות הניהוליות - לא כתחליף לאדם, אלא כהרחבה של האינטואיציה המקצועית וההיגיון ההנדסי שלו.

צעדים נוספים: מאחסון לניתוח ולחיזוי

גישות מודרניות לעבודה עם נתונים מתחילות לשנות את העקרונות לקבלת החלטות בתעשיית הבנייה. המעבר מהערכות אינטואיטיביות לניתוח אובייקטיבי של נתונים לא רק מעלה את הדיוק, אלא גם פותח הזדמנויות חדשות לאופטימיזציה של תהליכים. לסיכום חלק זה, יש להדגיש את הצעדים המעשיים העיקריים שיכולים לסייע ביישום השיטות שנדונו במשימות היומיומיות שלכם:

■ יצירת תשתית אחסון נתונים בת קיימא

☐ נסו לאחד מסמכים מפוזרים ונתוני פרויקטים למודל טבלאי אחד, תוך כדי אגרגציה של מידע מפתח במסגרת דאטה פריים לניתוח נוסף

☐ במקום Apache Parquet השתמשו בפורמטים יעילים לאחסון נתונים - לדוגמה, פורמטים עמודתיים כמו במיוחד עבור מערכות נתונים שעשויות לשמש בעתיד לאימון מודלים של למידת מכונה - XLSX או CSV

☐ צרו מערכת גרסאות לנתונים, המאפשרת לעקוב אחרי שינויים במהלך כל הפרויקט

■ הטמעת כלים לניתוח ואוטומציה

☐ התחילו לנתח נתונים היסטוריים של פרויקטים - לפי תיעוד, מודלים, אומדנים - כדי לחשוף תבניות, מגמות ואנומליות

☐ להעלאה אוטומטית והכנה של נתונים ETL (Extract, Transform, Load) שלטו בתהליכי

☐ שונות לוויזואליזציה Python למדו להציג מדדים מפתח באמצעות ספריות

☐ התחילו ליישם שיטות סטטיסטיות ודגימות אקראיות כדי לקבל מסקנות אנליטיות מייצגות וחוזרות

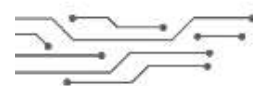
■ העלאת בשלות בעבודה עם נתונים

☐ למדו כמה אלגוריתמים בסיסיים של למידת מכונה על דוגמאות פשוטות וברורות, כמו מערך הנתונים של טיטניק

☐ נתחו את התהליכים הנוכחיים והגדירו היכן ניתן לעבור מלוגיקה סיבתית נוקשה לשיטות סטטיסטיות לחיזוי והערכה

☐ התחילו לראות בנתונים נכס אסטרטגי ולא תוצר לוואי: בנו תהליכי קבלת החלטות על בסיס מודלים של נתונים ולא סביב פתרונות תוכנה ספציפיים

חברות בנייה שהבינו את ערך הנתונים נכנסות לשלב חדש של התפתחות, שבו יתרון תחרותי מוגדר לא לפי כמות המשאבים, אלא לפי מהירות קבלת ההחלטות על בסיס אנליטיקה



מקסימום נוחות עם הגרסה המודפסת

לשימוש נוח יותר וגישה מהירה. Data-Driven Construction אתם מחזיקים בידיכם את הגרסה הדיגיטלית החינמית של לחומרים, אנו ממליצים לשקול את הגרסה המודפסת

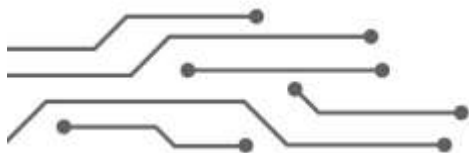


תמיד בהישג יד: ספר בפורמט מודפס יהפוך לכלי עבודה אמין, המאפשר למצוא ולהשתמש במהירות בויזואליזציות ובתרשימים בכל מצב עבודה איכות גבוהה של איורים: כל התמונות והגרפים בגרסה המודפסת מוצגים באיכות מקסימלית גישה מהירה למידע: ניווט

נוח, אפשרות לעשות הערות, סימניות ולעבוד עם הספר בכל מקום.

ברכישת הגרסה המודפסת המלאה של הספר, אתם מקבלים כלי נוח לעבודה נוחה ויעילה עם מידע: אפשרות להשתמש בחומרים ויזואליים במשימות יומיומיות, למצוא במהירות את התרשימים הנדרשים ולעשות הערות. בנוסף, רכישתכם תומכת בהפצת ידע פתוח.

datadrivenconstruction.io/books: ניתן להזמין את הגרסה המודפסת של הספר בכתובת



X חלק .תעשיית הבנייה בעידן הנתונים דיגיטליים

הזדמנויות ואתגרים

החלק העשירי והסופי מציע מבט כולל על עתיד תעשיית הבנייה בעידן של טרנספורמציה דיגיטלית. כאן נבחן המעבר מניתוח סיבתי לעבודה עם קורלציות של נתונים גדולים. נעשות, מקבילות בין התפתחות האומנות הוויזואלית לבין התפתחות העבודה עם נתונים בבנייה, המדגימות כיצד התעשייה מתקדמת מבקרת פרטים להבנה כוללת של תהליכים. נדונה הקונספט של "אובריזציה" של תעשיית הבנייה, שבה שקיפות הנתונים ואוטומציה של חישובים עשויות לשנות באופן רדיקלי מודלים עסקיים מסורתיים, תוך כדי חיסול הצורך במתווכים והפחתת אפשרויות לספקולציות. נדונות בעיות לא פתורות, כמו סיווג אוניברסלי של רכיבים, אשר נותנות לחברות הבנייה זמן להסתגל לתנאים החדשים. החלק מסתיים בהמלצות קונקרטיות לגבי גיבוש אסטרטגיית טרנספורמציה דיגיטלית, הכוללת ניתוח פגיעויות והרחבת מגוון השירותים לשמירה על תחרותיות בתעשייה המשתנה.

פרק 10.1.

אסטרטגיות הישרדות: יצירת יתרונות תחרותיים

קורלציות במקום חישובים: העתיד של אנליטיקה בבנייה

בשל הדיגיטציה המהירה של מידע, הבנייה המודרנית חווה טרנספורמציה יסודית, שבה הנתונים הופכים לא רק לכלי אלא לנכס אסטרטגי, המסוגל לשנות באופן דרמטי את הגישות המסורתיות לניהול פרויקטים ולעסקים.

במשך אלפי שנים, הפעילות הבנייתית התבססה על שיטות דטרמיניסטיות - חישובים מדויקים, פירוט ובקרת פרמטרים קפדנית. במאות הראשונות לספירה, מהנדסים רומיים השתמשו בעקרונות מתמטיים לבניית אקוודוקטים וגשרים. בימי הביניים, אדריכלים שאפו לפרופורציות אידיאליות של קתדרלות גותיות, ובעידן התיעוש של המאה ה-20, הוקמו מערכות של תקנים ורגולציות סטנדרטיים, שהפכו לבסיס לבנייה המונית.

כיום, כיוון ההתפתחות זו מהחיפוש אחר קשרים סיבתיים מדויקים לניתוח הסתברותי, חיפוש קורלציות ודפוסים סמויים הענף נכנס לשלב חדש - הנתונים הופכים למשאב מרכזי, והאנליטיקה המבוססת עליהם מחליפה גישות אינטואיטיביות וממוקדות מקומית.



פוטנציאל הנתונים הבנייתיים: החישובים הקיימים בחברה הם רק קצה הקרחון, הנגיש למנהלים לצורך ניתוח.

מערכת המידע של החברה דומה לקרחון: למנהלי החברה נראית רק חלק קטן מהפוטנציאל של הנתונים, בעוד שהערך העיקרי טמון בעמקי המידע. חשוב להעריך את הנתונים לא רק לפי השימוש הנוכחי שלהם, אלא גם לפי האפשרויות שהם יפתחו בעתיד. רק החברות שידעו לחלץ דפוסים סמויים וליצור ידע חדש מתוך הנתונים, יוכלו לבסס יתרון תחרותי בר קיימא.

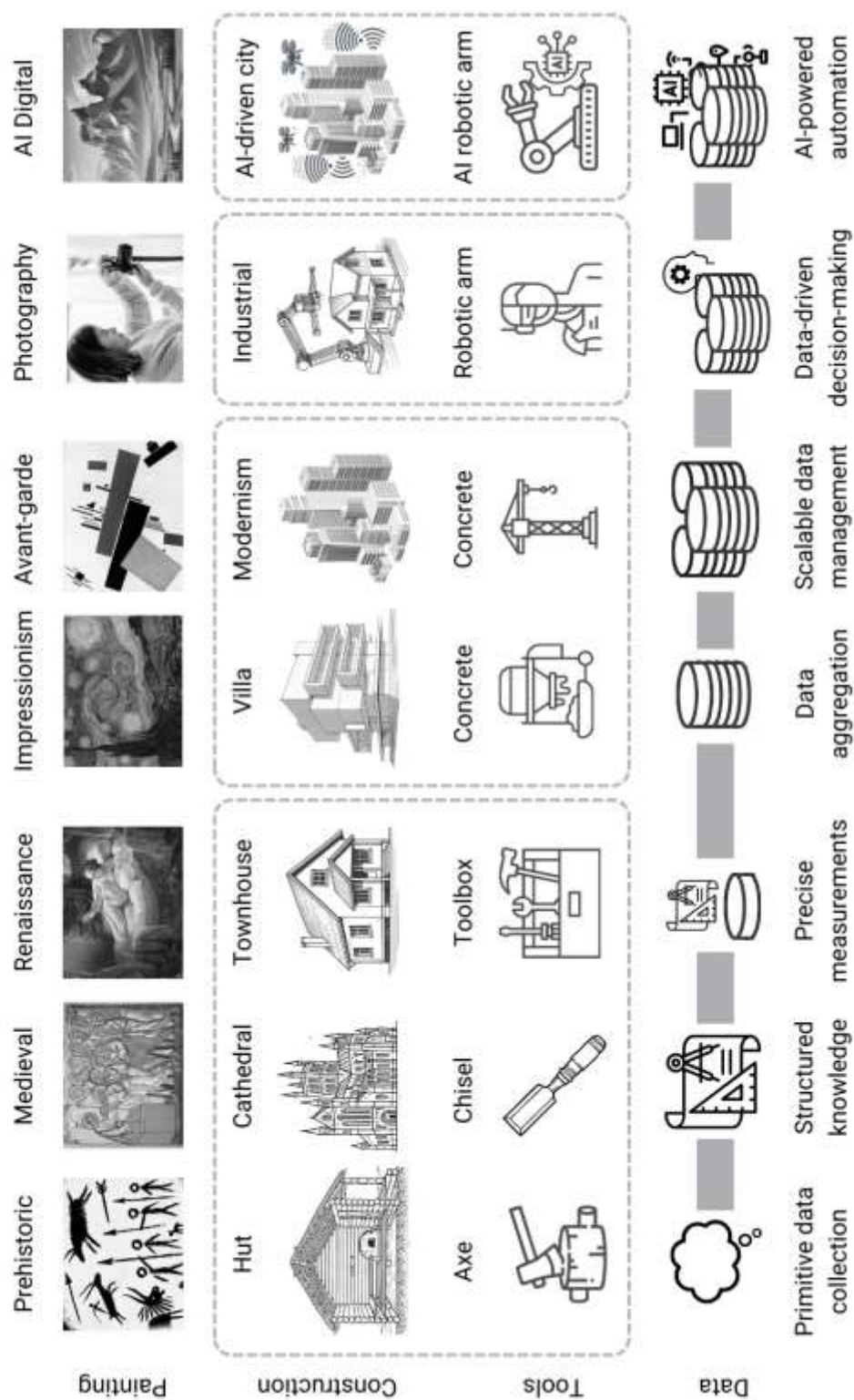
חיפוש דפוסים סמויים והבנת הנתונים אינם רק עבודה עם מספרים, אלא תהליך יצירתי, הדורש חשיבה מופשטת ויכולת לראות את התמונה הכוללת מאחורי אלמנטים מבודדים. במובן זה, התפתחות העבודה עם נתונים ניתנת להשוואה עם האבולוציה של האמנות הוויזואלית.

התפתחות הבנייה מזכירה באופן מפתיע את ההתקדמות של האמנות הוויזואלית. בשני המקרים, האנושות עברה דרך משיטות פרימיטיביות לטכנולוגיות מורכבות של ויזואליזציה וניתוח. בעידן הפרהיסטורי, אנשים השתמשו בציורים על קירות ובכלים פרימיטיביים כדי לפתור בעיות יומיומיות. בימי הביניים ובתקופת הרנסנס, רמת המורכבות בארכיטקטורה ובאמנות עלתה משמעותית. לקראת תחילת ימי הביניים, כלי הבנייה התפתחו מכלי פשוט כמו גרזן למערכות רחבות של כלים, המסמלים את עליית הידע הטכני.

עידן הריאליזם היה המהפכה הראשונה באמנות הוויזואלית: האמנים למדו לשחזר את הפרטים הקטנים ביותר, תוך השגת מקסימום אמינות בתמונה. בבנייה, המקבילה של תקופה זו הייתה שיטות הנדסה מדויקות, שרטוטים מפורטים וחישובים קפדניים, שהפכו לבסיס לפרקטיקה התכנונית במשך מאות שנים.

מאוחר יותר, האימפרסיוניזם שינה את תפיסת המציאות האמנותית: במקום העברה מילולית של הצורה, האמנים החלו לתעד את המצב רוח, האור והדינמיקה, במטרה לשקף את הרושם הכללי ולא את הדיוק האבסולוטי. באופן דומה, למידת מכונה באנליטיקה בנייה מתרחקת ממודלים לוגיים נוקשים לזיהוי תבניות וחוקי הסתברות, המאפשרים "לראות" תלות נסתרות בנתונים, שאינן נגישות בניתוח קלאסי. גישה זו מתכתבת עם רעיונות המינימליזם והפונקציונליות של באוהאוס, שבו המשמעות (הפונקציה) חשובה יותר מהצורה. באוהאוס שאף להסיר את המיותר, לוותר על קישוטים למען בהירות, שימושיות ומסחריות. הדברים היו צריכים להיות מובנים ושימושיים, ללא עודפים – האסתטיקה נולדה מתוך הלוגיקה של הבנייה והייעוד.

עם הופעת הצילום בסוף המאה ה-19, האמנות קיבלה כלי חדש לתיעוד המציאות בדיוק חסר תקדים והפכה את היחס לאמנות הוויזואלית. באופן דומה, בבנייה, המהפכה התעשייתית במאה ה-21 מביאה ליישום טכנולוגיות רובוטיות ורעיונות כמו "בנייה מחוברת", שבהם איסוף פרמטרים בודדים התפתח לתיעוד אינטליגנטי בקנה IoT, RFID, לייזרים מידה של המציאות המלאה של אתר הבנייה.



- תקופות האבולוציה של האמנות הוויזואלית מתאימות להתפתחות הגישות לעבודה עם נתונים בתעשיית הבנייה 10.12. ריש.

תעשיית הבנייה חווה, LLM ו-AI היום, בדיוק כמו שהאמנות הוויזואלית חווה מחדש את עצמה עם הגעת הכלים של

מאפשרות LLM-צ'אט (AI) קפיצה קוונטית נוספת: מערכות אינטליגנטיות, המנוהלות על ידי אינטליגנציה מלאכותית לחזות, ליעל וליצור פתרונות עם מינימום התערבות אנושית.

תפקיד הנתונים בתכנון ובניהול השתנה באופן רדיקלי. אם בעבר הידע הועבר בעל פה והיה בעל אופי אמפירי – בדומה לכך, עד המאה ה-19 המציאות תועדה בציורים שנכתבו ביד – היום במרכז תשומת הלב נמצאת התיעוד הדיגיטלי המלא של "תמונת" הבנייה. בעזרת אלגוריתמים של למידת מכונה, התמונה הדיגיטלית הזו מתמירה לתצוגה אימפרסיוניסטית של המציאות הבנייתית – לא העתק מדויק, אלא הבנה כללית, הסתברותית של התהליכים.

אנו מתקרבים במהירות לעידן שבו תהליכי התכנון, הבנייה וההפעלה של מבנים לא רק יתמכו, אלא גם ינוהלו במידה רבה על ידי מערכות אינטליגנציה מלאכותית. בדומה לכך, כיצד האמנות הדיגיטלית המודרנית נוצרת ללא מברשת – בעזרת הנחיות טקסטואליות ודגמים גנרטיביים – הפתרונות האדריכליים וההנדסיים של העתיד ייבנו על בסיס בקשות מרכזיות ופרמטרים שנקבעו על ידי המשתמש.

במאה ה-21, הגישה לנתונים, הפרשנות שלהם ואיכות האנליטיקה הופכים לתנאים בלתי נפרדים להצלחת הפרויקט. יתרה מכך, ערך הנתונים נקבע לא לפי כמותם, אלא לפי היכולת של המומחים לנתח, לבדוק ולהפוך אותם לפעולות.

גישה מבוססת נתונים בבנייה: תשתית ברמה חדשה

בהיסטוריה של האנושות, כל קפיצה טכנולוגית כזו הביאה לשינויים יסודיים בכלכלה ובחברה. היום אנו עדים לגל חדש של טרנספורמציה, שניתן להשוות אותו בהיקפו למהפכה התעשייתית של המאה ה-19. עם זאת, בעוד שלפני מאה שנה הכוח המניע העיקרי של השינויים היה כוחות מכניים וטכנולוגיות אנרגיה, כיום מדובר בנתונים ובבינה מלאכותית.

משנים את מהות היישומים עצמם, מה שהופך את ערכות התוכנה המסורתיות (שנדונו AI וסוכני LLM, למידת מכונה ולא בכללים עסקיים, AI בחלק השני של הספר) לבלתי נחוצות. כל הלוגיקה של עבודה עם נתונים מתרכזת בסוכני --מקודדים קשיחים

בעידן הנתונים, התפיסות המסורתיות של יישומים משתנות באופן דרמטי. אנו מתקדמים למודל שבו מערכות מודולריות ארגוניות כבדות יפנו את מקומן לפתרונות פתוחים, קלים ומיוחדים.

בעתיד תישאר רק התשתית הבסיסית של הנתונים, וכל האינטראקציה עימם תתרחש דרך סוכנים העובדים ישירות עם בסיס הנתונים. אני באמת מאמין שכל ערכת היישומים תיעלם, כי פשוט אין צורך בה, כאשר הבינה המלאכותית מתקשרת ישירות עם הקמתי חברות, עבדתי בהן, ואם להיות כנה, עכשיו כנראה לא הייתי - SaaS-בסיס הנתונים המרכזי. כל הקריירה שלי עבדתי ב כרגע. המצב לא ברור מדי. זה לא אומר שבעתיד לא SaaS חדש. וסביר להניח שלא הייתי משקיע בחברות SaaS מתחיל עסק יהיו חברות תוכנה, פשוט הן ייראו אחרת לגמרי. המערכות העתידיות יהיו בסיסי נתונים עם הלוגיקה העסקית שהועברה לסוכני סוכנים אלה יעבדו עם מספר מאגרי נתונים בו זמנית, מבלי להיות מוגבלים לבסיס אחד. כל הלוגיקה תעבור לרמת הבינה AI. המלאכותית. מתיו ברמן, מנכ"ל פורוורד פיוצ'ר

ההבדל המרכזי של הפרדיגמה החדשה הוא המינימליזם של העומס הטכנולוגי. במקום מערכות תוכנה מורכבות

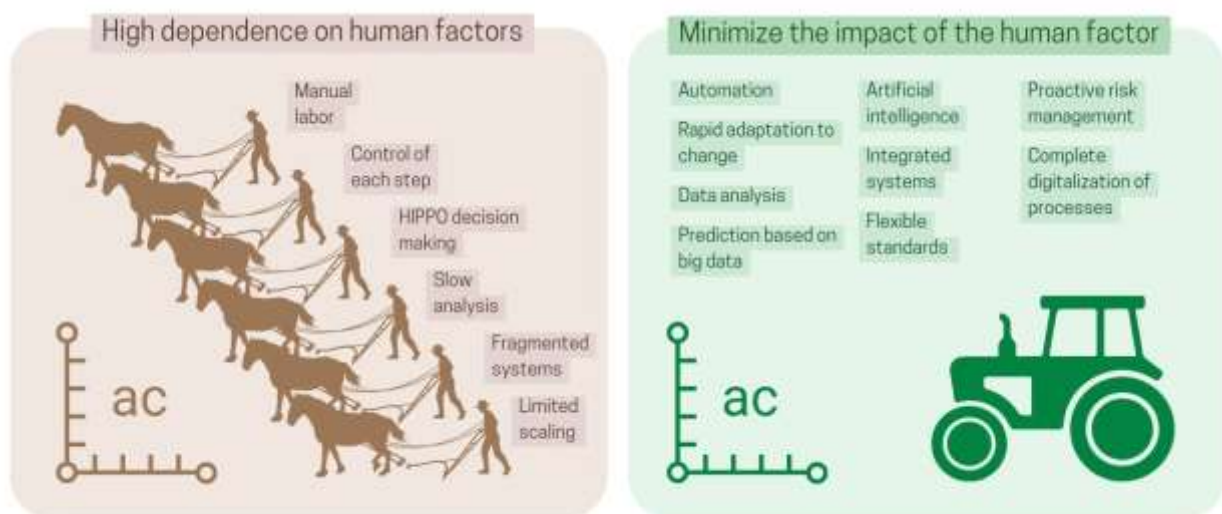
ומסיביות, נקבל מודולים גמישים, פתוחים ומהירים להתאמה, שחיים ממש בתוך זרם הנתונים. הארכיטקטורה של ERP, ניהול התהליכים בעתיד שימוש במיקרו-יישומים - כלים קומפקטיים וממוקדים, השונים באופן עקרוני ממערכות המסיביות והסגורות. הסוכנים החדשים יהיו מותאמים ככל האפשר, משולבים וממוקדים PMIS, CDE, CAFM. Low-Code/No-Code) למשל) במשימות עסקיות ספציפיות.

[יצירה, קריאה, עדכון ומחיקה] במספר CRUD והסוכנים הללו יבצעו פעולות [AI סוכני] כל הלוגיקה העסקית תעבור לאותם AI-ברמת מאגרי נתונים, כלומר הם לא יבחינו באיזה בקנה משתמשים. הם יעדכנו מספר בסיסי נתונים, וכל הלוגיקה תימצא תהפוך למקום שבו נמצאת כל הלוגיקה, אנשים יתחילו להחליף את הקנקנים. אנו כבר עדים לאחוז AI-הידועה. וכאשר רמת השימוש בסוכנים, ואנו נמשיך לפעול בכיוון זה, במטרה לאחד את כל זה Dynamics גבוה למדי של הצלחות בשוק הקנקנים. אלא גם הפתרונות שלנו בתחום הפיננסים והפעולות, CRM, בין אם זה בתחום שירות הלקוחות או בתחומים אחרים, כמו לא רק אחת AI. וסוכני AI שבהן רמת הלוגיקה יכולה להתנהל על ידי AI, כי אנשים רוצים יותר אפליקציות עסקיות ממוקדות, כלומר, מה שעשינו. עכשיו Copilot עם GitHub-שמשווה ל Python עם Excel מהדברים המרגשים ביותר עבורי היא ולהתחיל לשחק איתו. זה כבר לא רק הבנת המספרים Copilot כדאי פשוט לפתוח אותו, להפעיל את Excel כשיש לך יוצר תוכנית ולאחר מכן מבצע אותה, כך זה דומה GitHub Copilot Workspace-הקיימים - הוא יכין תוכנית בעצמו. כמו ש Excel-משתמש ב Copilot, ככלי ויזואליזציה של שורות ועמודות לניתוח. כך Excel-לעבודה של אנליסט נתונים שמשתמש ב סאטיה נדלה, מנכ"ל מיקרוסופט, ריאיון — Python. ככלי עם כל האפשרויות שלו, כי הוא יכול לייצר נתונים ויש לו מפרש דצמבר BG2 2024 לערוץ

העובדים ישירות AI השינוי שאנו עדים לו בלוגיקה של יישומי משרד — המעבר ממערכות מודולריות וסגורות לסוכני: עם נתונים פתוחים — הוא רק חלק מתהליך רחב הרבה יותר. מדובר לא רק בשינוי ממשקים או בארכיטקטורת תוכנה השינויים יגעו בעקרונות הבסיסיים של ארגון העבודה, קבלת החלטות וניהול עסקי. בבנייה, זה יביא להיווצרות לוגיקה מונעת נתונים, שבה הנתונים יהפכו לרכיב המרכזי של התהליכים — מעיצוב ועד ניהול משאבים ובקרת התקדמות הבנייה.

משרד דיגיטלי מהדור הבא: כיצד הבינה המלאכותית משנה את סביבת העבודה

כמעט מאה שנה לפני כן, האנושות כבר חוותה מהפכה טכנולוגית דומה. המעבר ממכונות קיטור למנועים חשמליים נמשך יותר מארבעה עשורים, אך בסופו של דבר הפך לקטליזטור לצמיחה חסרת תקדים בפרודוקטיביות — בראש ובראשונה בזכות דה-מרכזיזציה של מקורות האנרגיה וגמישות הפתרונות החדשים. שינוי זה לא רק שינה את מהלך ההיסטוריה והעביר את רוב האוכלוסייה מכפרים לערים, אלא גם הניח את היסודות לכלכלה המודרנית. היסטוריית הטכנולוגיה היא מסלול מהעבודה הפיזית לאוטומציה ולמערכות אינטליגנטיות. כמו שהטרקטור החליף עשרות חקלאים, הטכנולוגיות עוד בתחילת המאה ה-20, רוב (10.13 איור) הדיגיטליות המודרניות דוחקות את השיטות המסורתיות לניהול הבנייה האוכלוסייה על פני כדור הארץ עיבדה את האדמה ביד, עד שבשנות ה-30 החלה מכניזציה של העבודה בעזרת מכונות - וטרקטורים.



כמו שהטרקטור החליף עשרות אנשים בתחילת המאה ה-20, כך גם למידת מכונה תחליף את השיטות המסורתיות 10.13 איור לניהול עסקים ופרויקטים במאה ה-21.

בדיוק כמו שהאנושות לפני מאה שנה עברה מעיבוד שטחים בודדים עם כלים פרימיטיביים לחקלאות בקנה מידה גדול עם שימוש בטכנולוגיה, היום אנו עושים את המעבר מעיבוד "סילונים" מבודדים של מידע לעבודה עם מערכי נתונים ואלגוריתמים של אינטליגנציה מלאכותית ETL-pipeline – בעזרת "טרקטורים" רבי עוצמה

אנו עומדים על סף קפיצה דומה – אך כבר במישור הדיגיטלי: מהניהול המסורתי, הידני של עסקים למודלים מבוססי נתונים.

הדרך לארכיטקטורה מבוססת נתונים מלאה תדרוש זמן, השקעות ומאמצים ארגוניים. אך הדרך הזו פותחת פתח לא רק לשיפור הדרגתי, אלא לקפיצה איכותית – לעבר יעילות רבה יותר, שקיפות וניהוליות בתהליכי הבנייה. כל זה – בתנאי של יישום שיטתי של כלים דיגיטליים ונטישה של פרקטיקות עסקיות מיושנות

טוקניזציה, נתונים גדולים ולמידת מכונה, RFID, IoT, רכיבי ETL, LLM, פרמטריזציה של משימות יהפכו את הבנייה המסורתית לבניית נתונים, שבה כל פרט בפרויקט ובעסק הבנייה יישלט ויופעל בעזרת נתונים.

המהפכה LLM-בעבר, נדרשו אלפי שעות אדם לניתוח מידע. עכשיו המשימות הללו מבוצעות על ידי אלגוריתמים ומערכי נתונים מבודדים בעזרת פרומפטים למקורות אסטרטגיים. בעולם הטכנולוגי קורה מה שקרה עם החקלאות: מהמוט – וסיכומים ידניים Excel הישן אנו עוברים למערכת אוטומטית לחקלאות. כך גם העבודה המשרדית בבנייה – מקבצי עוברת למערכת אינטליגנטית, שבה הנתונים נאספים, מנוקים, מאורגנים והופכים לתובנות.

כבר היום חברות צריכות להתחיל "לגדל" שדות מידע בעזרת איסוף איכותי של נתונים ומבנה מידע, ול"להפרות" אותם בכלים של ניקוי ונורמליזציה, ולאחר מכן "לקטוף את היבול" – בצורה של ניתוחים חזויים ופתרונות אוטומטיים. אם חקלאי מודרני עם מכונה יכול להחליף מאה חקלאים, אז גם אלגוריתמים אינטליגנטיים יכולים להקל על העובדים. מהשגרה ולהעביר אותם לתפקידים של מנהלי אסטרטגיה של זרמי מידע

עם זאת, חשוב להבין שיצירת ארגון באמת מבוסס נתונים – זה לא תהליך מהיר. זהו כיוון אסטרטגי ארוך טווח, דומה ליצירת שטח חדש לגידול יער חדש, שבו כל "עץ" במערכת האקולוגית הזו – זהו תהליך, מיומנות או כלי נפרד, שדורש זמן לצמיחה ולפיתוח. וכמו במקרה של יער אמיתי, ההצלחה תלויה לא רק באיכות החומר הנשתל (טכנולוגיות, אלא גם - באדמה) (תרבות ארגונית, (באקלים) סביבה עסקית (ובטיפול) גישה מערכתית)

חברות כבר לא יוכלו להסתמך אך ורק על פתרונות סגורים "מוכנים". בניגוד לשלבים הקודמים של התפתחות טכנולוגית המעבר הנוכחי – לגישה פתוחה לנתונים, לשימוש באינטליגנציה מלאכותית ולהפצת קוד פתוח – סביר להניח שלא יקבל תמיכה מצד ספקים גדולים, מכיוון שהוא מאיים ישירות על המודלים העסקיים המושרשים שלהם ועל מקורות ההכנסה העיקריים.

כפי שהראה מחקר של בית הספר לעסקים של הרווארד, אשר כבר נדון בפרק על המהפכות הטכנולוגיות הרביעית והחמישית, העלויות של יצירת פתרונות קוד פתוח הנמצאים בשימוש הנרחב ביותר מאפס עבור כל החברות היו מגיעות לכ-4.15 מיליארד דולר. עם זאת, אם נניח שכל חברה תפתח חלופות משלה ללא גישה לכלים קיימים של קוד פתוח כפי שקרה בעשורים האחרונים, העלויות הכוללות של העסקים עשויות להגיע לסכום עצום של 8.8 טריליון דולר - זו המחיר של ביקוש לא רציונלי, אשר יכול להעריך את שוק התוכנה.

ההתקדמות הטכנולוגית תוביל בהכרח לבחינה מחדש של מודלים עסקיים מקובלים. אם בעבר חברות יכלו להרוויח על תהליכים מורכבים, לא שקופים ונתונים סגורים, עם התפתחות הבינה המלאכותית והאנליטיקה, גישה זו הופכת לפחות ופחות ברת קיימא.

כתוצאה מכך, הדמוקרטיזציה של הגישה לנתונים ולכלים עשויה לצמצם באופן משמעותי את השוק המסורתי של מכירת תוכנה. עם זאת, במקביל לכך, יקום שוק חדש - שוק המומחיות הדיגיטלית, ההתאמה האישית, האינטגרציה ועיצוב הפתרונות. כאן הערך ייווצר לא על ידי מכירת רישיונות, אלא על ידי היכולת לבנות תהליכים דיגיטליים גמישים, פתוחים ומותאמים. בדומה לאופן שבו החשמול והופעת הטרקטורים יצרו תעשיות חדשות, כך גם השימוש בנתונים גדולים, בינה מלאכותית ומודלים לשוניים גדולים פותח אופקים חדשים לחלוטין עבור העסקים בתעשיית הבנייה, אשר ידרשו לא רק השקעות טכנולוגיות, אלא גם טרנספורמציה עמוקה של החשיבה, התהליכים והמבנים הארגוניים. והחברות והמקצוענים שיבינו זאת ויתחילו לפעול כבר היום, יהיו המובילים של מחר.

בעולם שבו נתונים פתוחים הופכים לנכס המרכזי, זמינות המידע תשנה את כללי המשחק. משקיעים, לקוחות ומסדירים ידרשו יותר ויותר שקיפות, ואלגוריתמים של למידת מכונה יהיו מסוגלים לזהות אוטומטית אי-התאמות בהערכות, בזמנים ובעלויות. זה יוצר תנאים לשלב חדש של טרנספורמציה דיגיטלית, שמוביל אותנו בהדרגה ל"אובריזציה" של תעשיית הבנייה.

נתונים פתוחים ואובריזציה הם איום על עסקי הבנייה הקיימים

הבנייה הופכת לתהליך של ניהול מידע. ככל שהנתונים מדויקים, איכותיים ומלאים יותר, כך התכנון, החישובים, הערכות העלויות, הבנייה וההפעלה של הבניינים יהיו יעילים יותר. בעתיד, המשאב המרכזי לא יהיה נוכחות של מנוף בטון וברזל, אלא היכולת לאסוף, לנתח ולהשתמש במידע.

לקוחות חברות הבנייה - משקיעים ולקוחות המממנים את הבנייה, בעתיד ישתמשו בהכרח בערך של נתונים פתוחים ואנליטיקה של נתונים היסטוריים. זה יפתח אפשרויות לאוטומציה של חישובי זמני ועלויות הפרויקטים, מבלי לערב את חברות הבנייה בשאלות של הערכות, מה שיאפשר לשלוט בהוצאות ולזהות במהירות הוצאות מיותרות.

דמיינו אתר בנייה שבו סורקי לייזר, רחפנים ומערכות פוטוגרמטריה אוספים בזמן אמת נתונים מדויקים על כמויות הבטון CAD (BIM) בשימוש. מידע זה מומר אוטומטית למודלים שטוחים פשוטים עם מטא-נתונים, תוך דילוג על מערכות הנתונים שנאספים מאתר הבנייה מועברים באופן PMIS או ERP, מסורבלות, ללא תלות בליבות גיאומטריות מורכבות מרכזי לאחסון מאורגן אחד, הנגיש ללקוח לניתוח עצמאי, שבו מועמסים מחירים אמיתיים מחנויות בנייה שונות ודוגמאות, שונות – משיעור מימון האשראי ועד גורמים דינמיים משתנים כמו תנאי מזג האוויר, ציטוטי מניות של חומרים לבנייה, תעריפי לוגיסטיקה וסטטיסטיקות של תנודות עונתיות במחירי העבודה. בתנאים כאלה, כל סטייה בין כמויות החומרים המתוכננות לאלו בפועל הופכת מיד לגלויה, מה שהופך מניפולציות על חישובי התקציב לבלתי אפשריות, הן בשלב התכנון והן בעת מסירת הפרויקט. בסופו של דבר, שקיפות תהליך הבנייה מושגת לא בזכות צבא של בודקים ומנהלים, אלא בזכות נתונים דיגיטליים אובייקטיביים, שבהם יוקטן למינימום הגורם האנושי ואפשרות הספקולציות

(CQMS איור 1.24 מנהל) עבודה כזו על בקרת נתונים תתבצע בעתיד ככל הנראה על ידי מנהלי נתונים מצד הלקוח במיוחד כאשר מדובר בחישובים ובחישובי תקציב של פרויקטים: היכן שבעבר עבד מחלקה שלמה של מחשבים, כבר מחר יופיעו כלים של למידת מכונה וחזוי, שיקבעו לחברות הבנייה את המסגרות המחיריות שבהן עליהן להתאים את עצמן.

בהתחשב באופיו המפוצל של ענף הבנייה, כאשר רוב המערכות והמערכות המשניות מסופקות על ידי עסקים קטנים ובינוניים, האסטרטגיה הדיגיטלית צריכה לצאת מהלקוח. הלקוחות צריכים ליצור תנאים ומנגנונים לחשיפת האפשרויות הדיגיטליות של שרשרת האספקה.

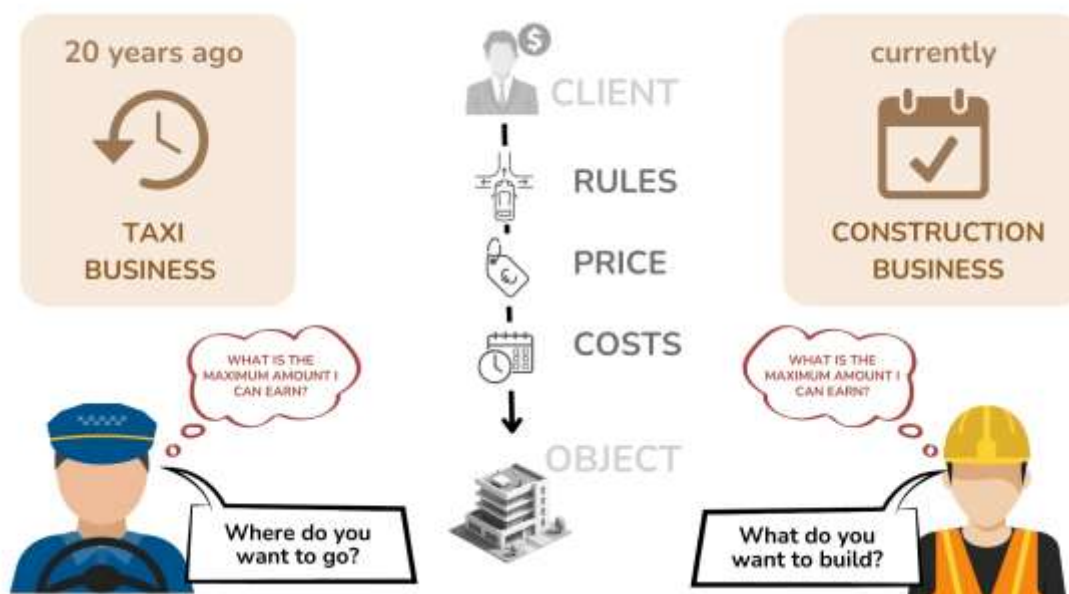
"יצירת ערך גדול יותר באמצעות פרויקטים קפיטליים" Accenture, אנדרו דיוויס וג'וליאנו דניקול

פתיחות ושקיפות כזו של נתונים מהווה איום על חברות הבנייה, שהתרגלו להרוויח מהעדר שקיפות בתהליכים ודוחות מסובכים, שבהם ניתן להסתיר ספקולציות והוצאות נסתרות כפורמטים מורכבים וסגורים ובפלטרמות נתונים מודולריות קנייניות. לכן, חברות הבנייה, כמו במקרה של קידום פתרונות קוד פתוח על ידי ספקים, סביר להניח שלא יתעניינו ביישום מלא של נתונים פתוחים בתהליכי העבודה שלהן. אם הנתונים יהיו זמינים וניתנים לעיבוד בקלות עבור הלקוח, ניתן יהיה לבדוק אותם אוטומטית, מה שיבטל את האפשרות להפריז בכמויות ולבצע מניפולציות על התקציבים

לפי דוח הפורום הכלכלי העולמי "עיצוב עתיד הבנייה" (2016), אחת הבעיות המרכזיות בענף היא התפקיד הפסיבי של הלקוח. עם זאת, הלקוחות עצמם צריכים לקחת על עצמם אחריות רבה יותר על תוצאות הפרויקטים – החל מתכנון מוקדם, בחירת מודלים בני קיימא של אינטראקציה ועד פיקוח על הביצוע. ללא מעורבות פעילה מצד בעלי הפרויקטים, טרנספורמציה מערכתית של ענף הבנייה אינה אפשרית.

אובדן השליטה על חישובי כמויות ועלויות כבר הפך לשינוי במהלך 20 השנים האחרונות בענפים אחרים בכלכלה, מה שאיפשר ללקוחות להשיג את מטרותיהם ישירות, ללא מתווכים. הדיגיטציה ושקיפות הנתונים שינו רבות מהמודלים

העסקיים המסורתיים, כפי שקרה עם נהגי המוניות לאחר הופעת אובר (איור 10.14), (עם בעלי המלונות לאחר הגעת ועם קמעונאים וחנויות לאחר התפתחות אמזון, כמו גם עם הבנקים - בזכות עליית הניאו-בנקים ואקוסיסטמות Airbnb הפינטק המופצות, שבהן הגישה הישירה למידע ואוטומציה של חישובי זמן ועלות הפחיתו באופן משמעותי את תפקיד המתווכים.



איור 10.14 עסקי הבנייה יתמודד עם "אוברייזציה", כפי שנאלצו להתמודד נהגי המוניות, בעלי המלונות והקמעונאים לפני 10 שנים.

תהליך הדמוקרטיזציה של הגישה לנתונים ולכלים לעיבודם הוא בלתי נמנע, ובסופו של דבר נתונים פתוחים בכל המרכיבים של הפרויקט יהפכו לדרישה של הלקוח ולסטנדרט חדש. לכן, שאלות של יישום פורמטים פתוחים וחישובים שקופים יקדמו בעיקר מצד משקיעים, מזמינים, בנקים וקרנות השקעה פרטיות - אלו שבסופו של דבר הם המשתמשים הסופיים של המבנים שנבנים ולאחר מכן מפעילים את המבנה במשך עשרות שנים.

יצירת "Accenture משקיעים גדולים, מזמינים ובנקים כבר דורשים שקיפות בענף הבנייה. לפי מחקר של ערך גדול יותר באמצעות פרויקטים הוניים" (2020), נתונים שקופים ואמינים הופכים לגורם מכריע בהחלטות השקעה בבנייה. כפי שמציינים המומחים, ניהול פרויקטים אמין ויעיל אינו אפשרי ללא שקיפות במיוחד בתנאי משבר. בנוסף, בעלי נכסים וקבלנים עוברים יותר ויותר לחוזים שמעודדים חילופי נתונים וניתוח משותף, מה שמשקף את הדרישות הגוברות מצד משקיעים, בנקים ומסדירים לאחריות ולשקיפות.

תנועת המשקיע, המזמין מהרעיון ועד הבניין המוגמר, בעתיד תהיה דומה למסע על טייס אוטומטי - ללא נהג בדמות חברת הבנייה, מבטיחה להיות בלתי תלויה בספקולציות ובחוסר ודאות.

עידן הנתונים הפתוחים והאוטומציה ישנה באופן בלתי נמנע את עסקי הבנייה כפי שכבר קרה בבנקאות, מסחר חקלאות ולוגיסטיקה. בענפים אלה תפקיד המתווכים והדרכים המסורתיות לעסוק בעסקים מפנים את מקומם לאוטומציה ולרובוטיזציה, מבלי להשאיר מקום להעלאות מחירים בלתי מוצדקות ולספקולציות.

נתונים ותהליכים בכל סוגי הפעילות הכלכלית של בני אדם אינם שונים מאלה שעליהם מתמודדים אנשי מקצוע בענף הבנייה. בטווח הארוך, חברות הבנייה ששולטות כיום בשוק, קובעות את הסטנדרטים של מחירים ואיכות שירותים. עשויות לאבד את תפקידן כמתווכים המרכזיים בין המזמין לבין פרויקט הבנייה שלו.

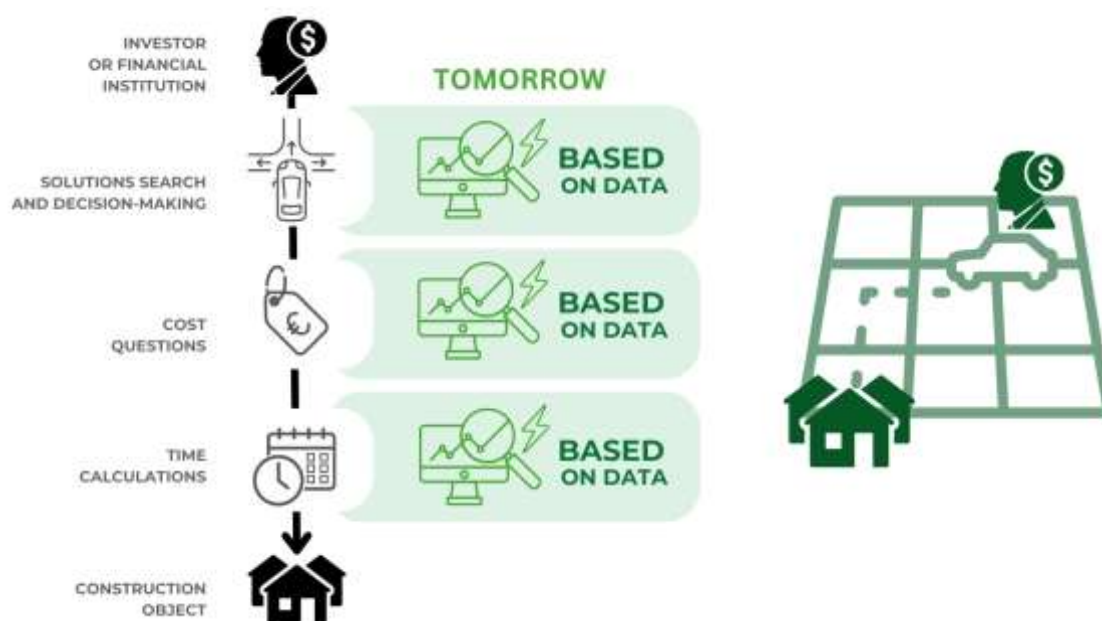
בעיות לא פתורות של אוברייזציה כהזדמנות האחרונה לנצל את הזמן עבור טרנספורמציה

אך נחזור למציאות של ענף הבנייה. בעוד שבחלקים מסוימים של הכלכלה מופיעים רכבים אוטונומיים, מערכות פיננסיות, מבוזרות ופתרונות מבוססי אינטליגנציה מלאכותית, חלק ניכר מחברות הבנייה עדיין נשארות ארגונים מבוססי נייר שבהם ההחלטות המרכזיות מתקבלות בעיקר על סמך אינטואיציה וניסיון של אנשי מקצוע בודדים.

בהקשר זה, ניתן להשוות חברת בנייה מודרנית עם תחנת מוניות מלפני 20 שנה, אשר שולטת במשאבים, במסלולים ובזמן ההגעה, נושאת באחריות על לוחות הזמנים ועל עלות "הנסיעה" – מהרעיון התכנוני (תהליך הלוגיסטיקה וההתקנה) ואלגוריתמים של למידת מכונה בחישובי זמן/עלות (IoT, RFID בעבודות בנייה) GPS-ועד למסירת הפרויקט. כפי ש יכולים לשנות את ניהול הבנייה – מהערכות AI שינו את תחום התחבורה, כך גם נתונים, אלגוריתמים וסוכני אינטואיטיביות למודל מנוהל וחזוי. בעשרים השנים האחרונות, בענפים רבים – פיננסים, חקלאות, קמעונאות, ולוגיסטיקה – האפשרות לספק ספקולציות על סמך חוסר שקיפות של נתונים הלכה והפכה לפחות אפשרית. מחירים עלויות משלוח או עסקאות פיננסיות מחושבים אוטומטית ובצורה סטטיסטית מבוססת – תוך מספר שניות בלבד. בפלטפורמות דיגיטליות.

בהביט לעתיד, חברות הבנייה צריכות להבין כי דמוקרטיזציה של הגישה לנתונים ולכלים לניתוחם תערער את הגישה המסורתית להערכת עלויות ולוחות זמנים של פרויקטים ותשלול את האפשרות לספק ספקולציות על סמך נתונים לא שקופים לגבי כמויות ומחירים.

בדומה לנסיעה בכביש מוסדר ללא התערבות של נהג, תהליכי הבנייה של העתיד יזכירו יותר ויותר מערכת "אובר" – עם הערכה אוטומטית של לוחות זמנים ועלויות, ניהול שקוף של משימות ותלות מינימלית בגורם האנושי. זה ישנה את מהות ה"נסיעה" מהרעיון ועד ההוצאה לפועל – ויהפוך אותה ליותר צפויה, מנוהלת וממוקדת נתונים.



עלות וזמן ה"דרך" בתהליך הבנייה ייקבעו בעזרת למידת מכונה וכלים סטטיסטיים 10.15. ריש

CAD- (BIM)- עם ההכנסה ההדרגתית של תקנות ודרישות חדשות כמעט בכל מדינה בעולם, המאלצות להעביר מודלים ללקוחות או לבנקים המממנים פרויקטי בנייה, הלקוח והזמין מקבלים את האפשרות להבטיח שקיפות בחישובי נתוני העלות וכמויות העבודה בעצמם. זה במיוחד רלוונטי עבור לקוחות גדולים ומשקיעים שיש להם את הכישורים והכלים – לניתוח מהיר של כמויות ומעקב אחר מחירים בשוק. עבור חברות המיישמות פרויקטים סטנדרטיים בקנה מידה גדול – חנויות, בנייני משרדים, קומפלקסים מגורים – פרקטיקות כאלה הופכות לסטנדרט

ככל שהמידע במודלים הופך להיות מלא יותר ומסודר יותר, האפשרות למניפולציות ולספקולציות כמעט נעלמת. טרנספורמציה דיגיטלית משנה בהדרגה את כללי המשחק בענף הבנייה, וחברות שלא יתאימו לשינויים הללו עלולות להתמודד עם אתגרים משמעותיים

חיזוק התחרות, הפער הטכנולוגי והירידה ברווחיות יכולים להשפיע על יציבות העסק. בתנאים של נזילות מוגבלת, יותר ויותר משתתפים בתעשייה פונים לאוטומציה, אנליטיקה וטכנולוגיות עיבוד נתונים כדרך לשיפור היעילות והשקיפות של התהליכים. כלים אלה הופכים למשאב חשוב לשמירה על התחרותיות בסביבה הכלכלית המשתנה

ייתכן שאין צורך לחכות עד שהנסיבות החיצוניות ידרשו לנקוט בצעדים דחופים - הרבה יותר יעיל להתחיל בהכנה כבר היום, לחזק את המיומנויות הדיגיטליות, ליישם פתרונות מודרניים ולבנות תרבות הממוקדת בעבודה עם נתונים אחד מהמכשולים הטכנולוגיים המרכזיים האחרונים בדרך לדיגיטציה רחבת היקף של תעשיית הבנייה, שתשפיע בשנים הקרובות על כל חברה, הוא בעיית הסיווג האוטומטי של רכיבי פרויקטים בנייתיים

ללא סיווג מהימן, מדויק וניתן להרחבה, אין אפשרות ליצור בסיס לאנליטיקה מלאה, אוטומציה של תהליכים וניהול מחזור חיי האובייקטים באמצעות בינה מלאכותית ודגמים חיזוי. כל עוד הסיווג של האובייקטים עדיין תלוי בפרשנות ידנית מצד מומחים מנוסים - מנהלי עבודה, מעצבים, ומערכי עלויות לתעשיית הבנייה נשאר חלון הזדמנויות. זמן זה ניתן לניצול כהכנה לשינויים הבלתי נמנעים: עלייה בדרישות לשקיפות, דמוקרטיזציה של כלים ונתונים, וכן הופעת מערכות סיווג אוטומטי שישנו באופן רדיקלי את כללי המשחק.

משימת הסיווג האוטומטי של רכיבי עולם הבנייה מבחינת מורכבותה דומה לזיהוי אובייקטים במערכות נהיגה אוטונומית, מה שמייצג את אחד האתגרים המרכזיים. נניח רכב אוטונומי הנע מנקודה א' לנקודה ב' (איור 10.15). (מערכות הנהיגה האוטומטיות המודרניות נתקלות בבעיה של סיווג אובייקטים, המוכרים באמצעות לייזרים ומצלמות. לרכב לא מספיק לראות "מכשול או סימן - עליו להבין ללא שגיאות מה עומד מולו: הולך רגל, שלט או פח אשפה".

בעיה יסודית דומה עומדת בפני כל תעשיית הבנייה. רכיבי הפרויקט - כמו חלונות, דלתות או עמודים - יכולים להיות מצולמים באתר הבנייה או מזוהים בענני נקודות מסריקות לייזר. עם זאת, CAD מתועדים במסמכים, מיוצגים במודלים כדי לבנות מערכת ניהול פרויקטים אוטומטית באמת, לא מספיק רק זיהוי ויזואלי או גיאומטרי גס. יש צורך להבטיח סיווג מדויק ועמיד של כל רכיב לפי סוג, שיהיה ניתן לזיהוי חד משמעי בכל התהליכים הבאים - מהערכות עלויות ומפרטים ועד לוגיסטיקה, ניהול מלאי והכי חשוב - תפעול (איור 4.26).

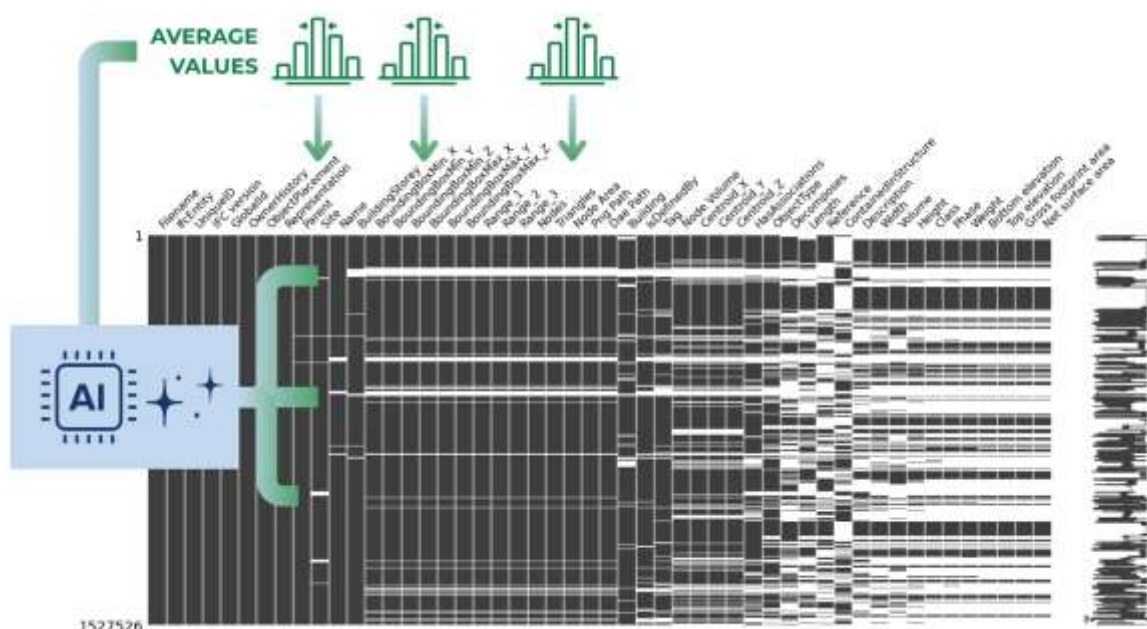
בשלב זה - המעבר מהכרה לסיווג משמעותי - מתעורר אחד המכשולים המרכזיים. גם אם מערכות דיגיטליות מסוגלות טכנית להבחין ולזהות אובייקטים במודלים ובאתרי בנייה, הקושי העיקרי טמון בהגדרה נכונה ועקבית של סוג האלמנט, "כאלמנט בקטגוריית "דלת CAD עבור סביבות תוכנה שונות. לדוגמה, דלת עשויה להיות מסומנת על ידי המהנדס במודל היא עשויה לקבל סיווג שגוי - במקרה של טעות מהמהנדס או בשל PMIS או ERP אך כאשר היא מועברת למערכות חוסר התאמה בין המערכות. יתרה מכך, האלמנט לעיתים קרובות מאבד חלק מהמאפיינים החשובים או נעלם לחלוטין מהמעקב המערכתי במהלך ייצוא וייבוא נתונים. זה מוביל לקרע בזרימת הנתונים ומערער את עקרון הדיגיטציה המלאה של תהליכי הבנייה. כך נוצר פער קריטי בין המשמעות "הנראית" ל"ברורה", מה שמערער את שלמות הנתונים ומקשה באופן משמעותי על אוטומציה של תהליכים לאורך כל מחזור חיי הפרויקט הבנייתי.

פתרון בעיית הסיווג האוניברסלי של אלמנטים בנייה באמצעות טכנולוגיות נתונים גדולים ולמידת מכונה (איור 10.16) יהפוך לקטליזטור של טרנספורמציה בכל התחום - ואולי גם גילוי מפתיע עבור רבות מחברות הבנייה. מערכת סיווג מאוחדת ולומדת תהפוך לבסיס לניתוח בקנה מידה רחב, ניהול דיגיטלי והטמעת בינה מלאכותית בפרקטיקה היומיומית - של ארגוני הבנייה.

ומובילים טכנולוגיים אחרים כבר היום מציעים פתרונות בתעשיות אחרות, המסוגלים לסווג ולארגן באופן NVIDIA אוטומטי כמויות עצומות של מידע טקסטואלי וויזואלי.

לדוגמה, מתמחה בסיווג אוטומטי והפצת נתונים לפי NVIDIA של [161] NeMo Curator מודל קטגוריות שנקבעו מראש, ומשחק תפקיד מרכזי באופטימיזציה של צינורות עיבוד מידע למשימות של כוונן מתאמנת על Cosmos עדין והכשרה מוקדמת של מודלים גנרטיביים של בינה מלאכותית. הפלטפורמה, סרטונים אמיתיים וסצנות תלת-ממדיות [162], [יצירת בסיס עבור מערכות אוטונומיות ודיגיטליים תאומים שהפכה עד 2025 לכלי המוביל לעבודה, NVIDIA Omniverse. NVIDIA. שכבר נבנים באקוסיסטם של בתהליכי העברת IFC תיאור אוניברסלי של סצנות, שיכול בעתיד להחליף את פורמט - USD עם פורמט NeMo סימולטור תהליכים רובוטיים [163] -] פתרונות כמו - Isaac Sim מידע פרויקט. בשילוב עם מייצגים רמה חדשה של אוטומציה: מניקוי וסינון נתונים ועד יצירת Omniverse ו-Cosmos Curator, מערכי נתונים להכשרה, סימולציה של תכונות אובייקטים והכשרת רובוטים באתר הבנייה. יתרה מכך, כל הכלים הללו מופצים בחינם ובגישה פתוחה, מה שמפחית באופן משמעותי את המחסומים להטמעה בפרקטיקות הנדסיות ובנייה.

סיווג אוטומטי של נתונים ברמת טבלאות מובנות הוא משימה שאינה כה מורכבת כפי שהיא עשויה להיראות במבט כאשר ישנם נתונים היסטוריים מצטברים, ניתן להשלים ערכים, (9.110 איור) ראשון. כפי שהראינו בפרק הקודם חסרים או לא נכונים של קטגוריות על בסיס פרמטרים דומים של אלמנטים אחרים. אם בפרויקטים שהושלמו מספר אלמנטים עם מאפיינים דומים כבר סווגו כראוי, המערכת יכולה להציע ערך מתאים לאלמנט חדש או חסר עם סבירות לוגיקה כזו, המבוססת על ערכים ממוצעים וניתוח הקשר, עשויה להיות יעילה במיוחד כאשר. (10.16 איור) גבוהה CAD. מעבדים כמות גדולה של נתונים טבלאיים המגיעים מהערכות. מפרטים או דגמים

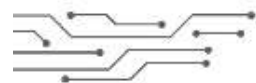


למידת מכונה תסייע למצוא אוטומטית ערכים ממוצעים עבור פרמטרים חסרים (בטבלאות על בסיס 10.16 איור - פרויקטים קודמים

על רקע ההתקדמות המהירה בתחום למידת המכונה, ברור כי בשנת 2025 יהיה נאיבי להניח שהבעיה של סיווג אוטומטי של אלמנטים בנייה תישאר לא פתורה זמן רב. נכון, האלגוריתמים המודרניים עדיין לא הגיעו לבשלות מלאה, במיוחד בתנאים של נתונים חסרים או לא אחידים, אך חלון ההזדמנויות להתאמה נסגר במהירות.

ימצאו ETL חברות שכבר היום משקיעות באיסוף, ניקוי וסידור הנתונים שלהן, כמו גם בלימוד כלים לאוטומציה של את עצמן במצב עדיף. השאר מסתכנים בכך שלא יצליחו להדביק את הפער – כמו שחברות מסוימות לא הצליחו להתמודד עם האתגרים של טרנספורמציה דיגיטלית בתחומי התחבורה והפיננסים.

אלו שימשיכו להסתמך על ניהול ידני של נתונים ושיטות מסורתיות להערכת עלויות ולוחות זמנים, מסתכנים בכך שיתפסו את עצמם במצב דומה למוניות של שנות ה-2000, שלא הצליחו להסתגל עד תחילת שנות ה-2020 לעידן האפליקציות הניידות ולחשובי מסלולים אוטומטיים.



פרק 10.2.

מדריך מעשי ליישום גישה מונחית נתונים

מעקרון למעשה: מפת דרכים להמרה דיגיטלית בענף הבנייה

תעשיית הבנייה נכנסת בהדרגה לשלב חדש של התפתחות, שבו תהליכים מוכרים מתווספים – ולעיתים מוחלפים – על ידי פלטפורמות דיגיטליות ודגמים שקופים של אינטראקציה. זה פותח בפני חברות לא רק אתגרים, אלא גם הזדמנויות משמעותיות. הארגונים שכבר היום בונים אסטרטגיה דיגיטלית ארוכת טווח, יוכלו לא רק לשמור על מעמדם בשוק, אלא גם להרחיב אותו, על ידי הצעת גישות מודרניות ופתרונות אמינים הנתמכים בטכנולוגיה.

עם זאת, חשוב להבין: ידע במושגים ובטכנולוגיות הוא רק נקודת התחלה. בפני מנהיגים ומומחים עומד השאלה המעשית: מאיפה להתחיל את היישום וכיצד להפוך רעיונות תיאורטיים לערך ממשי. בנוסף, עולה יותר ויותר השאלה: על מה תתבסס העסקה, אם השיטות המסורתיות לחישוב עלויות ולוחות זמנים עשויות להתבצע מחדש על ידי הלקוח בכל רגע.

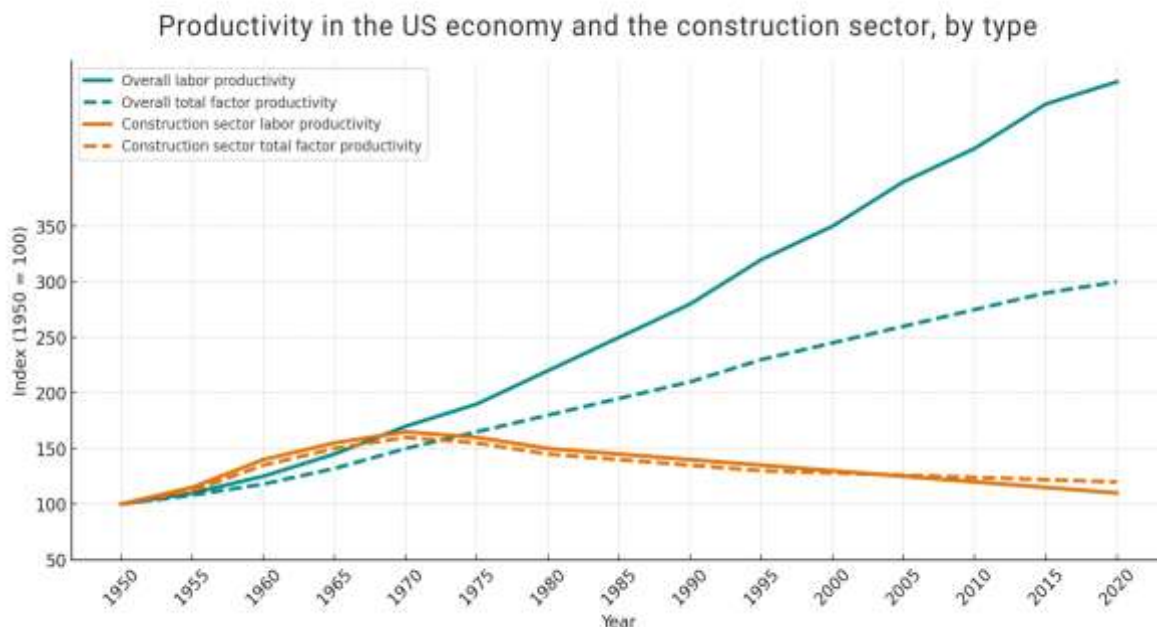
התשובה, ככל הנראה, טמונה לא כל כך בטכנולוגיות, אלא בעיצוב תרבות מקצועית חדשה, שבה עבודה עם נתונים נתפסת כחלק בלתי נפרד מהפרקטיקה היומיומית. חוסר תשומת הלב לטכנולוגיות דיגיטליות ולחדשנות הוא זה שהוביל לפער משמעותי בתעשיית הבנייה, שנמשך בעשורים האחרונים.

על פי נתוני מקינזי, ההוצאות על מחקר ופיתוח בתעשיית הבנייה מהוות פחות מ-1% מההכנסות, בעוד שבתעשיית הרכב והחלל שיעור זה מגיע ל-3.5%–4.5%. באופן דומה, ההוצאות על טכנולוגיות מידע בבנייה נשארות ברמה של פחות מ-1% מההכנסות הכוללות.

כתוצאה מכך, לא רק שהרמה של אוטומציה, אלא גם הפרודוקטיביות בעבודה בבנייה יורדת, ובשנת 2020 העובד הבנאי מייצר כבר פחות מאשר לפני חמישה עשורים.

בעיות דומות בפרודוקטיביות במגזר הבנייה מאפיינות את רוב המדינות המפותחות והמתפתחות (הפרודוקטיביות בבנייה ומצביעות לא רק על חוסר בטכנולוגיות, אלא גם על הצורך בשינויים, OECD-ירדה ב-16 מתוך 29 מדינות המערכתיים בגישות לניהול, הכשרה והטמעת חדשנות).

הצלחה של טרנספורמציה דיגיטלית תלויה לא רק בכמות ובזמינות הכלים, אלא ביכולת של הארגונים לבחון מחדש את התהליכים שלהם ולפתח תרבות פתוחה לשינויים. התפקיד המרכזי לא שייך לטכנולוגיות עצמן, אלא לאנשים ולתהליכים המובנים שמבטיחים את השימוש היעיל בהן, תומכים בלמידה מתמשכת ומקדמים את הקבלה של רעיונות חדשים.



איור 10.21 פרדוקס הפרודוקטיביות בעבודה ובפרודוקטיביות המשאבים הכוללת בכלכלה האמריקאית ובמגזר הבנייה (1950-2020).

בחלקים הראשונים של הספר, מודל הסביבה העסקית הושווה למערכת אקולוגית של יער. ביער בריא, שריפות תקופתיות, על כל כוחן ההרסני, משחקות תפקיד מרכזי בחידוש ארוך טווח. הן מנקות את האדמה מצמחייה ישנה, מחזירות את החומרים המזינים שנצברו ויוצרות מקום לחיים חדשים. כמה סוגי צמחים אף התפתחו כך שהזרעים שלהם נפתחים רק בהשפעת חום גבוה של שריפה – זהו מנגנון טבעי המבטיח את הזמן המושלם לנביטה.

באופן דומה בעסקים: משברים יכולים לשחק תפקיד של "שריפה מבוקרת", המעודדת את הופעתם של גישות וחברות חדשות, שאינן קשורות למערכות מיושנות. תקופות כאלה מאלצות לוותר על פרקטיקות לא יעילות, ומשחררות משאבים לחדשנות. כמו שהיער לאחר השריפה מתחיל עם צמחים חלוצים, כך גם העסק לאחר המשבר יוצר תהליכים חדשים וגמישים, המהווים את הבסיס לסביבה מידע בשלה.

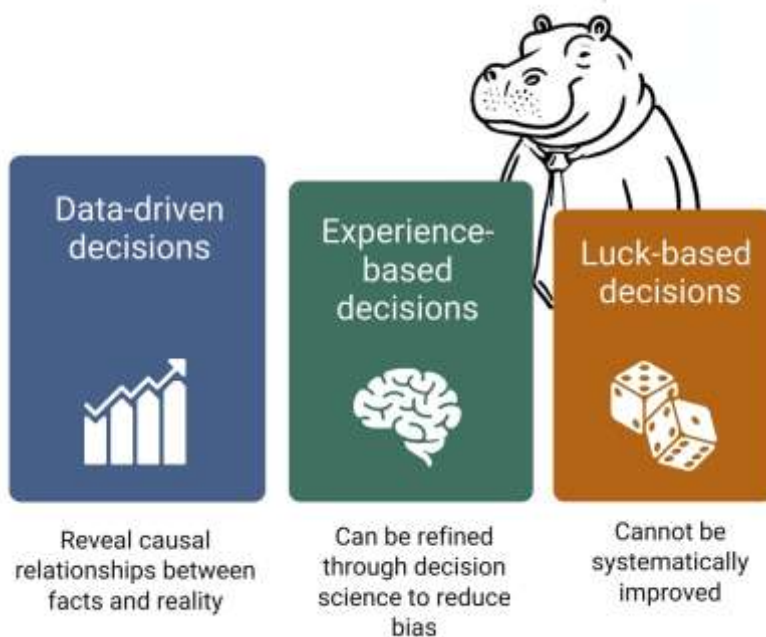
חברות שהצליחו לפרש נכון את "שריפות האות" הללו ולהמיר את האנרגיה ההרסנית שלהן לשינויים בונים, יגיעו לרמה חדשה של יעילות – עם תהליכים שקופים וגמישים יותר לעיבוד נתונים, המגבירים את היכולת הטבעית של הארגון להתחדש ולצמוח.

ההשפעה הגוברת של אינטליגנציה מלאכותית ולמידת מכונה על הסביבה העסקית כבר אינה מוטלת מסתכנות באובדן AI-בספק. זו לא רק מגמה זמנית, אלא צורך אסטרטגי. חברות שמתעלמות מהתחרויות בשוק, שמעודד יותר ויותר חדשנות וגמישות.

לא רק כלי, אלא הזדמנות לחשוב מחדש על כל היבט בפעילותם - מהאופטימיזציה של AI-העתיד שייך לאלה שרואים ב- תהליכים ועד קבלת החלטות ניהוליות.

מניחים את היסוד הדיגיטלי: 1-5 צעדים לבגרות דיגיטלית

בפרק זה נבחן את מפת הדרך של טרנספורמציה דיגיטלית ונגדיר את הצעדים המרכזיים הנדרשים ליישום גישה מונחית נתונים, שיכולה לסייע בשינוי הן את התרבות הארגונית והן את המערכת האקולוגית המידע של החברה.



עדכון מבוקר ובחירת אסטרטגיה: מקרה, ניסיון או נתונים 10.22 איור.

לפי מחקר של מקינזי "למה אסטרטגיות דיגיטליות נכשלות" (2018), יש לפחות חמישה סיבות [164] לכך שחברות אינן משיגות את מטרות הטרנספורמציה הדיגיטלית.

■ הגדרות מעורפלות: מנהיגים ומנהלים מבינים באופן שונה מהן "טכנולוגיות דיגיטליות", "מה שמוביל לאי הבנות וחוסר תיאום בפעולות

■ הבנה שגויה של כלכלת הטכנולוגיות הדיגיטליות: רבות מהחברות מעריכות נמוך את היקף השינויים (10.16 איור) שהדיגיטליזציה מביאה למודלים עסקיים ודינמיקות של תעשיות.

■ התעלמות מהמערכות האקולוגיות: חברות מתמקדות בפתרונות טכנולוגיים בודדים (סילואטים של נתונים) (4.112 איור, 2.22 איור) ומפספסות את הצורך בשילוב במערכות אקולוגיות דיגיטליות רחבות יותר.

■ הערכת יתר של הדיגיטליזציה מצד המתחרים: מנהיגים אינם לוקחים בחשבון כי המתחרים גם מיישמים טכנולוגיות דיגיטליות, מה שעלול להוביל לאובדן יתרון תחרותי.

■ החמצת הדואליות של הדיגיטליזציה: מנכ"לים מעבירים את האחריות על הטרנספורמציה הדיגיטלית למנהיגים אחרים, מה שמבצע בירוקרטיזציה של הבקרה ומאט את תהליך השינויים.

כדי לפתור בעיות אלו, יש צורך בהבנה ברורה ותיאום של האסטרטגיות הדיגיטליות בכל הרמות של הארגון. לפני שמבנים אסטרטגיה דיגיטלית, חשוב להבין את נקודת ההתחלה. רבות מהארגונים שואפים ליישם כלים ופלטפורמות

חדשות, מבלי שיש להם תמונה מלאה של המצב הנוכחי

שלב 1. ערכו אודיט של המערכות והנתונים הנוכחיים

לפני שמבצעים שינויים בתהליכים, חשוב להבין מה כבר קיים. עריכת אודיט מאפשרת לחשוף את החולשות בניהול הנתונים ולהבין אילו משאבים ניתן לנצל. אודיט כזה הוא מעין "רנטגן" של תהליכי העסקים. הוא יאפשר לכם לחשוף אזורי סיכון ולקבוע אילו נתונים הם קריטיים לפרויקט או לעסק שלכם, ואילו הם משניים

פעולות עיקריות:

■ רשמו את המערכות (Canva או Draw.io, Lucidchart, Miro, Visio) IT-צרו מפה של הסביבה המשתתפות בתהליכים שלכם ואשר דנו בהם בפרק (ועוד ERP, CAD, CAFM, CPM, SCM) בשימוש (1.24 איור) "טכנולוגיות ומערכות ניהול בבנייה המודרנית"

■ העריכו בעיות באיכות הנתונים עבור כל מערכת לפי תדירות קיומם של כפילויות, ערכים חסרים אפשריים ואי-התאמות בפורמטים בכל אחת מהמערכות

■ חשפו "נקודות כאב" - מקומות שבהם התהליכים עשויים להיקרע או לעיתים קרובות דורשים התערבות ידנית. ייבוא, ייצוא ותהליכי בדיקה נוספים -

אם אתם רוצים שהצוות יאמין בדו"ח, יש לעקוב אחרי נכונות הנתונים מההתחלה

ביקורת נתונים איכותית תראה אילו נתונים

■ (זקוקים לשיפור) דרושה הגדרה מחדש של תהליכי ניקוי אוטומטיים או טרנספורמציה נוספת)

■ מהווים "זבל", שמסבך את המערכות וניתן להיפטר ממנו, מבלי להשתמש בו יותר בתהליכים

ניתן לבצע ביקורת כזו באופן עצמאי. אך לפעמים מועיל להביא יועץ חיצוני - במיוחד מתחומים אחרים בכלכלה: מבט חדש ועצמאות מה"מאפיינים" של הבנייה יעזרו להעריך בצורה רציונלית את המצב הנוכחי ולהימנע ממלכודות טיפוסיות של הטיית דעת לגבי החלטות וטכנולוגיות מסוימות

שלב 2. קבעו סטנדרטים מרכזיים לאחידות הנתונים

לאחר ביצוע הביקורת יש ליצור כללים משותפים לעבודה עם נתונים. כפי שדנו בפרק "סטנדרטים: מקבצים אקראיים למודל נתונים מתוכנן", זה יעזור להעלים את הפיזור של זרמי המידע

ללא סטנדרט אחיד, כל צוות ימשיך לעבוד "בדרכו", ואתם תשמרו על "גן חיות" של אינטגרציות, שבהן נתונים הולכים לאיבוד בכל טרנספורמציה

פעולות עיקריות:

■ בחרו סטנדרטים לנתונים להחלפת מידע בין מערכות

□ או פורמטים יעילים CSV, XLSX עבור נתונים טבלאיים, אלו יכולים להיות פורמטים מובנים כמו

Parquet יותר כמו

☐ XML או JSON: להחלפת נתונים חלשים ומסמכים

■ למדו לעבוד עם מודלים של נתונים

☐ התחילו עם פרמטריזציה של משימות ברמת המודל הקונספטואלי של הנתונים - כפי שמתואר בפרק (4.32 איור) "מודל נתונים: מודל קונספטואלי, לוגי ופיזי"

☐ ככל שתעמיקו בלוגיקה של תהליכי העסקים, עברו לה של דרישות באמצעות פרמטרים במודלים - (4.36 איור) הלוגיים והפיזיים

☐ קבעו ישויות מרכזיות, את המאפיינים שלהן והקשרים במסגרת התהליכים, וכן ויזואליזציה של (4.37 איור) הקשרים הללו - הן בין ישויות והן בין פרמטרים

■ כפי שדנו בפרק, (4.47 איור) לאימות וסטנדרטיזציה של נתונים (Regex) השתמשו בביטויים רגולריים נושא לא מסובך, אך קריטי בעבודת יצירת דרישות - Regex. "Regex דרישות מובנות וביטויים רגולריים" - ברמת המודלים הפיזיים של הנתונים

ללא סטנדרטים ברמת הנתונים וויזואליזציה של תהליכים, בלתי אפשרי להבטיח סביבה דיגיטלית עקבית וניתנת להרחבה. זכרו: נתונים גרועים עולים ביוקר. "והמחיר של טעות עולה ככל שהפרויקט או הארגון הופכים מורכבים יותר. אחדות בפורמטים, קביעת כללי שמות, מבנה ואימות - אלו השקעות ביציבות וביכולת ההרחבה של הפתרונות העתידיים.

ואוטומטו תהליכים DataOps שלב 3. יישמו

ללא ארכיטקטורה ברורה של החברה, בלתי נמנע להיתקל בנתונים מפוזרים, כלואים במערכות מידע מבודדות. הנתונים לא יתממשקו, יוכפלו במקומות שונים וידרשו הוצאות משמעותיות לתמיכה

דמיינו שדאטה היא מים, ואדריכלות דאטה היא מערכת מורכבת של צינורות, דרכם המים מועברים ממקור האחסון למקום השימוש. אדריכלות הדאטה קובעת כיצד המידע נאסף, נשמר, מומר, מנותח ומועבר למשתמשים הסופיים או ליישומים

(פעולות דאטה היא מתודולוגיה המאחדת את האיסוף, הניקוי, הבדיקה והשימוש בדאטה לזרם תהליכים DataOps אוטומטי אחד, כפי שדיברנו עליו בפירוט בפרק השמיני של הספר

פעולות עיקריות:

■ כדי לאוטומט את התהליכים ETL צרו והגדירו צינורות

☐ CAD, מודלים, Excel, גיליונות, PDF, ארגנו איסוף אוטומטי של דאטה מתוך מסמכי Extract: ---ומקורות אחרים שאתם עובדים איתם ERP מערכות

☐ הגדרו תהליכי המרה אוטומטיים לדאטה בפורמט מובנה אחיד ואוטומטו חישובים Transform: -שיתבצעו מחוץ ליישומים סגורים

□ ---נסו ליצור ייצוא אוטומטי של דאטה לטבלאות סופיות, מסמכים או מחסנים מרכזיים: Load

■ QTO Quantity "כפי שדיברנו בפרק, QTO (Quantity Take-Off) אוטומטו את תהליכי החישוב ו
"קבוצת נתוני פרויקט לפי תכונות: Take-Off".

□ תוספים או כלים של הנדסה, API, באמצעות CAD, הגדרו חילוץ אוטומטי של כמויות ממודלי
-הפוכה.

□ צרו כללים לקיבוץ פריטים עבור קטגוריות שונות לפי תכונות בצורה של טבלאות.

□ נסו לאוטומט חישובים חוזרים ונשנים של כמויות ועלויות מחוץ למערכות סגורות מודולריות.

■ כלי: Python Pandas "לעיבוד דאטה, כפי שראינו בפרק Pandas ו-Python-התחילו להשתמש ב
"חיוני לעבודה עם דאטה".

■ -ואוטומציה של עיבוד נתונים טבלאיים XLSX לעבודה עם קבצי DataFrame-השתמשו ב

■ Python. אוטומטו את האגרציה וההמרה של מידע דרך ספריות שונות של

■ -כדי לפשט את כתיבת קטעי קוד מוכנים וצינורות שלמים LLM-השתמשו ב

■ ,שמאתר שגיאות או רואה אנומליות ושולח הודעה לאחראי) למשל Python-נסו לבנות צינור ב
-.(מנהל פרויקט

מאפשרת לעבור מעבודה ידנית ומפוצלת עם דאטה לתהליכים יציבים וחוזרים DataOps אוטומציה על בסיס עקרונות
זה, לא רק מפחית את העומס על עובדים העוסקים באותם המרות יום-יום, אלא גם מעלה באופן דרמטי את האמינות
הסקלריות והשקיפות של כל מערכת המידע

שלב 4. צרו מערכת אקולוגית לניהול דאטה פתוחה

למרות ההתפתחות של מערכות מודולריות סגורות ואינטגרציה שלהן עם כלים חדשים, חברות מתמודדות עם בעיה
רצינית – עליית המורכבות של מערכות אלו מקדימה את התועלת שלהן. הרעיון הראשוני של יצירת פלטפורמה
פרופירטית אחת המכסה את כל תהליכי העסק, הביא למרכזיות יתר, שבה כל שינוי דורש משאבים וזמן משמעותיים
להתאמה.

כפי שדנו בפרק "מיצילום תאגידי: כיצד נתונים מחברים תהליכי עסקים, "עבודה יעילה עם נתונים דורשת יצירת מערכת
אקולוגית פתוחה ואחידה, המאחדת את כל מקורות המידע

מרכיבי מפתח של המערכת האקולוגית

■ בחרו מאגר נתונים מתאים

□ (אזור) MySQL או PostgreSQL, עבור טבלאות וחשובים השתמשו בבסיסי נתונים - לדוגמה
- 3.17)

□ או במערכות (Google Drive, OneDrive) עבור מסמכים ודוחות ניתן להשתמש באחסון בענן
JSON התומכות בפורמט

□ וכלים אחרים לאחסון מרכזי Data Warehouse, Data Lakes הכירו את האפשרויות של

- (8.18 איור) וניתוח כמויות גדולות של מידע

■ יישמו פתרונות לגישה לנתונים קנייניים

☐ כדי לקבל נתונים SDK או API אם אתם משתמשים במערכות קנייניות, הגדרו גישה אליהן דרך (4.12 איור) לעיבוד חיצוני

☐ (4.113 איור) CAD הכירו את הפוטנציאל של כלים להנדסה הפוכה עבור פורמטים

☐ אשר אוספים באופן תקופתי נתונים מהיישומים או השרתים, ממירים אותם, ETL-Pipeline הגדרו (7.23 איור) לפורמטים פתוחים ומסודרים ושומרים במאגרים

☐ דונו בתוך הצוות בנושאים הקשורים להבטחת גישה לנתונים מבלי הצורך להשתמש בתוכנה קניינית

☐ זכרו: נתונים חשובים יותר מממשקים. הערך לטווח הארוך טמון במבנה ובזמינות המידע, ולא בכלים ספציפיים של ממשק משתמש

■ "במידול נתונים (CoE) בתחום הנתונים, כפי שדנו בפרק "מרכז מצוינות (CoE) שקלו להקים מרכז מצוינות (4.39 איור) או בדרכים אחרות להבטיח מומחיות בעבודה עם נתונים

מערכת אקולוגית לניהול נתונים יוצרת מרחב מידע אחיד, שבו כל המשתתפים בפרויקט עובדים עם מידע מעודכן, מהימן ומוסכם. זהו הבסיס לתהליכים דיגיטליים ניתנים להרחבה, גמישים ואמינים

מגלה את הפוטנציאל של נתונים: 5-10 צעדים לבגרות דיגיטלית

מעבר לכך, אינטגרציה טכנית, גורם חשוב נוסף להצלחה ביישום פתרונות דיגיטליים הוא קבלת המשתמשים הסופיים. גיוס לקוחות או משתמשים לשאלות הערכת האפקטיביות - זו משימה לשיפור חוויית המשתמש וניהול השינויים בחברה אם הפתרון לא משתלב בתהליך העבודה הרגיל או לא פותר בעיות אמיתיות של משתמשים או לקוחות, הוא לא יעשה בו שימוש, ולא יועילו שום צעדים או תמריצים נוספים

טרנספורמציה היא תהליך איטרטיבי, המבוסס על ניתוח נתונים על אינטראקציות משתמשים עם תהליכים חדשים, עם מחזורי בדיקה תכופים, משוב מתמשך ושיפורים

שלב 5. פתחו תרבות עבודה עם נתונים, הכשירו את הצוות ואספו משוב

אפילו המערכת המתקדמת ביותר לא תעבוד ללא מעורבות העובדים. יש ליצור סביבה שבה הנתונים משמשים מדי יום והצוות מבין את ערכם

במימוש פרויקטים ממשלתיים "לשנת AI-בדו"ח שפורסם על ידי ממשלת בריטניה "אנליטיקת נתונים ו מצוין [83], [כי הכשרת אנשי מקצוע בעלי הכישורים הנדרשים לעיבוד ופרשנות נתונים היא קריטית 2024 AI-להצלחה ביישום אנליטיקת נתונים ו

חוסר ידע בתחום ניתוח הנתונים הוא אחת הבעיות המרכזיות המגבילות את הדיגיטליזציה. מנהיגים רגילים להליכי עבודה מקובלים: מחזורי רבעון, יוזמות עדיפות ודרכים מסורתיות לקידום פרויקטים. לשינויים נדרש מנהיג מיוחד - בדרגה גבוהה מספיק כדי להשפיע, אך לא גבוהה מדי כדי שיהיה לו זמן ומוטיבציה להוביל פרויקט טרנספורמציה ארוך טווח

פעולות עיקריות:

(HiPPO), ההכרה בצורך לעבור מהחלטות סובייקטיביות, המבוססות על דעתו של עובד בעל שכר גבוה או הסכנה של דעות "HiPPO" לתרבות קבלת החלטות המסתמכת על עובדות ונתונים, כפי שנדון בפרק "-. בקבלת החלטות

■ ארגנו הכשרה מערכתית

☐ ערכו סדנאות לשימוש בנתונים מובנים, והזמינו מומחים מתחומים אחרים בכלכלה, שאין להם הטיות כלפי מוצרים ומושגים פופולריים כיום בתעשיית הבנייה

☐ דונו בגישות ובכלים לניתוח נתונים עם עמיתים, וכן למדו באופן עצמאי לעבוד עם כלים כמו Python, pandas ו-LLM.-

☐ צרו ספריית חומרי לימוד (עדיף עם סרטונים קצרים (בנושא מבנה נתונים ויצירת מודלים של --. נתונים

■ השתמשו בטכנולוגיות למידה מודרניות

☐ לתמיכה בעבודה עם קוד ונתונים, כולל יצירה, ריפקטורינג (LLM) השתמשו במודלים לשוניים וניתוח קוד, כמו גם עיבוד ופרשנות של מידע טבלאי.

☐ מלא בעבודה Pipeline יכול להיות מותאם ומוטמע בפתרון LLM למדו כיצד קוד שנוצר באמצעות (IDE).- בסביבת פיתוח לא מקוונת

כאשר מנהיג ממשיך לקבל החלטות "בדרך הישנה", אף סדנה לא תשכנע אנשים להתייחס לניתוח הנתונים ברצינות.

יצירת תרבות עבודה עם נתונים אינה אפשרית ללא משוב מתמשך. המשוב מאפשר לחשוף חסרונות בתהליכים, בכלים פורמליים. הערות חיוביות של משתמשים בפתרונות KPI ובאסטרטגיות, שאי אפשר לגלות דרך דוחות פנימיים או מדדי שלכם לא יביאו ערך מעשי. הערך טמון דווקא במשוב ביקורתי, במיוחד אם הוא מבוסס על תצפיות ועובדות ספציפיות עם זאת, קבלת מידע כזה דורשת מאמצים: יש לבנות תהליכים שבהם משתתפים - הן פנימיים והן חיצוניים - יכולים (לשתף הערות) אולי יש מקום לעשות זאת בעילום שם (ללא עיוותים וללא חשש שהדעה שלהם עשויה להשפיע על עבודתם האישית. חשוב שהם יעשו זאת ללא עיוותים וללא פחד מהשלכות שליליות על עצמם

כל למידה היא בסופו של דבר למידה עצמית

מילטון פרידמן, בלכלן וסטטיסטיקאי אמריקאי -

דבר שניתן להשיג רק באמצעות (ROI, KPI) יישום כלי ניתוח צריך להתלוות לאימות קבוע של היעילות שלהם בפועל משוב מובנה מעובדים, לקוחות ושותפים. זה מאפשר לחברות לא רק להימנע מחזרה על טעויות, אלא גם להסתגל במהירות לשינויים בסביבה. קיומו של מנגנון לאיסוף וניתוח משוב הוא אחד הסימנים לבגרות של הארגון, שעובר מיזמים דיגיטליים אקראיים למודל יציב של שיפור מתמשך (איור 2.25).

שלב 6. מפרויקטים פיילוט להרחבה

בחרו קרב מספיק גדול כדי שיהיה לו ערך, ומספיק קטן כדי לנצח. – ג'ונתן קוזול

השקת טרנספורמציה דיגיטלית "מיד ובכל מקום" היא מסוכנת מאוד. גישה יעילה יותר היא להתחיל מפרויקטים פיילוט ולהרחיב בהדרגה את הפרקטיקות המוצלחות.

פעולות עיקריות:

■ בחרו פרויקט מתאים לפיילוט

- ☐ - (איור 7.15) (KPI, ROI) הגדירו משימת עסקים ספציפית או תהליך עם תוצאות מדידות
- ☐ (QTO) כגון בדיקה אוטומטית של נתונים או חישוב כמויות עבודה, ETL בחרו תהליך אוטומציה של - (איור 5.210) Python ו-Pandas באמצעות
- ☐ קבעו מדדים ברורים להצלחה (למשל - לצמצם את הזמן הנדרש להכנת מפרטי בדיקות או דוחות על (בדיקות נתונים משבוע ליום אחד)

■ השתמשו בגישות איטרטיביות

- ☐ התחילו עם תהליכים פשוטים של המרת נתונים ויצירת המרה של נתונים בפורמטים שונים --לפורמטים הנדרשים לתהליכים שלכם) איור 4.12, איור 4.15
- ☐ בהדרגה הגדילו את מורכבות המשימות והרחיבו את האוטומציה של התהליכים, תוך יצירת -על בסיס בלוקים של קוד מתועד) איור 4.17, איור 7.218 (IDE) מלא בסביבת פיתוח Pipeline -
- ☐ תעדו והקליטו) עדיף באמצעות סרטונים קצרים (פתרונות מוצלחים ושתפו אותם עם קולגות או בקהילות מקצועיות

■ פתחו תבניות ודוקומנטציה נלווית להעתקת פתרונות דומים, כך שעמיתים) או משתתפי הקהילה המקצועית כולל משתמשים ברשתות החברתיות (יוכלו להשתמש בהם ביעילות

גישה הדרגתית מאפשרת לשמור על איכות גבוהה של שינויים ולא ליפול לכאוס של יישומים מקבילים. אסטרטגיית מקטן לגדול "ממזער סיכונים ומאפשרת ללמוד מטעויות קטנות, מבלי לאפשר להן להתפתח לבעיות קריטיות"

(CoE - למשל, מרכזי מומחיות) המעבר מגישה פרויקטלית, שבה העובדים מעורבים רק חלקית, להקמת צוותים קבועים מאפשר להבטיח פיתוח מתמשך של המוצר גם לאחר השקת הגרסה הראשונה שלו. צוותים אלה לא רק תומכים בפתרונות הקיימים, אלא גם ממשיכים לשפר אותם

זה מפחית את התלות באישורים ממושכים: חברי הצוות מקבלים את הזכות לקבל החלטות במסגרת תחום האחריות

שלהם. כתוצאה מכך, מנהלים משתחררים מהצורך בניהול מיקרו, והצוותים יכולים להתמקד ביצירת ערך אמיתי

פיתוח פתרונות חדשים הוא לא ספרינט, אלא מרתון. מצליחים בו אלה שמתמקדים מראש בעבודה ארוכת טווח ועקבית.

חשוב להבין שהטכנולוגיות דורשות פיתוח מתמשך. השקעה בפיתוח טכנולוגי ארוך טווח היא הבסיס לעבודה מוצלחת

שלב 7. השתמשו בפורמטים פתוחים של נתונים ופתרונות

חשוב להתמקד בפורמטים, (ועוד ERP, PMIS, CAFM, CDE) כפי שדנו בפרקים המוקדשים לפלטפורמות מודולריות פתוחים ואוניברסליים של נתונים, אשר יבטיחו עצמאות מפתרונות ספקים ויגבירו את זמינות המידע לכל המשתתפים בתהליך.

פעולות עיקריות:

■ המעבר מפורמטים סגורים לפורמטים פתוחים

☐ השתמשו בפורמטים פתוחים במקום פורמטים קנייניים, או מצאו אפשרות להגדיר ייצוא אוטומטי או המרה של פורמטים סגורים לפורמטים פתוחים.

☐ שהם סטנדרטים להחלפת נתונים בין Parquet, CSV, JSON, XLSX הטמיעו כלים לעבודה עם -רוב המערכות המודרניות

☐ אם העבודה עם גיאומטריה תלת-ממדית משחקת תפקיד חשוב בתהליכים שלכם, שקלו להשתמש OBJ או USD, glTF, DAE בפורמטים פתוחים כמו.

■ השתמשו בבסיסי נתונים וקטוריים לניתוח וחיפוש מידע יעיל

☐ -ושיטות אחרות כדי לפשט את העבודה עם גיאומטריה תלת-ממדית Bounding Box-השתמשו ב

☐ חשבו היכן ניתן להטמיע וקטורזציה של נתונים - המרת טקסטים, אובייקטים או מסמכים לייצוגים מספריים.

■ השתמשו בכלים לניתוח נתונים גדולים

☐ בפורמטים המתאימים (PDF, XLSX, CAD כגון) ארגנו את אחסון הנתונים ההיסטוריים שנצברו - (Apache Parquet, CSV, ORC) לניתוח

☐ התחילו ליישם שיטות סטטיסטיות בסיסיות ולעבוד עם דגימות מייצגות - או לפחות, הכירו את העקרונות הבסיסיים של הסטטיסטיקה.

☐ הטמיעו ולמדו כלים להמחשת נתונים וקשרים בין נתונים כדי להציג בצורה ברורה את תוצאות הניתוח. ללא המחשה איכותית, אי אפשר להבין באופן מלא את הנתונים עצמם ואת התהליכים המבוססים עליהם.

המעבר לפורמטים פתוחים של נתונים והטמעת כלים לניתוח, אחסון והמחשה של מידע מניחים את הבסיס לניהול דיגיטלי עצמאי ועמיד. זה לא רק מפחית את התלות בספקים, אלא גם מבטיח גישה שווה לנתונים לכל המשתתפים בתהליך

שלב 8. התחילו להטמיע למידת מכונה לחיזוי

בחברות רבות נאספו כמויות גדולות של נתונים - מעין "גייזרים מידעיים", אשר עד כה נותרו לא מנוצלים. נתונים אלו נאספו במסגרת מאות ואלפי פרויקטים, אך לעיתים קרובות נעשה בהם שימוש חד פעמי בלבד או שהם לא היו מעורבים בתהליכים נוספים. מסמכים ודגמים, המאוחסנים בפורמטים ובמערכות סגורות, לעיתים קרובות נתפסים כמשא כבד מיושן וחסר תועלת. עם זאת, למעשה הם מהווים משאב יקר ערך - בסיס לניתוח טעויות שנעשו, אוטומציה של פעולות שגרתיות ופיתוח פתרונות חדשניים למיון אוטומטי וזיהוי רכיבים בפרויקטים עתידיים.

המשימה המרכזית היא ללמוד כיצד לחלץ את הנתונים הללו ולהמירם לתועלת מעשית. כפי שכבר נדון בפרק "למידת מכונה וחיזויים", שיטות למידת מכונה יכולות לשפר באופן משמעותי את דיוק ההערכות והחיזויים בתהליכים שונים הקשורים לבנייה. השימוש המלא בנתונים שנצברו פותח את הדרך לשיפור היעילות, הפחתת הסיכונים ובניית תהליכים דיגיטליים עמידים.

פעולות עיקריות:

■ התחילו עם אלגוריתמים פשוטים

☐ לחיזוי מדדים חוזרים בסטי נתונים שבהם - LLM-נסו ליישם רגרסיה ליניארית - תוך שימוש בהנחות מ התלות במספר רב של גורמים אינה קיימת או מינימלית.

☐ (k- שכנים הקרובים-k ליישם את אלגוריתם theorectically שקלו באילו שלבים בתהליכים שלכם ניתן לדוגמה, למשימות סיווג, הערכת דמיון בין אובייקטים או חיזוי בהתבסס על אנלוגים היסטוריים - NN).

■ אספו ומסדו נתונים לצורך הכשרת מודלים

☐ אספו נתונים היסטוריים על פרויקטים במקום אחד ובפורמט אחיד.

☐ - אוטומטיים ETL עבדו על איכות וייצוגיות של דגימות ההכשרה, באמצעות

☐ - למדו לחלק את הנתונים לסטים של אימון ובדיקה, כפי שעשינו בדוגמה עם סט הנתונים של הטיטניק

■ שקלו אפשרויות להרחבת השימוש בשיטות למידת מכונה לפתרון מגוון רחב של בעיות - מחיזוי זמני ביצוע פרויקטים ועד אופטימיזציה של לוגיסטיקה, ניהול משאבים וזיהוי מוקדם של בעיות פוטנציאליות

למידת מכונה היא כלי המאפשר להפוך נתונים ארכיוניים לנכס יקר ערך לחיזוי, אופטימיזציה וקבלת החלטות מבוססות התחילו עם סטים קטנים של נתונים ומודלים פשוטים, והגדילו בהדרגה את המורכבות.

וטכנולוגיות מודרניות לאיסוף נתונים IoT שלב 9. אינטגרציה של

כבר חלק מתהליך גדול של - Teams-עולם הבנייה הופך במהירות לדיגיטלי: כל תמונה מהאתר, כל הודעה ב ואיסוף נתונים IoT, RFID שינה את הלוגיסטיקה, כך טכנולוגיות GPS-פרמטריזציה וטוקניזציה של המציאות. כפי ש אינטרנט של הדברים וחוזים חכמים, "אתר בנייה דיגיטלי IoT" אוטומטי משנות את תעשיית הבנייה. כפי שנדון בפרק עם חיישנים ומעקב אוטומטי - הוא העתיד של התעשייה

פעולות עיקריות:

■ **נופירוט תהליכים הקשורים אליהם RFID תוויות, IoT, הטמיעו מכשירי**

- ☐ העריכו באילו אזורים או שלבים בפרויקט התקנת חיישנים יכולה להניב את התשואה הגבוהה ביותר לדוגמה, למעקב אחר טמפרטורה, רעידות, לחות או תנועת ציוד - (ROI)
- ☐ למעקב אחר חומרים, כלים וציוד בכל שלב בשרשרת הלוגיסטית RFID-שקלו את השימוש ב
- ☐ לצורך עיבוד Apache NiFi חשבו כיצד ניתן לשלב את הנתונים שנאספו במערכת מידע אחת, כמו (7.45 איור)- וניתוח אוטומטי בזמן אמת

■ **צרו מערכת מעקב בזמן אמת**

- ☐ פתחו לוחות מחוונים למעקב אחר מדדים מרכזיים של תהליך או פרויקט באמצעות כלי ויזואליזציה Power BI או Streamlit, Flask כמו
- ☐ - (7.42 איור) הגדרו התראות אוטומטיות שיתריעו על סטיות קריטיות מהתוכנית או מהנורמות
- ☐ העריכו את הפוטנציאל של תחזוקה חיזויית של ציוד בהתבסס על הנתונים שנאספו ודפוסים שנחשפו (9.36 איור)-

■ **שילבו נתונים ממקורות שונים**

- ☐ התחילו בוויזואליזציה של מודל הנתונים ברמה הפיזית - שיקפו את מבנה זרמי המידע והפרמטרים - (4.31 איור) ERP ופלטפורמות IoT מכשירי, CAD המרכזיים המגיעים ממערכות
- ☐ התחילו ביצירת תיאור גולמי של פלטפורמה אחת, המיועדת לניתוח נתונים ולתמיכה בקבלת החלטות ניהוליות. תעדו את הפונקציות המרכזיות, מקורות הנתונים, המשתמשים, כמו גם תרחישי השימוש (4.37 איור) המיועדים

ככל שתחברו תהליכים אמיתיים לעולם הדיגיטלי מוקדם יותר, כך תוכלו לנהל אותם באמצעות נתונים - ביעילות, שקיפות ובזמן אמת.

שלב 10. התכנון לעתיד של שינויים בתעשייה

חברות בנייה נמצאות כל הזמן תחת לחצים מהסביבה החיצונית: משברים כלכליים, קפיצות טכנולוגיות, שינויים רגולטוריים. בדומה ליער, שנאלץ לעמוד בגשם, שלג, בצורת ושמש קופחת, חברות חיות בתנאים של התאמה מתמשכת וכפי שעצים רוכשים עמידות לקור ולבצורת בזכות מערכת שורשים עמוקה, כך רק אותן ארגונים שיש להם בסיס יציב של תהליכים אוטומטיים, יכולת לחזות שינויים וגמישות בהתאמת אסטרטגיות, שומרים על חיוניותם ותחרותיותם כפי שנאמר בפרק "אסטרטגיות הישרדות: יצירת יתרונות תחרותיים", תעשיית הבנייה נכנסת לשלב של טרנספורמציה דיסקלית. האינטראקציה בין הלקוח למבצע מתקדמת למודל של אובריזציה, שבו שקיפות, ניבוי וכלים דיגיטליים דוחקים את הגישות המסורתיות. במציאות החדשה הזו, המנצחים אינם הגדולים ביותר, אלא הגמישים והטכנולוגיים ביותר

פעולות עיקריות:

■ **נתחו את הפגיעויות של העסק בהקשר של נתונים פתוחים**

- ☐ העריכו כיצד דמוקרטיזציה של גישה לנתונים במסגרת האובריזציה יכולה להשפיע באופן הרסני על

- (10.15 איור) יתרונות התחרות שלכם ועל העסק שלכם

☐ שקלו אסטרטגיה למעבר מתהליכים לא שקופים ומבודדים למודלים עסקיים המבוססים על פתרונות (2.25 איור) פתוחים, תאימות מערכות ושקיפות נתונים

■ פתחו אסטרטגיה דיגיטלית ארוכת טווח

☐ קבעו אם אתם שואפים להיות מובילי חדשנות או מעדיפים תרחיש "מגיב", שבו תחסכו את המשאבים שלכם

☐ (ETL ו-LLM הטמעת) פרטו את השלבים: קצרים) אוטומציה של תהליכים, מבנה נתונים, (בינוניים ארוכים) מערכות אקולוגיות דיגיטליות, מאגרי נתונים מרכזיים)

■ חשבו על הרחבת תיק השירותים

☐ נדון. (שירותי עיבוד נתונים, ESG, ממוקדים ביעילות אנרגטית) שקלו להציע שירותים חדשים במודלים עסקיים חדשים בפרק הבא

☐ שאפו למצב את עצמכם כשותף טכנולוגי אמין, המלווה את כל מחזור חיי האובייקט – מעיצוב ועד הפעלה. האמון בכם צריך להתבסס על גישה מערכתית, שקיפות תהליכים ויכולת לספק פתרונות טכנולוגיים ברי קיימא

בתנאי טרנספורמציה, המנצחים אינם אלו המגיבים לשינויים, אלא אלו הפועלים לפניהם. גמישות, פתיחות ובגרות דיגיטלית – הם הבסיס לעמידות בבניית המחר

מפת דרכים לשינוי: מהכאוס לחברה מונחית נתונים

התוכנית הבאה יכולה לשמש כנקודת ייחוס ראשונית – נקודת מוצא לעיצוב אסטרטגיית טרנספורמציה דיגיטלית מבוססת נתונים:

■ ביקורת וסטנדרטים: נתחו את המצב הנוכחי, אחידו את הנתונים

■ מבנה וסיווג נתונים: אוטומטו את המרת נתונים לא מובנים וחלשים

■ אוטומציה של קבוצות, חישובים וחשבונות: השתמשו בכלים ובספריות פתוחות לאוטומציה

■ הקימו צוות פנימי שיבנה מערכת אקולוגית נתונים אחידה בחברה: COE-מערכת אקולוגית ו

■ להחלטות מבוססות נתונים HiPPO תרבות ולמידה: התרחקו מהחלטות

■ פיילוטים, משוב והרחבה: פעלו באופן איטרטיבי: בדקו שיטות חדשות בהיקף מוגבל, אספו משוב מבוסס והרחיבו בהדרגה את הפתרונות

■ פורמטים פתוחים: השתמשו בפורמטים אוניברסליים ופתוחים לעצמאות מספקי תוכנה

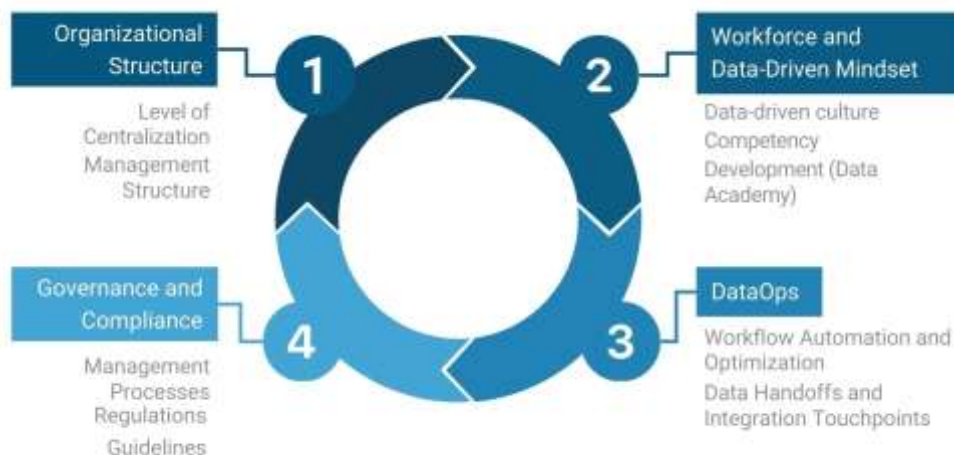
■ לחיזוי ואופטימיזציה ML למידת מכונה: הטמיעו בתהליכים אלגוריתמים של

■ ואתר בנייה דיגיטלי: אינטגרציה טכנולוגית מודרנית לאיסוף נתונים בתהליכים IoT

■ התאמה אסטרטגית: התכוננו לשינויים עתידיים בתעשייה

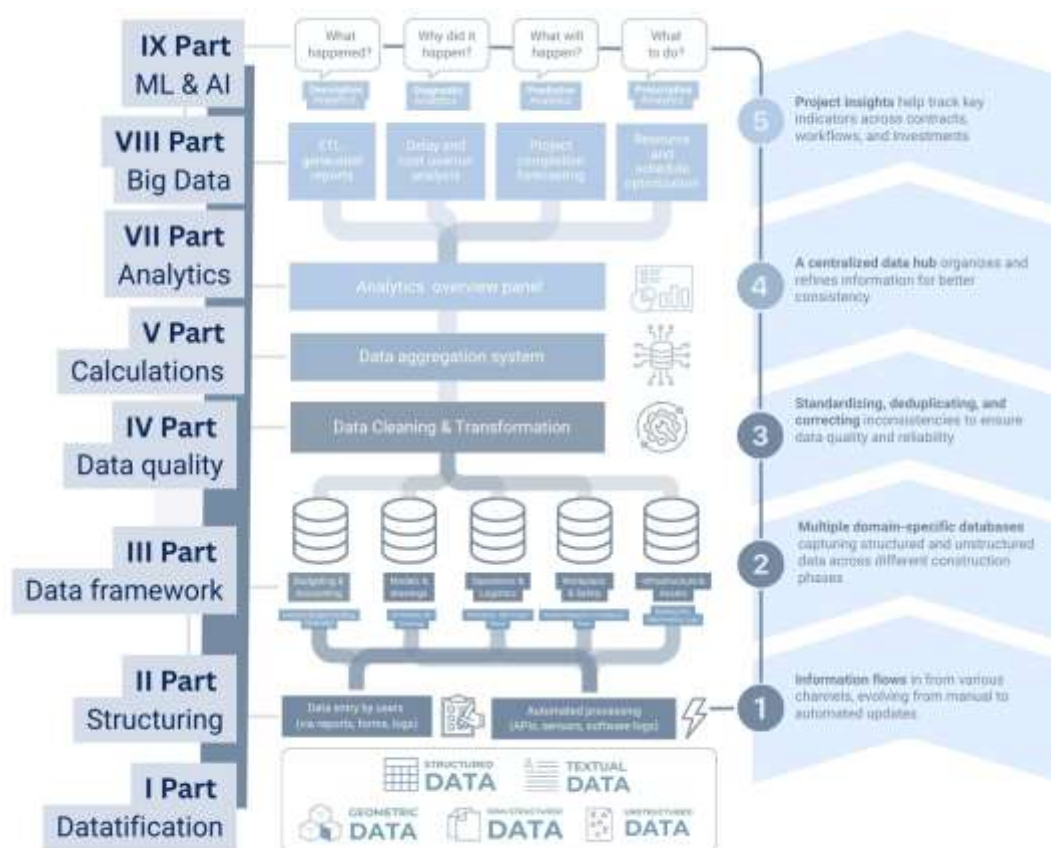
העיקר – זכרו ש"נתונים בפני עצמם אינם משנים חברה: אנשים המיומנים בעבודה עם נתונים הם אלו שמשנים אותה. "השקיעו בתרבות, תהליכים שקופים ושאיפה לשיפורים מתמשכים

הגישה המערכתית מאפשרת לעבור מיוזמות דיגיטליות מפוזרות למודל ניהול מונחה נתונים שלם, שבו ההחלטות מתקבלות לא על סמך אינטואיציה או השערות, אלא על סמך נתונים, עובדות והסתברויות מחושבות מתמטית טרנספורמציה דיגיטלית של תעשיית הבנייה אינה רק יישום טכנולוגיות, אלא יצירת מערכת אקולוגית עסקית שבה המידע על הפרויקט מועבר בצורה חלקה ואיטרטיבית בין מערכות שונות. אלגוריתמים של למידת מכונה מספקים ניתוח אוטומטי, מתמשך, חיזוי ואופטימיזציה של תהליכים. בסביבה כזו, ספקולציות ונתונים מוסתרים מאבדים את הרלוונטיות שלהם - נשארים רק מודלים מוכחים, חישובים שקופים ותוצאות צפויות



מרכיבי מפתח לניהול נתונים מוצלח ברמת החברה 10.23. ריש.

אם תרצו לחזור לאחת (2.25. ריש) כל חלק בספר תואם לשלב ספציפי בעיבוד ובניתוח נתונים בפרויקט בנייה מהנושאים שנדונו קודם לכן ולהסתכל עליה מנקודת מבט כוללת של זרימת השימוש בנתונים - תוכלו להתמקד בכותרות 10.24. החלקים המצוינות בריש



-מהדיגיטציה של מידע ועד לניתוח ובינה מלאכותית: (2.25. ריש) חלקי הספר בהקשר של צינור עיבוד הנתונים 10.24. ריש

ללא קשר לגודל הארגון שלכם, רמת הבשלות הטכנולוגית או התקציב, אתם יכולים להתחיל לנוע לעבר גישה מונחית נתונים כבר היום. אפילו צעדים קטנים בכיוון הנכון יובילו לתוצאות עם הזמן.

טרנספורמציה מונחית נתונים אינה פרויקט חד פעמי, אלא תהליך מתמשך ואיטרטיבי של שיפור, הכולל יישום כלים חדשים, סקירת תהליכים ופיתוח תרבות קבלת החלטות על סמך נתונים.

בנייה בתעשייה 5.0: כיצד להרוויח כאשר כבר אי אפשר להסתיר יותר

במשך זמן רב, חברות הבנייה הרוויחו מהיעדר שקיפות בתהליכים. המודל העסקי העיקרי הפך להיות ספקולציה - העלאת סגורות, שאינן נגישות לביקורת חיצונית. גישה מוגבלת ERP, PMIS מחירי חומרים, כמויות עבודה ורווחים במערכות של לקוחות ואנשיהם המהימנים לנתוני הפרויקט המקוריים יצרה קרקע פורייה לתוכניות שבהן בדיקת אמינות החישובים הפכה כמעט לבלתי אפשרית.

הגעת נתונים, LLM עם זאת, מודל זה מאבד במהירות את הרלוונטיות שלו. עם דמוקרטיזציה של גישה לנתונים, הופעת התעשייה עוברת לסטנדרט עבודה חדש, ETL פתוחים, כמו גם כלים לאוטומציה.

כתוצאה מכך, היעדר שקיפות מפסיק להיות יתרון תחרותי - בקרוב הוא יהפוך לנטל שיהיה קשה להיפרד ממנו. שקיפות הופכת מאופציה לדרישה הכרחית כדי להישאר בשוק.

עם מי יעסקו הלקוחות - בנקים, משקיעים, לקוחות פרטיים, הון פרטי, לקוחות ממשלתיים - במציאות הדיגיטלית החדשה? התשובה ברורה: עם מי שיכול לספק לא רק תוצאה, אלא גם הסבר לכל צעד בדרך אליה. בתנאים של עלייה בנפח הנתונים הפתוחים, שותפים ולקוחות יבחרו בחברות שמבטיחות שקיפות, דיוק ותוצאות צפויות.

על הרקע הזה מתהווים מודלים עסקיים חדשים, שבבסיסם לא ספקולציה, אלא ניהול נתונים ואמון.

מכירת תהליכים במקום מטרים רבועים: הנכס המרכזי אינו הסכמים על הנחות על בטון, אלא אמון ויעילות הערך העיקרי יהיה תחזיות תוצאה מבוססות על נתונים מהימנים ומאומתים. חברות מודרניות ימכרו לא את אובייקט הבנייה כשלעצמו, אלא:

- לוחות זמנים מדויקים וגרפים שקופים של עבודות;
- אומדנים מבוססים, מאומתים על ידי חישובים;
- אפשרות למעקב דיגיטלי מלא ושליטה בכל שלב של הפרויקט.

שיטת אספקת נתונים מוכנים למשתמשים (דרך "Data-as-a-Service" הנדסה וניתוח כשירות: מודל האינטרנט, כשירות, (שבו כל פרויקט הופך לחלק משרשרת נתונים דיגיטלית, והערך של העסק הוא ביכולת ות לפלטפורמות אינטליגנטיות המציעות פתרונות על בסיס Transform לנהל את השרשרת הזו. חברות מת אוטומציה וניתוח:

- הכנת אומדנים ותוכניות אוטומטיות ושקופות;
- הערכת סיכונים ולוחות זמנים על בסיס אלגוריתמים של למידת מכונה;
- (יעילות אנרגטית, CO₂, ESG) חישוב מדדים סביבתיים;
- יצירת דוחות ממקורות פתוחים ניתנים לאימות.

- מוצריזציה של ניסיון הנדסי: הידע של החברה יכול לשמש שוב ושוב בתוך החברה ולהתפשט כמוצר נפרד מה שמייצר מקור הכנסה נוסף באמצעות שירותים דיגיטליים. בתנאים החדשים, חברות יוצרות לא רק פרויקטים, אלא גם נכסים דיגיטליים.

- ספריות של רכיבים ותבניות אומדנים;

- מודולים אוטומטיים לאימות

- תוספים וסקריפטים בקוד פתוח) מכירת ייעוץ (לעבודה עם נתונים

■ סוג חדש של חברה: אינטגרטור מונחה נתונים: משתתף בשוק שאינו תלוי בספקי תוכנה או מערכות מודולריות ספציפיות ואינו "סגור" בממשק של תוכנה אחת. הוא פועל בחופשיות עם נתונים - ועל כך הוא בונה את התחרותיות שלו. חברת הבנייה של העתיד אינה רק קבלן, אלא אינטגרטור מידע, המסוגל עבור הלקוח לבצע את הפונקציות הבאות:

- ;לאחד נתונים ממקורות שונים ולבצע ניתוח

- ;להבטיח שקיפות ומהימנות של תהליכים

- ;לייעץ על אופטימיזציה של תהליכי עסקים

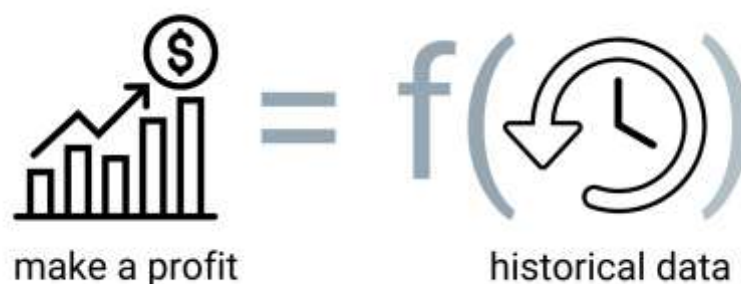
- LLM, ETL ו-Pipelines, לפתח כלים הפועלים במערכת אקולוגית של נתונים פתוחים

מסמנת את סוף "עידן הממוצעים הידניים" ופגישות הערב של מנכ"לים עם מחלקת (2.112 איור) תעשייה 5.0 האומדנים והחשבונאות. כל מה שהיה קודם מוסתר - חישובים, אומדנים, כמויות - הופך לגלוי, נבדק וברור אפילו לא מומחה. המרוויחים יהיו אלו שיתמקדו מחדש ראשונים. כל השאר - יישארו מאחור בכלכלה הדיגיטלית החדשה של תחום הבנייה.

סיכום

ענף הבנייה נכנס לעידן של שינויים מהותיים. מההקלטות הראשונות על לוחות חימר ועד למאגרים של נתונים דיגיטליים המגיעים משרתי תכנון ואתרי בנייה, ההיסטוריה של העבודה עם מידע בבנייה תמיד שיקפה את רמת הבשלות של הטכנולוגיות בזמנה. כיום, עם הגעת האוטומציה, פורמטים פתוחים ומערכות ניתוח אינטליגנטיות, הענף מתמודד לא עם אבולוציה הדרגתית, אלא עם טרנספורמציה דיגיטלית מהירה.

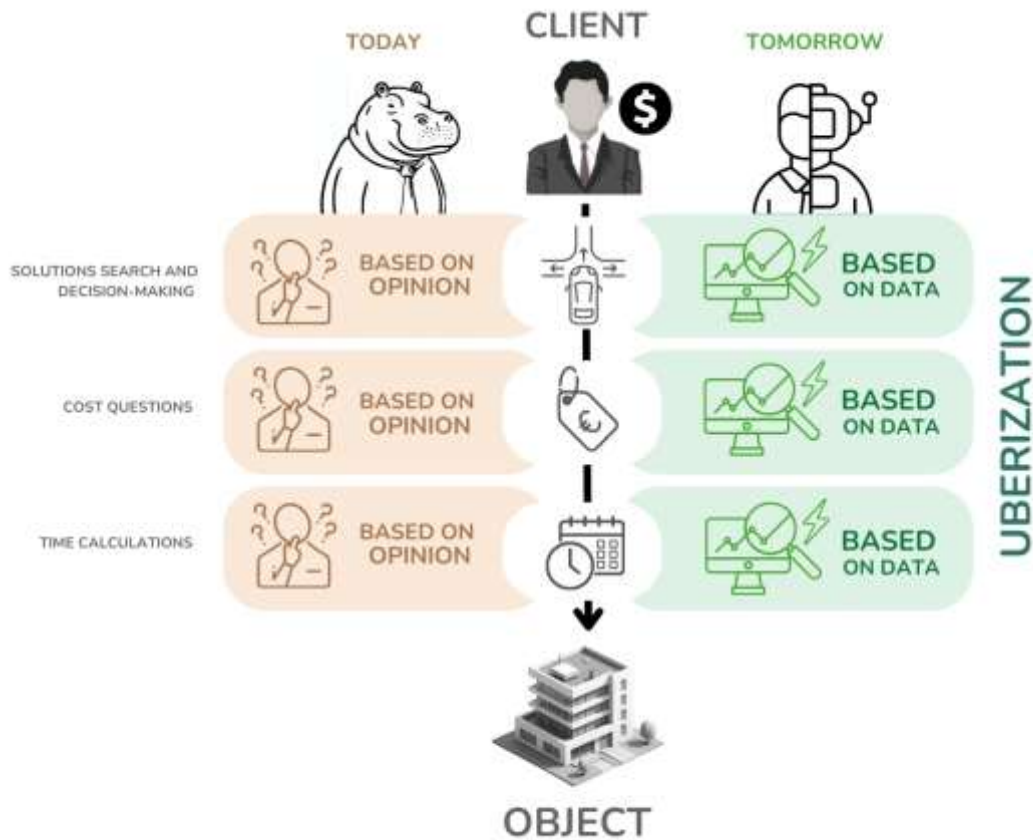
כמו בענפים אחרים בכלכלה, לבנייה יש צורך לחשוב מחדש לא רק על הכלים, אלא גם על עקרונות העבודה. חברות שבעבר קבעו את השוק ושימשו כמתווכים העיקריים בין הלקוח לפרויקט, מאבדות את מעמדן הייחודי. האמון ויכולת העבודה עם נתונים עולים למרכז הבמה: מהאיסוף והמבנה שלהם ועד לניתוח, חיזוי ואוטומציה של החלטות.



- נתונים היסטוריים מובנים - דלק לעסק יעיל ומנוהל 10.21. ריס

בספר זה נידונו בהרחבה העקרונות המרכזיים של ניהול נתונים בענף הבנייה - מהאודיט והסטנדרטיזציה ועד לאוטומציה של תהליכים, שימוש בכלי ויזואליזציה והטמעת אלגוריתמים אינטליגנטיים. בדקנו כיצד גם עם משאבים מוגבלים ניתן לבנות ארכיטקטורת נתונים פועלת ולהתחיל לקבל החלטות לא על סמך אינטואיציה, אלא על סמך בלבד - היא הופכת לבסיס IT-עובדות ניתנות לבדיקה. העבודה עם נתונים מפסיקה להיות משימה של מחלקת הלתרבות ניהולית, שעליה תלויה הגמישות של החברה, יכולת ההתאמה שלה ועמידותה לטווח הארוך.

השימוש בטכנולוגיות של למידת מכונה, מערכות עיבוד אוטומטיות, דיגיטליות כפולים ופורמטים פתוחים כבר היום, מאפשר להוציא את הגורם האנושי מהתהליך היכן שבעבר הוא היה קריטי. הבנייה מתקדמת לעבר אוטונומיה וניהול, שבו המעבר מרעיון ליישום פרויקט ניתן להשוואה לניווט במצב טייס אוטומטי: ללא תלות בהחלטות סובייקטיביות. (ריס) 10.22. ללא צורך בהתערבות ידנית בכל שלב, אך עם מעקב דיגיטלי מלא ושליטה



לניתוח נתונים יתקדם בראש ובראשונה על (HiPPO) המעבר מקבלת החלטות על סמך דעות של מומחים חשובים **10.22**. ריס - ידי הלקוח.

לאחר שלמדנו את השיטות, העקרונות והכלים המוצגים בספר זה, תוכלו להתחיל לקבל החלטות בחברתכם המבוססות מוכנים לסביבת ETL Pipelines להעתיק, LLM-על נתונים ולא על אינטואיציה. תוכלו גם להפעיל רצפי מודולים בולעדי נתונים אוטומטית, תוך קבלת מידע בצורה הנדרשת לכם. בהמשך, בהתבסס על פרקים (IDE) הפיתוח שלכם בספר העוסקים בנתונים גדולים ולמידת מכונה, תוכלו ליישם תרחישים מורכבים יותר - לחלץ ידע חדש מנתונים היסטוריים וליישם אלגוריתמים של למידת מכונה לחיזוי ואופטימיזציה של התהליכים שלכם.

נתונים ופרוצדורות פתוחים יהפכו לבסיס להערכות מדויקות יותר של עלויות ולוחות זמנים של פרויקטים, מה שימנע מחברות הבנייה את האפשרות לסחור בנתונים לא שקופים. זהו גם אתגר וגם הזדמנות עבור הענף לחשוב מחדש על תפקידו ולהתאים את עצמו לסביבה החדשה, שבה שקיפות ויעילות יהפכו לגורמים המרכזיים להצלחה.

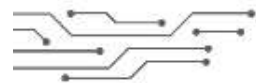
מוכנות לקבל ידע וליישם אותו בפועל היא המפתח להצלחה בעידן של טרנספורמציה דיגיטלית.

חברות שמבינות זאת ראשונות, יזכו ליתרון בתנאי התחרות הדיגיטלית החדשה. אך חשוב להבין: נתונים בפני עצמם לא משנים דבר. רבים יצטרכו לשנות את דרך החשיבה הרגילה שלהם, ולצורך כך נדרש תמריץ. החברה שלכם צריכה לבחון מחדש את הגישה שלה להחלפת נתונים.

את החברה משנים אנשים שמסוגלים לעבוד עם הנתונים הללו, לפרש אותם, להשתמש בהם לאופטימיזציה וליצור על

בסיסם ארכיטקטורה חדשה של תהליכים

אם אתם קוראים את השורות הללו, אתם מוכנים לשינויים ואתם כבר צעד אחד קדימה. תודה שבחרתם בדרך זו. ברוכים
הבאים לעידן של טרנספורמציה דיגיטלית



על המחבר



שמי ארטיום בויקו. הדרך שלי באתרי בנייה החלה בשנת 2007 - בעבודתי ככורה במכרה שיחים בעיר הולדתי, בזמן שלמדתי באוניברסיטת סנט פטרבורג להנדסה מינרלית ובנייה תת-קרקעית. בצד האחורי של כריכת הספר הזה תוכלו לראות מפוצץ במכרה, שם חצבנו ופוצצנו מאות קוביות של שיחים דליקים. הקריירה שלי התפתחה בכיוונים שונים - מעבודה כעובד מכרה ובבניית רכבת תחתית ועד לטיפול תעשייתי מתקין גגות וציוד מעליות. זכיתי להשתתף בפרויקטים במגוון רחב של היקפים: מבניית בתים פרטיים ועד פרויקטים תעשייתיים גדולים באזורים שונים בעולם.

עם הזמן, העבודה שלי עברה מבנייה פיזית לניהול מידע ותהליכים דיגיטליים. מאז 2013 עבדתי בתפקידים שונים, בחברות בנייה קטנות, בינוניות וגדולות בכמה אזורים בגרמניה, מתכנן ועד מנהל ניהול נתונים. בתחום ניהול הנתונים, עסקתי באופטימיזציה ERP, CAD (BIM), MEP, FEM, CMS. הניסיון שלי כולל עבודה עם נתונים במערכות אוטומציה של תהליכים, כמו גם בניתוח, למידת מכונה, ועיבוד נתונים בשלבי תכנון, חישובים וביצוע עבודות בנייה. בחברות העוסקות בבניית פרויקטים תעשייתיים, מגורים, תשתיות ושירותים ציבוריים.

- מאז 2003 אני עובד עם תוכנה פתוחה ונתונים פתוחים. במהלך הזמן הזה יישמתי מספר פרויקטים אינטרנטיים פתוחים CMS-מאתרים וחנויות אינטרנטיות ועד לאפליקציות אינטרנט שלמות, תוך שימוש בפתרונות קוד פתוח ו בנייה מודרניים, בעלות ארכיטקטורה מודולרית, התאמה גבוהה ונגישות ERP-הפלטפורמות הללו, דומות במידה רבה ל ניסיון זה הניח את היסוד לגישה המקצועית שלי - התמקדות בטכנולוגיות פתוחות ובתרבות של פיתוח משותף. אני שואף לקדם את הכבוד לקוד פתוח ולחלוקת ידע חופשית בתעשיית הבנייה. העבודה שלי לשיפור נגישות הנתונים בתעשיית הבנייה התבטאה בהקמת מספר קהילות ברשתות החברתיות לדיון בשאלות של שקיפות נתונים ושימוש בקוד פתוח בבנייה, כמו גם בהשקת מספר סטארט-אפים המפתחים פתרונות להבטחת גישה לנתונים ממערכות ופלטפורמות סגורות שונות.

CAD התרומה שלי לקהילה המקצועית מתבטאת בהשתתפות כדובר בכנסים העוסקים בנושאים של אינטרופרוביליות של למידת מכונה ובינה מלאכותית, כמו גם במאמרים שפורסמו בפרסומים אירופיים (BIM), ERP, 4D-5D, LLM מפה מקיפה של פתרונות תוכנה, "BIM המוקדשים לתחום הבנייה. אחד מהישגים הבולטים שלי הוא יצירת "היסטוריית ומשחקי לובי", שתורגמה BIM חשוכים לניהול נתונים בתעשיית הבנייה. סדרת המאמרים שלי בת 7 חלקים "התפתחות למספר שפות, זכתה להכרה רחבה כניסיון להאיר את הדינמיקה הנסתרת של התפתחות הסטנדרטים הדיגיטליים.

כך עברתי מהפקת חומרי גלם - להפקת וסידור נתוני בנייה. אני תמיד פתוח לדיאלוג מקצועי, רעיונות חדשים ופרויקטים משותפים. אקבל בברכה כל משוב ואשמח לקבל את הודעותיכם או לראותכם בין העוקבים ברשתות החברתיות. תודה רבה לכם על כך שקראתם את הספר הזה עד הסוף! אשמח אם ספר זה יעזור לכם להבין טוב יותר את הנושא של נתונים בתעשיית הבנייה.

משוב

דעת הקוראים משחקת תפקיד חשוב בהתפתחות העתידית של הפרסומים ובבחירת הנושאים המועדפים הערות על אילו רעיונות היו מועילים ואילו עוררו ספקות ודורשים הבהרות נוספות או הפניות למקורות, הן בעלות ערך רב. הספר כולל מגוון רחב של חומרים והערות אנליטיות, חלקם עשויים להיראות שנויים במחלוקת או סובייקטיביים. אם במהלך הקריאה תגלו אי דיוקים, מקורות שצוינו לא נכון, חוסר עקביות: לוגית או שגיאות הקלדה - אשמח לקבל את תגובותיכם, מחשבותיכם או ביקורתכם, שתוכלו לשלוח לכתובת: boikoartem@gmail.com. או דרך הודעות בלינקדאין: [linkedin.com/in/boikoartem](https://www.linkedin.com/in/boikoartem)

ברשתות החברתיות - החלפת חוויות קריאה Data-Driven Construction אודה לכל אזכור של הספר מסייעת להפיץ מידע על נתונים פתוחים וכלים ותומכת בעבודתי

הערה על התרגום

ספר זה תורגם בעזרת טכנולוגיות בינה מלאכותית. זה אפשר להאיץ משמעותית את תהליך התרגום. עם זאת, כמו בכל פעולה טכנולוגית, עשויות להתרחש שגיאות או אי דיוקים. אם תבחינו במשהו שנראה לא נכון או מתורגם בצורה לא מדויקת, אנא כתבו לי. ההערות שלכם יסייעו לשפר את איכות התרגום

DATA-DRIVEN CONSTRUCTION קהילות

זהו מקום שבו תוכלו לשאול שאלות ולשתף את הבעיות והפתרונות שלכם

DataDrivenConstruction.io: <https://datadrivenconstruction.io>

LinkedIn: <https://www.linkedin.com/company/datadrivenconstruction/>

Twitter: <https://twitter.com/datadrivenconst>

Telegram: <https://t.me/datadrivenconstruction>

YouTube: <https://www.youtube.com/@datadrivenconstruction>

נוספים ומושגים מיומנויות

מגוון נדונה DataDrivenConstruction בספר, הבנייה בתעשיית נתונים עם עבודה של המרכזיים לעקרונות בנוסף הם אך, כללי באופן רק מוצגים חלקם. נתונים עם העובד למומחה הנדרשות ומיומנויות תוכניות, מושגים של רחב בפרקטיקה קריטי תפקיד משחקים

שבו מוצגות קישורים לחומרים נוספים על מיומנויות DataDrivenConstruction.io, קורא מעוניין יכול לבקר באתר דוגמאות לעיבוד נתונים בפרויקט ETL, בניית תהליכי Python ו-Pandas מפתח. חומרים אלה כוללים עבודה עם מערכות לעיבוד נתוני גדולים, כמו גם גישות מודרניות להמחשה ואנליטיקה של נתוני בנייה, CAD, בנייה

וכל הדוגמאות המעשיות נעשה שימוש במגוון כלים ותוכנות קוד "DataDrivenConstruction" במהלך הכנת הספר פתוח. המחבר מביע תודה למפתחים ולשותפים של הפתרונות הבאים

- הבסיס לעבודה עם נתונים ואוטומציה – Python ו-Pandas
- ספריות לניתוח נתונים ולמידת מכונה – Scikit-Learn ו-NumPy, Matplotlib
- כלים לאחסון ועיבוד כמויות גדולות של נתוני בנייה – SQL ו-Apache Parquet
- כלים פתוחים לעבודה עם נתונים בפורמטים פתוחים (BIM) Open Source CAD
- מערכות לתיאום ואוטומציה של תהליכי עבודה – N8n, Apache Airflow, Apache NiFi
- DeepSeek, LLaMa, Mistral – Open Source LLM

תודה מיוחדת לכל המשתתפים בדיונים בנושא נתונים פתוחים וכלים בקהילות מקצועיות וברשתות חברתיות, שביקורתם הערותיהם ורעיונותיהם סייעו לשפר את התוכן והמבנה של ספר זה

שבו מתפרסמים לא רק עדכונים על הספר DataDrivenConstruction.io, עקבו אחרי התפתחות הפרויקט באתר ותיקונים, אלא גם פרקים חדשים, חומרים לימודיים ודוגמאות מעשיות ליישום השיטות המתוארות

מקסימום נוחות עם הגרסה המודפסת

לשימוש נוח יותר וגישה מהירה. Data-Driven Construction אתם מחזיקים בידיכם את הגרסה הדיגיטלית החינמית של לחומרים, אנו ממליצים לשקול את הגרסה המודפסת



■ תמיד בהישג

יד: ספר

בפורמט מודפס

יהפוך לכלי

עבודה אמין,

המאפשר

למצוא

ולהשתמש

במהירות

בויזואליזציות

ובתרשימים

בכל מצב

עבודה

■ איכות גבוהה

של איורים: כל

התמונות

והגרפים

בגרסה

המודפסת

מוצגים

באיכות

מקסימלית

■ גישה מהירה

למידע: ניווט

נוח, אפשרות לעשות הערות, סימניות ולעבוד עם הספר בכל מקום

ברכישת הגרסה המודפסת המלאה של הספר, אתם מקבלים כלי נוח לעבודה נוחה ויעילה עם מידע: אפשרות להשתמש בחומרים ויזואליים במשימות יומיומיות, למצוא במהירות את התרשימים הנדרשים ולעשות הערות. בנוסף, רכישתכם תומכת בהפצת ידע פתוח.

datadrivenconstruction.io/books: ניתן להזמין את הגרסה המודפסת של הספר בכתובת

הזדמנות ייחודית למיצוב אסטרטגי



אנו מציעים לכם לפרסם חומרים פרסומיים בגרסה החנימית של המהדורה הגרסה בתשלום של המהדורה בשנה הראשונה לאחר DataDrivenConstruction. הפרסום משכה את תשומת הלב של מומחים מיותר מ-50 מדינות ברחבי העולם – מאמריקה הלטינית ועד אזור אסיה-פסיפיק. לדיון על תנאי שיתוף פעולה אישיים ולקבלת מידע מפורט על אפשרויות פרסום, אנא מלאו את טופס המשוב באתר הרשמי datadrivenconstruction.io הספר

DATA-DRIVEN CONSTRUCTION פרקי הספר זמינים באתר



באתר, שבו מתפרסמים Data-Driven Construction אתם יכולים לקרוא את פרקי הספר בהדרגה חלקי הספר, כך שתוכלו למצוא במהירות את המידע הנדרש ולהשתמש בו בעבודתכם בנוסף, באתר תמצאו פרסומים נוספים בנושאים דומים, כמו גם דוגמאות ליישומים ולפתרונות שיסייעו לכם לפתח את המיומנויות שלכם וליישם נתונים בבנייה

הורידו את הגרסאות האחרונות של הספר מהאתר הרשמי



זמינות להורדה DataDrivenConstruction הגרסאות העדכניות והאחרונות של הספר אם אתם רוצים לקבל עדכונים עם פרקים חדשים. datadrivenconstruction.io באתר של הספר, טיפים מעשיים או סקירות של יישומים חדשים, הירשמו לניוזלטר

■ תהיו הראשונים להכיר את הפרקים החדשים של הספר

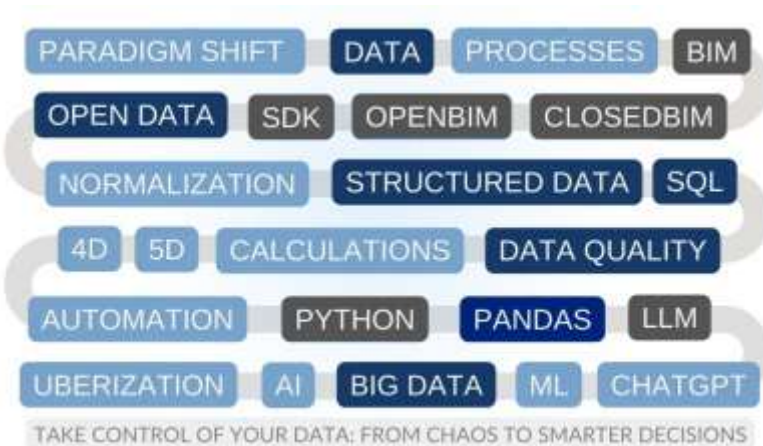
■ קבלו מקרים מעשיים וטיפים על אנליטיקה ואוטומציה בבנייה

■ לעקוב אחרי מגמות, פרסומים ודוגמאות של יישומים

!כדי להירשם datadrivenconstruction.ioעברו ל

ייעוץ, סדנאות והדרכה: DATA DRIVEN CONSTRUCTION

וייעוץ סייעו לעשרות חברות בנייה מובילות ברחבי העולם לשפר Data Driven Construction תוכניות ההדרכה של נמצאים Data Driven Construction את היעילות, להפחית עלויות ולשפר את איכות ההחלטות. בין לקוחות השחקנים הגדולים בשוק עם מחזור של מיליארדי יורו, כולל חברות בנייה, ייעוץ וטכנולוגיה.



למה לבחור בנו?

רלוונטיות: אנו

מדברים על המגמות

וההבנות המרכזיות

בתעשייה

פרקטיקה: אנו עוזרים למומחים לפתור

בעיות יומיומיות בצורה יעילה דרך

PoC.

גישה מותאמת

אישית: אנו

מתחשבים בצרכים

הייחודיים של העסק שלך, ומספקים את התועלת המרבית מההדרכה והייעוץ

Data Driven Construction תחומי הפעולה המרכזיים של צוות

ניהול איכות נתונים: אנו עוזרים לפרמטריזציה של משימות, איסוף דרישות, בדיקה והכנה של נתונים לעיבוד אוטומטי.

PDF, Excel, ומחלקים נתונים מדוא"ל ETL כריית נתונים - חילוף ומבנה נתונים: אנו מגדירים תהליכי תמונות ומקורות נוספים.

CAD (BIM) ומבני RVT, IFC, DWG אנו אוספים, מסדרים ומנתחים מידע מקבצי BIM ו-CAD ניתוח נוספים.

ניתוח והמרת נתונים: אנו הופכים מידע מפוזר לנתונים מובנים, ניתוחים, מסקנות והחלטות.

אינטגרציית נתונים ואוטומציה של תהליכים: מהפקת מסמכים אוטומטית ועד אינטגרציה עם מערכות פנימיות ומסדי נתונים חיצוניים.

כדי לגלות כיצד השימוש באוטומציה יכול לעזור לחברה שלכם DataDrivenConstruction.io צרו קשר עם להשיג תוצאות עסקיות משמעותיות.

מילון מונחים

(בינה מלאכותית -) היכולת של מערכות מחשב לבצע משימות שדורשות בדרך כלל אינטליגנציה אנושית, כמו זיהוי AI דפוסים, למידה וקבלת החלטות.

פלטפורמת קוד פתוח לאורקסטריציה של תהליכי עבודה, המאפשרת ליצור, לתכנן ולעקוב אחר - Apache Airflow (cyclic directed גרפים א) DAG באמצעות ETL-תהליכי עבודה ו

כלי לאוטומציה של זרימת נתונים בין מערכות, המתמחה בהכוונה והמרת נתונים - Apache NiFi

פורמט קובץ יעיל לאחסון נתונים בעמודות, מותאם לשימוש במערכות ניתוח נתונים גדולים - Apache Parquet מספק דחיסה משמעותית ועיבוד מהיר

(ממשק תכנות יישומים -) ממשק פורמלי המאפשר לתוכנית אחת לתקשר עם אחרת מבלי לגשת לקוד המקור API exchanging data and functionality through standardized requests and responses.

מאפיין - תכונה או איכות של אובייקט, המתארת את המאפיינים שלו (כגון שטח, נפח, עלות, חומר)

מסדי נתונים - מבנים מאורגנים לאחסון, ניהול וגישה למידע, המשמשים לחיפוש ועיבוד נתונים בצורה יעילה

תוכנית ליישום מודל מידע ביינים, המגדירה מטרות, שיטות ותהליכים ליישום - (BEP) תוכנית ביצוע מודל מידע בניין בפרויקט BIM

מערכות מידע בהיקף משמעותי, מגוון ומהירות עדכון, אשר דורשות טכנולוגיות מיוחדות - (Big Data) נתונים גדולים לעיבוד וניתוח

ניתוח עסקי; תהליכים, טכנולוגיות וכלים להמרת נתונים למידע משמעותי לקבלת החלטות - (BI) אינטליגנציה עסקית

תהליך יצירה וניהול של ייצוגים דיגיטליים של תכונות פיזיות ופונקציונליות של אובייקטים - (BIM) מודל מידע בניין בבנייה, הכולל לא רק מודלים תלת-ממדיים אלא גם מידע על תכונות, חומרים, לוחות זמנים ועלויות

קופסה שחורה/קופסה לבנה - גישות להבנת מערכת: במקרה הראשון הלוגיקה הפנימית מוסתרת, ורק הקלטות והפלטות נראים; במקרה השני - תהליך העיבוד שקוף ונגיש לניתוח

מבנה גיאומטרי המתאר את גבולות האובייקט במרחב תלת-ממדי דרך - (Bounding Box) קופסה מגבילה ויוצר "קופסה" סביב האובייקט, X, Y ו-Z קואורדינטות מינימליות ומקסימליות על צירי

ייצוג גיאומטרי של אובייקטים, המגדיר אותם דרך גבולות פני השטח - (BREP) ייצוג גבול

מערכת תכנון ממוחשבת, המשמשת ליצירה, עריכה וניתוח של שרטוטים מדויקים - (CAD) תכנון בעזרת מחשב ומודלים תלת-ממדיים בארכיטקטורה, בנייה, הנדסה ועוד תחומים

תוכנה לניהול נכסי נדל"ן ותשתיות, הכוללת תכנון חלל, ניהול נכסים, תחזוקה - (CAFM) ניהול מתקנים בעזרת מחשב ומעקב עלויות

מרחב דיגיטלי מרכזי לניהול, אחסון, חילופי מידע ועבודה משותפת עם מידע פרויקט - (CDE) סביבה נתונים משותפת בכל שלבי מחזור חיי האובייקט

מבנה מיוחד בארגון, האחראי על פיתוח תחום ידע מסוים, פיתוח - (Center of Excellence, CoE) מרכז מצוינות סטנדרטים ושיטות עבודה מיטביות, הכשרת עובדים ותמיכה ביישום חדשנות

מערכת סיווג מודרנית של רכיבי בניין מהדור השלישי - CoClass

מודל נתונים קונספטואלי - ייצוג ברמה גבוהה של ישויות עיקריות וקשריהן ללא פירוט של תכונות, המשמש בשלב ההתחלתי של תכנון בסיסי נתונים

מערכת לניהול אינטראקציה עם לקוחות, המשמשת לאוטומציה של תהליכי מכירה – (CRM) ניהול קשרי לקוחות ושירות.

(Airflow, NiFi כגון) גרף מכוון שאינו מעגלי, המשמש במערכות תיאום נתונים – (DAG) מכוון cyclic גרף א לקביעת רצף ותלות בין משימות.

מסגרת פייתון ליצירת ויזואליזציות נתונים אינטראקטיביות באינטרנט – Dash

פאנל מידע, המייצג באופן חזותי את המדדים המרכזיים של ביצועים ומדדים בזמן אמת – (Dashboard) לוח מחוונים מתודולוגיה הממקדת את תשומת הלב בנתונים ולא ביישומים או בקוד תוכנה – (Data-Centric) גישה מונחת נתונים מה שהופך את הנתונים לנכס המרכזי של הארגון.

מכלול של פרקטיקות, תהליכים ומדיניות, המבטיחים שימוש נאות ויעיל – (Data Governance) ניהול נתונים בנתונים בארגון, כולל בקרת גישה, איכות וביטחון.

אגם נתונים – מאגר המיועד לאחסון כמויות גדולות של נתונים גולמיים בפורמט המקורי שלהם עד לרגע השימוש בהם עם (Data Lake) אגם בית נתונים – גישה ארכיטקטונית המשלבת את הגמישות והסקאלביליות של אגמי נתונים (DWH). הניהוליות והביצועים של מחסני נתונים.

בנייה מונעת נתונים – גישה אסטרטגית שבה כל שלב במעגל חיי הפרויקט – מעיצוב ועד תפעול – נתמך על ידי מערכות אוטומטיות ומקושרות. גישה זו מבטיחה למידה מתמשכת על בסיס עובדות, מפחיתה אי-ודאות ומאפשרת לחברות להשיג יתרון תחרותי בר קיימא בענף.

אינטגרטור מונע נתונים – חברה המתמחה באיחוד נתונים ממקורות שונים וניתוחם לצורך קבלת החלטות ניהוליות. גישה מונעת נתונים (גישה מרכזית לנתונים) – מתודולוגיה שבה נתונים נחשבים לנכס אסטרטגי, וההחלטות מתקבלות על בסיס ניתוח אובייקטיבי של מידע ולא על בסיס דעות סובייקטיביות.

מינימליזם נתוני – גישה לצמצום הנתונים לנתונים החשובים והמשמעותיים ביותר, המאפשרת לפשט את עיבוד וניתוח המידע.

ביצת נתונים – אוסף מפוזר של נתונים לא מובנים, הנוצר כתוצאה מאיסוף ואחסון מידע ללא שליטה וללא ארגון נאות. נתונים וניתוח, הממוקדת בשיפור שיתוף הפעולה, DevOps, מתודולוגיה המשלבת את עקרונות DataOps – האינטגרציה ואוטומציה של זרמי נתונים.

דיגיטציה של מידע – תהליך המרה של כל היבטי הפעילות הבנייתית לצורה דיגיטלית, המתאימה לניתוח, פרשנות ואוטומציה.

שבו השורות מייצגות רשומות או (Pandas) מסגרת נתונים – מבנה נתונים דו-ממדי בספריית DataFrame אובייקטים נפרדים, והעמודות מייצגות את המאפיינים או התכונות שלהם.

ניתוח תיאורי – ניתוח נתונים היסטוריים להבנת מה קרה בעבר.

ניתוח אבחוני – ניתוח נתונים לקביעת הסיבות לכך שמהו קרה.

תרשים גנט – כלי לתכנון פרויקטים, המייצג משימות בצורה של פסים אופקיים על ציר זמן, ומאפשר להציג בצורה ברורה את הרצף והמשך של העבודות.

(מחסן נתונים) – מערכת מרכזית לאחסון נתונים, המאגדת מידע ממקורות רבים, מארגנת אותו ומאפשרת גישה DWH לניתוח ודיווח.

(סביבתי, חברתי, ממשלתי) – סט קריטריונים להערכת ההשפעה הסביבתית, החברתית והממשלתית של חברה ESG או פרויקט.

(חילוף, טעינה, המרה) – תהליך שבו נתונים קודם כל נחלצים ממקורות ונטענים למחסן, ולאחר מכן מומרת ELT לצורכי ניתוח.

(חילוף, המרה, טעינה) – תהליך של חילוף נתונים ממקורות שונים, המרתם לפורמט הנדרש וטעינתם למחסן היעד ETL לצורך ניתוח.

(ישות-קשר) – תרשים חזותי הממחיש ישויות, את המאפיינים שלהן ואת הקשרים ביניהן, המשמש במידול ER דיאגרמת נתונים.

(תכנון משאבי ארגון) – מערכת מודולרית מקיפה לתכנון משאבי הארגון, המשמשת לניהול ואופטימיזציה של ERP היבטים שונים בתהליך הבנייה.

תכונות – בלמידת מכונה, משתנים בלתי תלויים או תכונות, המשמשות כקלט למודל.

מודל פיזי של נתונים – ייצוג מפורט של מבנה בסיס הנתונים, כולל טבלאות, עמודות, סוגי נתונים, מפתחות ואינדקסים. מותאם למערכת ניהול בסיסי נתונים ספציפית.

PDF ספריית פייתון ליצירת מסמכי – FPDF

ליבת גיאומטריה – רכיב תוכנה המספק אלגוריתמים בסיסיים ליצירה, עריכה וניתוח של אובייקטים גיאומטריים ב-CAD, BIM ויישומים הנדסיים אחרים.

(הדעת של האדם המרוויח ביותר) – גישה לקבלת החלטות המבוססת על דעתו של האדם המרוויח ביותר HiPPO בארגון, ולא על נתונים אובייקטיביים.

(PyCharm, VS Code, Jupyter Notebook), סביבת פיתוח משולבת – (כלי כולל לכתיבה, בדיקה ודיבוג של קוד) למשל IDE (Notebook).

(מפרט העברת מידע) – מפרט העברת מידע, המגדיר את הדרישות לנתונים בשלבים שונים של הפרויקט IDS.

(המבטיח תאימות בין פתרונות תוכנה שונים, BIM) מחלקות יסוד בתעשייה – (פורמט להחלפת נתוני IFC

תעשייה 5.0 – מושג לפיתוח התעשייה, המאחד את יכולות הדיגיטציה, האוטומציה והבינה המלאכותית עם הפוטנציאל האנושי והקיימות הסביבתית.

אינטגרציית נתונים – תהליך של שילוב נתונים ממקורות שונים למערכת אחת, שלמה, כדי להבטיח ייצוג אחיד של המידע.

סילו מידע – מערכות אחסון נתונים מבודדות, שאינן משתפות מידע עם מערכות אחרות, מה שיוצר מחסומים לשימוש יעיל בנתונים.

(האינטרנט של הדברים) – מושג לחיבור אובייקטים פיזיים לאינטרנט לצורך איסוף, עיבוד והעברת נתונים IoT.

אלגוריתם למידת מכונה, המדרג אובייקטים על סמך דמיון עם השכנים הקרובים – (שכנים הקרובים ביותר-k) k-NN (ביותר במדגם האימון).

פלטפורמה לניתוח נתונים ותחרויות בלמידת מכונה – Kaggle

חישוב – חישוב עלות עבודות בנייה או תהליכים לפי יחידת מידה מסוימת) למשל, עבור 1 מ² של קיר גבס, 1 מ³ של (בטון).

(מדדי ביצוע מרכזיים) – מדדים כמותיים הניתנים למדידה, המשמשים להערכת הצלחת פעילות החברה או פרויקט KPI ספציפי.

תוויות – בלמידת מכונה, משתנים או תכונות יעד, שהמודל צריך לחזות.

אלגוריתם למידה – תהליך חיפוש ההשערה הטובה ביותר במודל, העונה על הפונקציה היעד, תוך שימוש בקבוצת נתוני אימון.

רגרסיה ליניארית – שיטה סטטיסטית למידול הקשר בין משתנה תלוי ואחד או יותר משתנים בלתי תלויים.
(מודל שפה גדול – מודל שפה גדול, בינה מלאכותית, שהוכשרה להבין וליצור טקסט על בסיס כמויות עצומות LLM של נתונים, המסוגלת לנתח הקשר ולכתוב קוד תוכנה.

(רמת פירוט/פיתוח – רמת הפירוט של המודל, המגדירה את מידת הדיוק הגיאומטרי ואת התוכן המידע LOD מודל לוגי של נתונים - תיאור מפורט של ישויות, תכונות, מפתחות וקשרים, המייצג את המידע העסקי והכלליים, שלב ביניים בין המודל הקונספטואלי למודל הפיזי.

למידת מכונה - קבוצה של שיטות אינטליגנציה מלאכותית, המאפשרות למערכות מחשב ללמוד ולבצע תחזיות על בסיס נתונים ללא תכנות מפורש.

מאסטרפורמט - מערכת סיווג מהדור הראשון, המשמשת לארגון מפרטי בנייה לפי סעיפים ודיסציפלינות.
(מכני, חשמלי, אינסטלציה -) מערכות הנדסה של מבנים, הכוללות רכיבים מכניים, חשמליים ואינסטלציה MEP רשת - ייצוג רשתי של אובייקטים תלת-ממדיים, המורכב מקודקודים, קווים ופנים.

מודל - בלמידת מכונה, אוסף של היפותזות שונות, אחת מהן מתקרבת לפונקציה היעד שיש לחזות או לאפר מודל נתונים - תהליך יצירת ייצוג מסודר של נתונים וקשרים ביניהם לצורך יישום במערכות מידע, הכולל רמות קונספטואליות, לוגיות ופיזיות.

כלי קוד פתוח לאוטומציה של תהליכי עבודה ואינטגרציה של יישומים באמצעות גישה עם קוד נמוך - n8n.
נורמליזציה - בלמידת מכונה, תהליך הבאת נתונים מספריים שונים לאותו סקאלה כדי להקל על עיבוד וניתוחם.
הנדסה הפוכה - תהליך חקר המבנה, הפונקציה והטכנולוגיה של אובייקט על ידי ניתוח המבנה, הפונקציות והעבודה שלו.
בהקשר של נתונים - חילוץ מידע מפורמטים קנייניים לשימוש במערכות פתוחות.

(זיהוי תווים אופטי -) טכנולוגיה לזיהוי תווים אופטי, המאפשרת להמיר תמונות טקסט (מסמכים סרוקים, תמונות OCR) לפורמט טקסט קריא על ידי מכונה.

תקן סיווג בינלאומי מהדור השני לניהול מידע על אובייקטי בנייה - OmniClass.
אוטולוגיה - מערכת של קשרים בין מושגים, המנסחת תחום ידע מסוים.
קוד פתוח - מודל לפיתוח והפצת תוכנה עם קוד פתוח, הזמין לשימוש חופשי, חקר ושינוי.

פתוח, המניח שימוש בתקנים ובפורמטים פתוחים להחלפת נתונים בין פתרונות תוכנה שונים BIM מושג - Open BIM.
תקנים פתוחים - מפרטים זמינים לציבור להשגת משימה מסוימת, המאפשרים למערכות שונות לתקשר ולהחליף נתונים.
לעבודה DataFrame ו-Series ספריית פייתון עם קוד פתוח לעיבוד וניתוח נתונים, המספקת מבני נתונים - Pandas יעילה עם מידע טבלאי.

פרדיגמת נתונים פתוחים - גישה לעיבוד נתונים, שבה המידע זמין לשימוש חופשי, שימוש חוזר והפצה לכל אדם.
שיטה פרמטרית - שיטת הערכת פרויקטי בנייה, המשתמשת במודלים סטטיסטיים להערכת עלויות על בסיס פרמטרים של הפרויקט.

(מודל מידע פרויקט -) מערכת דיגיטלית, המיועדת לארגון, אחסון והחלפת כל המידע הפרויקטלי PIMS.
צינור – רצף של תהליכי עיבוד נתונים, החל מהפקה והמרה ועד לניתוח והצגה.

מערכת לניהול פרויקטים, המיועדת לשליטה מפורטת על ביצוע משימות – (PMIS) מערכת ניהול מידע פרויקטים ברמת פרויקט בנייה ספציפי

תחום אנליטיקה המשתמש בשיטות סטטיסטיות ולמידת מכונה כדי – (Predictive Analytics) אנליטיקה חיזוית לחזות תוצאות עתידיות על בסיס נתונים היסטוריים

תחום אנליטיקה שאינו רק חוזה תוצאות עתידיות, אלא גם – (Prescriptive Analytics) אנליטיקה פרסקריפטיבית מציע פעולות אופטימליות להשגת תוצאות רצויות

פורמטים קנייניים – פורמטים סגורים של נתונים, הנשלטים על ידי חברה מסוימת, המגבילים את אפשרויות החלפת המידע ומגדילים את התלות בתוכנה ספציפית

תהליך של הפקת מאפיינים כמותיים של רכיבים מתוך מסמכי תכנון לצורך חישוב – (QTO) תהליך חישוב כמויות כמויות החומרים הנדרשים לביצוע הפרויקט

מערכת ניהול איכות – מערכת לניהול איכות, המבטיחה עמידה בדרישות שנקבעו עבור תהליכים ותוצאות

(הפקת מידע מוגברת) – שיטה המשלבת את יכולות ההפקה של מודלים לשוניים עם הפקת מידע רלוונטי ממסדי RAG נתונים ארגוניים, מה שמגביר את הדיוק והעדכניות של התשובות

מערכת לניהול מסדי נתונים רלציוניים, המארגנת מידע בצורה של – (RDBMS) מערכת ניהול מסדי נתונים רלציוניים טבלאות מקושרות

שפה פורמלית לחיפוש ועיבוד מחרוזות, המאפשרת לקבוע תבניות לבדוק נתוני טקסט – (RegEx) ביטויים רגולריים בהתאם לקריטריונים מסוימים

רגרסיה – שיטה סטטיסטית לניתוח תלות בין משתנים

שיטה להערכת פליטות פחמן דו-חמצני הקשורות לייצור ושימוש בחומרי בניין ובתהליכים – CO₂ חישובי

שיטת משאבים – שיטה להכנת אומדנים, המבוססת על ניתוח מפורט של כל המשאבים הנדרשים (חומרים, עבודה, ציוד) לביצוע עבודות בנייה

טכנולוגיה לזיהוי אוטומטי של אובייקטים באמצעות אותות רדיו, המשמשת למעקב – (RFID) זיהוי באמצעות תדר רדיו אחר חומרים, ציוד ואנשים

מדד המייצג את היחס בין רווחים להוצאות, המשמש להערכת יעילות ההשקעות – (ROI) החזר על השקעה

מודל לספקת תוכנה כשירות, שבו היישומים מאוחסנים על ידי ספק ונגישים למשתמשים – (SaaS) תוכנה כשירות דרך האינטרנט

ניהול שרשרת אספקה, הכולל תיאום ואופטימיזציה של כל התהליכים החל מרכישת – (SCM) ניהול שרשרת אספקה חומרים ועד אספקת מוצרים מוגמרים

סילו נתונים – מחסני מידע מבודדים בארגון, שאינם משולבים עם מערכות אחרות, מה שמקשה על החלפת נתונים ומפחית את היעילות

שפה לשאילתות מובנות, המשמשת לעבודה עם מסדי נתונים רלציוניים – (SQL) שפת שאילתות מובנית

מערכת ניהול מסדי נתונים קלה, משולבת, חוצת פלטפורמות, שאינה דורשת שרת נפרד ותומכת בפונקציות – SQLite בשימוש נרחב באפליקציות ניידות ובמערכות משולבות, SQL בסיסיות של

נתונים מובנים – מידע מאורגן בפורמט מסוים עם מבנה ברור, לדוגמה, במסדי נתונים רלציוניים או בטבלאות

שבו אלמנטים שונים עשויים להכיל, XML או JSON נתונים חלשים-מובנים - מידע עם ארגון חלקי ומבנה גמיש, כמו קבוצות שונות של תכונות

היא אובייקט קונקרטי או מופשט מהעולם האמיתי, שניתן להבחין בו, לתאר אותו ולהציגו (entity: באנגלית) ישות כנתונים.

למידה מפקחת - סוג של למידת מכונה שבו האלגוריתם מתאמן על נתונים מסומנים, שבהם לכל דוגמה ידוע התוצאה הרצויה.

טקסונומיה - מערכת סיווג היררכית, המשמשת לחלוקה שיטתית של אלמנטים לקטגוריות על בסיס תכונות משותפות.

מעריך הנתונים של טיטניק - מעריך נתונים פופולרי ללמידה ובחינת מודלים של למידת מכונה.

אימון - תהליך שבו אלגוריתם למידת מכונה מנתח נתונים כדי לזהות דפוסים וליצור מודל.

למידת העברה - שיטה של למידת מכונה שבה מודל, שהוכשר למשימה אחת, משמש כנקודת מוצא למשימה אחרת.

המרת נתונים - תהליך שינוי הפורמט, המבנה או התוכן של נתונים לשימוש עתידי.

דרישות נתונים - קריטריונים פורמליים המגדירים את המבנה, הפורמט, השלמות והאיכות של המידע הנדרש לתמיכה בתהליכי עסקים.

אובריזציה של תעשיית הבנייה - תהליך של טרנספורמציה של מודלים עסקיים מסורתיים בבנייה בהשפעת פלטפורמות דיגיטליות, המאפשרות אינטראקציה ישירה בין לקוחות לביצועיסטים ללא מתווכים.

יוניקלאס - מערכת סיווג של רכיבי בנייה מהדור השני והשלישי, בשימוש נרחב בבריטניה.

(תיאור סצנה אוניברסלי -) פורמט נתונים שפותח עבור גרפיקה ממוחשבת, אך נמצא בשימוש במערכות הנדסה USD בזכות המבנה הפשוט שלו ועצמאותו מגיאומטריות שונות.

אימות נתונים - תהליך של בדיקת מידע על מנת לוודא שהוא עומד בקריטריונים ובדרישות שנקבעו, המבטיח דיוק, שלמות ועקביות של הנתונים.

מסד נתונים וקטורי - סוג מיוחד של מסד נתונים המאחסן נתונים בצורה של וקטורים רב-ממדיים לחיפוש סמנטי יעיל והשוואת אובייקטים.

ייצוג וקטורי (הטמעה -) שיטה להמרת נתונים לווקטורים מספריים רב-ממדיים, המאפשרת לאלגוריתמים של מכונה לעבד ולנתח מידע ביעילות.

מתודולוגיה המתמקדת בעיבוד, אחסון וניתוח נתונים וקטוריים רב-ממדיים, במיוחד בתחומים כמו - VectorOps דיגיטליים כפולים וחיפוש סמנטי.

ויזואליזציה - ייצוג גרפי של נתונים לצורך תפיסה וניתוח מידע בצורה יעילה יותר.

חלוקת המונחים באלפבית בוצעה לפי שמותיהם באנגלית.

רשימת ספרות וחומרים מקוונים

- [1] אינדקס מסמכים ומדדים שפורסמו "12 דצמבר 2016]. באינטרנט: IT 2017 גארטנר", נתוני מדדי מפתח [תאריך גישה: 1 מרץ 2025]. <https://www.gartner.com/en/documents/3530919>. [זמין]
- [2] אתגרים מוכרים - גישות חדשות. סקר הבנייה הגלובלי 2023, "1 ינואר 2023]. באינטרנט. KPMG, [זמין], <https://assets.kpmg.com/content/dam/kpmgsites/xx/pdf/2023/06/familiar-challenges-new-solutions-1.pdf>. [תאריך גישה: 5 מרץ 2025].
- [3] פ. ר. ברנרד, תמונה שווה אלף מילים, "10 במאי 1927]. באינטרנט. [זמין] https://en.wikipedia.org/wiki/A_picture_is_worth_a_thousand_words. [תאריך גישה: 15 במרץ 2025].
- [4] [4]
- [5] "היא הונאת מדדים חסרת טעם AGI מ. באסטיאן", מנכ"ל מיקרוסופט סאטיה נדלה אומר שהצהרה עצמית על <https://the-decoder.com/microsoft-ceo-satya-nadella-says-self-claiming-agi-is-nonsensical-benchmark-hacking/>. [בפברואר 2025]. באינטרנט. [זמין 21 תאריך גישה: 15 במרץ 2025].
- [5] ו. א. פורום, פורום מעצב את עתיד הבנייה – נוף במעבר, "1 בינואר 2016]. באינטרנט. [זמין] https://www3.weforum.org/docs/WEF_Shaping_the_Future_of_Construction.pdf. [תאריך גישה: 2 במרץ 2025].
- [6] ס. ד. גילספי, חימר: הסתבכות כדור הארץ בעידן החימר, "2024]. באינטרנט. [זמין] <https://ufl.pb.unizin.org/imos/chapter/clay/>.
- [7] פפירוס מהמאה ה-3 לפני הספירה. שפה – יוונית, "2024]. באינטרנט. [זמין] <https://www.facebook.com/429710190886668/posts/595698270954525>.
- [8] ניטור: שימוש בכלים הזמינים, "1980]. באינטרנט. [זמין] <https://pubmed.ncbi.nlm.nih.gov/10246720/>. [תאריך גישה: 15 במרץ 2025].
- [9] מנוהל על ידי נתונים מה הסטודנטים צריכים להצליח בעולם העסקים המשתנה במהירות, "15", PWC, <https://www.pwc.com/us/en/faculty-resource/assets/PwC-Data-driven-paper-Feb2015.pdf>. [בפברואר 2015]. באינטרנט. [זמין] [תאריך גישה: 15 במרץ 2025].
- [10] סקנסקה ארה"ב", מגמות שוק הבנייה בסתיו "2, בנובמבר 2023]. באינטרנט. [זמין]

<https://x.com/SkanskaUSA/status/1720167220817588714>.

- [11] ציטוטים חיוניים מאוקספורד (מהדורה 4, "הוצאת אוקספורד, 2016). באינטרנט. [זמין]
<https://www.oxfordreference.com/display/10.1093/acref/9780191826719.001.0001/q-oro-ed4-00006236>. [תאריך גישה: 1 במרץ 2025].
- [12] ציטוט: סונדרגארד על ניתוח נתונים, "באינטרנט. [זמין]
<https://www.causeweb.org/cause/resources/library/r2493>. [תאריך גישה: 15 במרץ 2025].
- [13] <https://iot-analytics.com/how-global-ai-interest-is-boosting-data-management-market/>. [תאריך גישה: 15 במרץ 2025]
 "מחזק את שוק ניהול הנתונים", 28 במאי 2024. באינטרנט. [זמין AI-איך העניין הגלובלי ב" [תאריך גישה: 15 במרץ 2025].
- [14] [באינטרנט. [זמין. "ERP, 2024. י. מקיאו", היסטוריה של
<https://www.netsuite.com/portal/resource/articles/erp/erp-history.shtml>.
- [15] <https://erp scout.de/en/erp-costs/>. [באינטרנט. [זמין, "ERP? כמה עולה מערכת ERP: מחיר", erpscout, [תאריך גישה: 15 במרץ 2025].
- [16] 18, "ERP 2022 דו"ח) ERP מלמדים אותנו על בחירת ERP מה 1,384 פרויקטי", softwarepath, [תאריך גישה: 15 במרץ 2025]
<https://softwarepath.com/guides/erp-report>. [באינטרנט. [זמין, [תאריך גישה: 15 במרץ 2025].
- [17] דלויט, "ניהול מונחה נתונים בפרויקטים דיגיטליים של הון", 16 בדצמבר 2016. באינטרנט. [זמין]
<https://www2.deloitte.com/content/dam/Deloitte/us/Documents/Real%20Estate/us-engineering-construction-data-driven-management-digital-capital-projects.pdf>. [תאריך גישה: 1 במרץ 2025].
- [18] [תאריך גישה: 15 מרץ 2025]
https://en.wikipedia.org/wiki/Moore%27s_law. [זמין ויקיפדיה, חוק מור], "באינטרנט. [זמין]
- [20] אקסנטור, בניית ערך נוסף עם פרויקטים של הון, 1 ינואר 2020. באינטרנט. [זמין]
<https://www.accenture.com/content/dam/accenture/final/a-com-migration/r3-3/pdf/pdf-143/accenture-industryx-building-value-capital-projects-highres.pdf>. [תאריך גישה: 3 מרץ 2024].
- [21] ב. מאר, כמה נתונים אנו מייצרים בכל יום? הסטטיסטיקות המדהימות שכולם צריכים לקרוא, 2018.
<https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-> [באינטרנט. [זמין]

data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read.

- [22] כמה נתונים מיוצרים בכל יום, [2024?]. באינטרנט. [זמין]
<https://graduate.northeastern.edu/resources/how-much-data-produced-every-day/>.
- [23] ת. סאליבן, אי. אי. ו. "הדאטספירה" הגלובלית: כמה מידע יהיה לאנושות עד לשנת 2025, [2024?]. באינטרנט. [זמין]
<https://www.datauniverseevent.com/en-us/blog/general/AI-and-the-Global-Datasphere-How-Much-Information-Will-Humanity-Have-By-2025.html>.
- [24] סטטיסטיקה, מספר כולל של ספרים מודפסים שהופקו באזורים שונים של מערב אירופה בכל חצי מאה בין השנים 1800-ל-2025, [זמין 1454].
<https://www.statista.com/statistics/1396121/europe-book-production-half-century-region-historical/>. [תאריך גישה: 1 מרץ 2025].
- [25] https://cloud.google.com/storage/pricing-examples. דוגמאות לקביעת מחירים, [2024]. באינטרנט. [זמין]
- [26] מ. אשאר, עסקים מעבירים את אחסון הנתונים החוצה ככל שהמורכבות עולה, 10 במאי 2024. באינטרנט. [זמין]
<https://www.ciodive.com/news/enterprises-outsource-data-storage-complexity-rises/715854/>. [תאריך גישה: 15 במרץ 2025].
- [27] מת? ניתוח התחזית המפתיעה של מנכ"ל מיקרוסופט, "13 ינואר 2025 SaaS האם", JETSOFTPRO, [זמין]
<https://jetsoftpro.com/blog/saas-is-dead/>. [באינטרנט. [זמין]
- [28] עם: ביל גורלי וברד גרסטנר, 12 בדצמבר 2024. באינטרנט. [זמין BG2 | פוד, סאטיה נאדלה BG2]
https://www.youtube.com/watch?v=9NtsnzRFJ_o. [תאריך גישה: 15 במרץ 2025].
- [29] https://www.goodreads.com/quotes/8644920-data-is-a-precious-thing-and-will-last-longer-than. גודרידס, טים ברנרס-לי, באינטרנט. [זמין]
 [תאריך גישה: 15 מרץ 2025].
- [30] קיו בנייה 4.0: זמן להכרעה, 1 ינואר 2023. באינטרנט. [זמין] KPMG,
<https://kpmg.com/ca/en/home/insights/2023/05/cue-construction-make-or-break-time.html>. [תאריך גישה: 5 מרץ 2025].
- [31] א. דיינינגר, ב. קוך, ר. באוקנהכט ומ. לאנגהנס, שימוש במודלים דיגיטליים לצורך דקרבוניזציה של אתר ייצור: דוגמה לחיבור בין מודל הבניין, מודל הייצור ומודל האנרגיה, [2024]. באינטרנט. [זמין]
https://www.researchgate.net/publication/374023998_Using_Digital_Models_to_Decarbonize_a_Production_Site_A_Case_Study_of_Connecting_the_Building_Model_Production_Model_and_Energy_Model.
- [32] מכנסי", שיקום הבנייה: דרך לעלייה בפרודוקטיביות, "1 בפברואר 2017. באינטרנט. [זמין]

<https://www.mckinsey.com/~media/mckinsey/business%20functions/operations/our%20insights/reinventing%20construction%20through%20a%20productivity%20revolution/mgi-reinventing-construction-a-route-to-higher-productivity-full-report.pdf>.

- [33] כוח המשימה לבניית תשתיות לראש ממשלת המשנה, מחשבה מחדש על הבנייה, 1 באוקטובר 2014. https://constructingexcellence.org.uk/wp-content/uploads/2014/10/rethinking_construction_report.pdf. [באינטרנט. זמין]
- [34] פורבס, בלי דעה, אתה רק אדם נוסף עם נתונים, 15 מרץ 2016. [באינטרנט. זמין] <https://www.forbes.com/sites/silberzahnjones/2016/03/15/without-an-opinion-youre-just-another-person-with-data/>. [תאריך גישה: 15 מרץ 2025].
- [35] ויקיטקסט, צ'ארלס בבג, https://en.wikiquote.org/wiki/Charles_Babbage. [באינטרנט. זמין] [תאריך גישה: 15 מרץ 2025].
- [36] מחקר חדש מגלה שכמעט חצי מהמנהלים סומכים על בינה מלאכותית יותר מאשר על עצמם, 12 במרץ, SAP, <https://news.sap.com/2025/03/new-research-executive-trust-ai/>. [באינטרנט. זמין] [תאריך גישה: 15 במרץ 2025].
- [37] בקנדה, 2021, בנייה בעידן הדיגיטלי, 1 במאי 2021. [באינטרנט KPMG-האסוציאציה הקנדית לבנייה ו- <https://assets.kpmg.com/content/dam/kpmg/ca/pdf/2021/05/construction-in-the-digital-age-report-en.pdf>. [באינטרנט. זמין] [תאריך גישה: 5 במרץ 2025].
- [38] <https://www.pwc.in/decoding-the-fifth-industrial-revolution.html>. [באינטרנט. זמין] [תאריך גישה: 15 מרץ 2025].
- [39] מ. ק. זכויות פרטיות ובעיות ציבוריות: הכלכלה הגלובלית של, מכון פיטרסון לכלכלה בינלאומית, 2012.
- [40] F. N. a. Y. Z. 2024 ינואר 24, הפתוח, ערך התוכנה הקוד הפתוח, מנואל הופמן, בית הספר לעסקים של הרווארד: <https://www.hbs.edu/faculty/Pages/item.aspx?num=65230>. [באינטרנט. זמין] [תאריך גישה: 15 מרץ 2025].
- [41] מרכז ניתוח עלויות ימי סדנאות ניתוח עלויות של חיל האוויר, מדריך להערכת עלויות פיתוח תוכנה, 1 בספטמבר 2008. [באינטרנט. זמין] <https://www.dau.edu/sites/default/files/Migrated/CopDocuments/SW%20Cost%20Est%20Manual%20Vol%20I%20rev%2010.pdf>.
- [42] מכינסיי, שיפור פרודוקטיביות הבנייה, [באינטרנט. זמין] <https://www.mckinsey.com/capabilities/operations/our-insights/improving->

- construction-productivity. [תאריך גישה: 15 מרץ 2025].
- [43] א. ג. א. ס. סייברסון, הדרך המוזרה והנוראה של הפרודוקטיביות במגזר הבנייה בארצות הברית, 19 ינואר 2023. [תאריך גישה: 15 מרץ 2025]. <https://bfi.uchicago.edu/insight/research-summary/the-strange-and-awful-path-of-productivity-in-the-us-construction-sector/>.
- [44] מכנסים, "המסירה על פרודוקטיביות הבנייה אינה אופציה יותר", 9 באוגוסט 2024. [תאריך גישה: 5 במרץ 2025]. <https://www.mckinsey.com/capabilities/operations/our-insights/delivering-on-construction-productivity-is-no-longer-optional>.
- [45] הפריון הנמוך בבנייה מעלה את עלויות הבנייה, "12 בדצמבר 2022. [תאריך גישה: 15 במרץ 2025]. <https://think.ing.com/articles/lagging-productivity-drives-up-building-costs-in-many-eu-countries/>.
- [46] מ. ברמן, מנכ"ל מיקרוסופט, תחזית מפתיעה: סוכנים יחליפו את כל התוכנה, "19 בדצמבר 2024. [תאריך גישה: 15 במרץ 2025]. <https://www.youtube.com/watch?v=uGOLYz2pgr8>.
- [47] ביזנס אינסיידר, מנהל אנתרופיק אומר כי בתוך 3 עד 6 חודשים, הבינה המלאכותית תכתוב 90% מהקוד שהיה באחריות המפתחים, 15 מרץ 2025. [תאריך גישה: 30 מרץ 2025]. <https://www.businessinsider.com/anthropic-ceo-ai-90-percent-code-3-to-6-months-2025-3>.
- [48] ברחבי העולם נכון ליוני 2024, לפי (DBMS) סטטיסטיקה, השוואת פופולריות של מערכות ניהול מסדי נתונים [קטגוריה, יוני 2024]. [תאריך גישה: 15 מרץ 2025]. <https://www.statista.com/statistics/1131595/worldwide-popularity-database-management-systems-category/>.
- [49] DB-Engines, דירוג DB-Engines, [תאריך גישה: 15 מרץ 2025]. <https://db-engines.com/en/ranking>.
- [50] [תאריך גישה: 15 מרץ 2025]. <https://survey.stackoverflow.co/2023/>.
- [51] SQL, 2024. [תאריך גישה: 15 מרץ 2025]. <https://en.wikipedia.org/wiki/SQL>.
- [52] נתונים מובנים ולא מובנים: מה ההבדל, 2024?. [תאריך גישה: 15 מרץ 2025]. <https://www.ibm.com/blog/structured-vs-unstructured-data/>.
- [53] DataDrivenConstruction, 2024 אפריל 23, השוואת פורמטים של נתונים עבור פרויקטים בבנייה.

- [זמין]: <https://datadrivenconstruction.io/wp-content/uploads/2024/10/COMPARISON-OF-DATA-FORMATS-FOR-CONSTRUCTION-PROJECTS-1.pdf>.
- [54] מודל מידע בנייה, מסמך מידע, [2003]. באינטרנט. [זמין]
<https://web.archive.org/web/20030711125527/http://usa.autodesk.com/adsk/servlet/item?id=2255342&siteID=123112>.
- [55] חלק 5: בלאקרוק – מאסטר כל הטכנולוגיות. כיצד תאגידים BIM. א. בויוק, מלחמות לוביסטים ופיתוח
 [זמין]. באינטרנט. [2024, פתוח בקוד פתוח]: <https://bigdataconstruction.com/autodesk-oracle-blackrock-open-source/>.
- [56] ד. אושקוב, מודל ישיר - מי ולמה צריך את זה? סקירה של טכנולוגיות מתחרות, 14 נובמבר 2011
 תאריך גישה: 02: [זמין]. באינטרנט. [2025].
https://isicad.net/articles.php?article_num=14805.
- [57] צ. איסטמן ואחרים, איסטמן, צ'אק; ואחרים, ספטמבר 1974]. באינטרנט. [זמין]
 [תאריך גישה: 15 מרץ 2025].
<https://files.eric.ed.gov/fulltext/ED113833.pdf>.
- [58] ד. אושקוב, מודל ישיר - מי ולמה צריך את זה? סקירה של טכנולוגיות מתחרות, 11 נובמבר 2011
 תאריך גישה: 15 מרץ]. באינטרנט. [2025].
https://isicad.net/articles.php?article_num=14805.
- [59] דצמבר 2022]. באינטרנט. [זמין 12, CAD ד. וייסברג, היסטוריה של
https://www.shapr3d.com/blog/history-of-cad?utm_campaign=cadhistorynet.
 [תאריך גישה: 15 מרץ 2025].
- [60] מסמך מידע על מודל מידע בנייה, [2002]. באינטרנט. [זמין, ADSK]
https://web.archive.org/web/20060512180953/http://images.adsk.com/apac_sapac_main/files/4525081_BIM_WP_Rev5.pdf#expand. [תאריך גישה: 15 מרץ 2025].
- [61] מסמך מידע על מודל מידע בנייה בפועל], באינטרנט. [זמין, ADSK]
https://web.archive.org/web/20060512181000/http://images.adsk.com/apac_sapac_main/files/4525077_BIM_in_Practice.pdf. [תאריך גישה: 15 מרץ 2025].
- [62] סגור. אירופה מול שאר העולם, BIM 2024 פתוח מול BIM: חלק 2. BIM. א. בויוק, לוחמת לוביסטים ופיתוח
 [זמין]. באינטרנט. [2025].
<https://bigdataconstruction.com/lobbyist-wars-and-the-development-of-bim-part-2-open-bim-vs-closed-bim-revit-vs-archicad-and-europe-vs-the-rest-of-the-world/>.

- [63] BIMs, 2024. א. בויוקו, לובי מלחמות על נתונים בתחום הבנייה | טכנולוגיה-פיאודליזם וההיסטוריה של [זמין]. https://youtu.be/S-TNdUgfHxk?si=evM_v28KQbGOG0k&t=1360.
- [64] אדסק, מסמך מידע על מודל מידע בנייה, 2002. באינטרנט. [זמין]. https://web.archive.org/web/20060512180953/http://images.autodesk.com/apac_sapac_main/files/4525081_BIM_WP_Rev5.pdf#expand. [תאריך גישה: 15 מרץ 2025].
- [65] אינטגרציה בעיצוב-דרך-ייצור: יתרונות והיגיון, באינטרנט. [זמין]. https://web.archive.org/web/20010615093351/http://www3.adsk.com:80/adsk/files/734489_Benefits_of_MAI.pdf. [תאריך גישה: 15 מרץ 2025].
- [66] מ. שקלט, נתונים מובנים ולא מובנים: הבדלים מרכזיים, 2024. באינטרנט. [זמין]. <https://www.datamation.com/big-data/structured-vs-unstructured-data/>.
- [67] ק. וולארד, הסקת מסקנות על עליית נתונים לא מובנים, 2024. באינטרנט. [זמין]. <https://automationhero.ai/blog/making-sense-of-the-rise-of-unstructured-data/>.
- [68] א. ס. א. ג. ל. ד. ג. א. ל. ט. ג. מייקל פ. גלהר, ניתוח עלויות של חוסר אינטראופרביליות, 2004. באינטרנט. [זמין]. <https://nvlpubs.nist.gov/nistpubs/gcr/2004/nist.gcr.04-867.pdf>. [תאריך גישה: 02: 2025].
- [69] קרואדפלוור, דוח מדע הנתונים 2016, 2016. באינטרנט. [זמין]. https://visit.figure-eight.com/rs/416-ZBE-142/images/CrowdFlower_DataScienceReport_2016.pdf. [תאריך גישה: 15 מרץ 2025].
- [70] אנליטיקס אינדיה מג, 6 המשימות הגוזלות את הזמן הרב ביותר מדאטה סיינטיסטים, 15 במאי 2019. באינטרנט. [זמין]. <https://analyticsindiamag.com/ai-trends/6-tasks-data-scientists-spend-the-most-time-doing/>.
- [71] ביזרפורט, דו"ח: מדעני נתונים מבלים את רוב הזמן בניקוי, 06 יולי 2015. באינטרנט. [זמין]. <https://web.archive.org/web/20200824174530/http://www.bizreport.com/2015/07/report-data-scientists-spend-bulk-of-time-cleaning-up.html>. [תאריך גישה: 5 מרץ 2025].
- [72] במדע: תשובות של סטיבן הוקינג, 27 ביולי 2015. באינטרנט. [זמין]. https://www.reddit.com/r/science/comments/3nyn5i/science_ama_series_stephen_hawking_ama_answers/. [תאריך גישה: 15 במרץ 2025].
- [73] ב. סייפרס וק. דוקטורוב, פרטיות ללא מונופול: הגנת נתונים והתאמה, 2024. באינטרנט. [זמין].

<https://www.eff.org/wp/interoperability-and-privacy>.

- [74] מכון מקינזי העולמי, נתונים פתוחים: שחרור חדשנות וביצועים עם מידע נוזלי "1 באוקטובר 2013. [באינטרנט. זמין]: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information>. [תאריך גישה: 15 במרץ 2025].
- [75] AUTOLISP, intelliCAD, א. בוויקו, המאבק למען נתונים פתוחים בתעשיית הבנייה. ההיסטוריה של openDWG, ODA ו-openCASCADE, 15 05 2024. [זמין]: <https://boikoartem.medium.com/the-struggle-for-open-data-in-the-construction-industry-2b97200e6393>. [תאריך גישה: 16 02 2025].
- [76] ויקיפדיה, מיקרוסופט וקוד פתוח, «באינטרנט. [זמין]: https://en.wikipedia.org/wiki/Microsoft_and_open_source. [תאריך גישה: 15 מרץ 2025].
- [77] זמן, הפער בין מודלים של בינה מלאכותית פתוחים וסגורים עשוי להתכווץ. הנה למה זה חשוב, 5 בנובמבר 2024. [באינטרנט. זמין]: <https://time.com/7171962/open-closed-ai-models-epoch/>. [תאריך גישה: 15 במרץ 2025].
- [78] הוורג, 'יותר מרבע מהקוד החדש בגוגל נוצר על ידי בינה מלאכותית', 29 אוקטובר 2024. [באינטרנט. זמין]: <https://www.theverge.com/2024/10/29/24282757/google-new-code-generated-ai-q3-2024>. [תאריך גישה: 15 מרץ 2025].
- [79] להאצת עיבוד אנליטיקה, "15 בדצמבר 2020 GPUs-מכונני מקינזי דיגיטל", המקרה העסקי לשימוש ב. [באינטרנט. זמין]: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/the-business-case-for-using-gpus-to-accelerate-analytics-processing>. [תאריך גישה: 15 במרץ 2025].
- [80] טרברס סמית, הסוד הפתוח: תוכנה בקוד פתוח "2024. [באינטרנט. זמין]: <https://www.pwc.de/de/digitale-transformation/open-source-monitor-research-report-2019.pdf>. [תאריך גישה: 15 מרץ 2025].
- [81] טרברס סמית, הסוד הפתוח: תוכנה בקוד פתוח "2024. [באינטרנט. זמין]: <https://www.traverssmith.com/knowledge/knowledge-container/the-open-secret-open-source-software/>. [תאריך גישה: 15 מרץ 2025].
- [82] זמין. [באינטרנט]. 2021, תאגידים בשינויים הנתונים העברת תהליך, דלויט: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/finance/us-the-data-transfer-process-in-corporate-transformations.pdf>. [גישה תאריך: 15 מרץ 2025].

- [83] זמין. [באינטרנט]. 20 מרץ 2024, ממשלתיים פרויקטים במימוש מלאכותית ובינה אנליטיקה נתוני gov.uk, <https://www.gov.uk/government/publications/data-analytics-and-ai-in-government-project-delivery/data-analytics-and-ai-in-government-project-delivery>. [5: תאריך גישה: 2025 מרץ].
- [84] זמין. [באינטרנט]. 13 במאי 2011. יותר פשוט לא אך, האפשר ככל פשוט להיות צריך הכל: מקור ציטוט <https://quoteinvestigator.com/2011/05/13/einstein-simple/>. [15: במרץ]. [2025].
- [85] זמין. [באינטרנט], (עמוקה למידה ארכיטקטורת) טרנספורמר [https://en.wikipedia.org/wiki/Transformer_\(deep_learning_architecture\)](https://en.wikipedia.org/wiki/Transformer_(deep_learning_architecture)). [תאריך: 15 מרץ 2025].
- [86] סטטיסטיקות הורדת חבילות פייתון, 2024. באינטרנט. [זמין] <https://www.pepy.tech/projects/pandas>.
- [87] זמין. [באינטרנט]. 2023, "בפייתון המובילות הספריות 10", ביט ראיון <https://www.interviewbit.com/blog/python-libraries/#:~:text=With%20more%20than%20137%2C000%20libraries,data%20manipulation%2C%20and%20many%20more>. [30: מרץ 2025].
- [88] 2025. במרץ 7, עבודה תחנות על הגנרטיבית המלאכותית הבינה ואת הנתונים מדע את מזרזות HP ו-NVIDIA <https://nvidianews.nvidia.com/news/nvidia-hp-supercharge-data-science-generative-ai-workstations>. [15: במרץ 2025].
- [89] עם: מיליוני שורות בשניות, 2024. באינטרנט. [זמין DataFrame ר. אורק, כיצד לעבד <https://towardsdatascience.com/how-to-process-a-dataframe-with-millions-of-rows-in-seconds>.
- [90] זמין. [באינטרנט]. 2024, קגל זה מה, אוסלו. צ. <https://www.datacamp.com/blog/what-is-kaggle>.
- יוני 2024. באינטרנט. [זמין 2, COMPUTEX 2024-ג'נסן חואנג, נאום מפתח ב NVIDIA נשיא <https://www.youtube.com/live/pKXDVSWZmUU?si=Z3Rj1Las8wiPII2w>. [תאריך גישה: 15 מרץ 2025].
- חברים: מייסדים וחברים תאגידיים, 2024. באינטרנט. [זמין] <https://www.opendesign.com/member-showcase>.

עבור IFC אחרים מוכנים לוותר על CAD וספקי Autodesk הוא דבר מהעבר או מדוע IFC א: בוויקו, עידן השינוי
ב-14 עובדות מפתח, 24 נובמבר 2024]. באינטרנט. USD

<https://boikoartem.medium.com/the-age-of-change-ifc-is-a-thing-of-the-past-or-why-adsk-and-other-cad-vendors-are-willing-to-3f9a82ccd10a>. [תאריך גישה: 23 פברואר 2025].

המעבר לנתונים ולתהליכים והאם תעשיית הבנייה זקוקה לסמנטיקה, פורמטים BIM. א. בוויקו, העולם שלאחר
[תאריך גישה: 23 פברואר 2025]. <https://boikoartem.medium.com/the-post-bim-world-7e35b7271119>. [20 דצמבר 2024]. באינטרנט. [זמין]

למדע הנתונים, 2016]. באינטרנט. NIH תכנית אסטרטגית של N. I. o. Health, https://datascience.nih.gov/sites/default/files/NIH_Strategic_Plan_for_Data_Science_Final_508.pdf. [תאריך גישה: 23 פברואר 2025].

נתונים גרועים עולים לארה"ב 3 טריליון דולר בשנה, 22 ספטמבר 2016, Harvard Business Review, <https://hbr.org/2016/09/bad-data-costs-the-u-s-3-trillion-per-year>. [זמין]

<https://delpha.io/impacts-of-data-quality/>. דלפה, השפעות איכות הנתונים, 1 ינואר 2025]. באינטרנט. [זמין]

עיצוב לתחזוקה: החשיבות של שיקולי תפעול ותחזוקה במהלך שלב העיצוב של פרויקט, W. B. D. Guide, <https://www.wbdg.org/resources/design-for-maintainability>. [תאריך גישה: 15 מרץ 2025]. באינטרנט. [זמין]

מדריך לתכנון מניעת קורוזיה ושליטה עבור מערכות וציוד צבאיים, אפריל, O. o. D. C. P. a. Oversight, 2014. באינטרנט. [זמין]. <https://www.dau.edu/sites/default/files/Migrated/CopDocuments/CPC%20Planning%20Guidebook%204%20Feb%2014.pdf>. [תאריך גישה: 15 מרץ 2025].

גארטנר, איכות נתונים: שיטות עבודה מומלצות לתובנות מדויקות, 1 ינואר 2025]. באינטרנט. [זמין]. <https://www.gartner.com/en/data-analytics/topics/data-quality>.

https://en.wikipedia.org/wiki/For_Want_of_a_Nail. עבור רצון של מסמר, באינטרנט. [זמין]. [תאריך גישה: 15 מרץ 2025].

מכון מקינזי העולמי, נתונים פתוחים: שחרור חדשנות וביצועים עם מידע נוזלי, אוקטובר 2013]. באינטרנט. [זמין]. <https://www.mckinsey.com/~media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/open%20data%20unlocking%20innovation%20and%20per>

- formance%20with%20liquid%20information/mgi_open_data_fullreport_oct2013.pdf
[תאריך גישה: 15 מרץ 2025].
- יולי 2024]. באינטרנט. PWC, "ESG מודעות", 1 זמין. <https://www.pwc.com/kz/en/assets/esg-awareness/kz-esg-awareness-rus.pdf>.
[תאריך גישה: 15 מרץ 2025].
- [105 :באינטרנט. [זמין. 2024. (ICE), ג. המונד", פחמן מגולל - המלאי של פחמן ואנרגיה
] <https://greenbuildingencyclopaedia.uk/wp-content/uploads/2014/07/Full-BSRIA-ICE-guide.pdf>.
- [106 חישוב הפחמן המגולל, 2024]. באינטרנט. CO2_זמין. https://github.com/datadrivenconstruction/CO2_calculating-the-embodied-carbon.
- [107 מקינזי", דמיון העתיד הדיגיטלי של הבנייה, "24 יוני 2016]. באינטרנט. [זמין
] <https://www.mckinsey.com/capabilities/operations/our-insights/imagining-constructions-digital-future>. [תאריך גישה: 25 פברואר 2025].
- [108 בונד של משלמי המסים בגרמניה בע"מ", הספר השחור, "10 אוקטובר 2024]. באינטרנט. [זמין
] <https://steuerzahler.de/aktuelles/detail/das-schwarzbuch-202425/>. [תאריך גישה: 15 מרץ 2025].
- [109 סאס", אגם נתונים ומחסן נתונים – הכירו את ההבדל], באינטרנט. [זמין
] https://www.sas.com/en_is/insights/articles/data-management/data-lake-and-data-warehouse-know-the-difference.html. [תאריך גישה: 15 מרץ 2025].
- [110 מודל מידע בנייה, 2002]. באינטרנט. ADSK, [זמין. https://www.laiserin.com/features/bim/autodesk_bim.pdf. [תאריך גישה: 15 מרץ 2025].
- [111 :באינטרנט. [זמין. 2024. BIM, א. בויוק, מפה של היסטוריית
] <https://bigdataconstruction.com/history-of-bim/>.
- [112 על ידי ארגונים וסטנדרטים, 27 דצמבר 2023]. באינטרנט. [זמין. BIM א. ס. בורקובסקי, הגדרות של
] <https://encyclopedia.pub/entry/53149>. [תאריך גישה: 5 מרץ 2025].
- [113 :באינטרנט. [זמין. 2012. OPEN BIM, תוכנית CAD, ספק
] https://web.archive.org/web/20140611075601/http://www.graphisoft.com/archicad/open_bim/. [תאריך גישה: 30 מרץ 2025].
- [114 ויקיפדיה", קלאסות יסוד תעשייה], באינטרנט. [זמין

- תאריך גישה: 15 מרץ. https://en.wikipedia.org/wiki/Industry_Foundation_Classes. [2025].
- תאריך גישה: 30 מרץ. <https://en.wikipedia.org/wiki/IGES>. [2025].
- בדצמבר 2021. באינטרנט. [זמין 15, CAD (BIM) א. בוויקו, היסטוריה של [https://miro.com/app/board/o9J_laML2cs=/. \[2025\]](https://miro.com/app/board/o9J_laML2cs=/) בפברואר 2025].
- T. K. K. A. O. F. B. C. E. L. H. H. E. L. P. N. S. H. T. J. v. L. H. G. D. H. T. K. C. L. A. W. חלק א, '8 ינואר - GeoBIM 2019 הבנצ'מרק: IFC פרנססקה נוארדו, מחקר הפניה של תמיכת תוכנת J. S. [תאריך גישה: 5 מרץ 2025]. <https://arxiv.org/pdf/2007.10951>. [2021].
- מקומי "13 אפריל 2021. באינטרנט [BIM | IFC | מקסים נצ'פורנקו | רגה: BIM א. רוגצ'וב", נדבר על <https://www.youtube.com/watch?t=3000&v=VO3Y9uuzF9M&feature=youtu.be>. [2025].
- בנדל"ן: מדוע זה חיוני ליעילות ולצמיחה "17 בדצמבר 2024. באינטרנט. [זמין RETS", ד. ארס: <https://www.realpha.com/blog/rets-importance-in-real-estate-explained>. [תאריך גישה: 5 במרץ 2025].
- <https://www.adsk.com/buying/flex?term=1-YEAR&tab=flex>: עלות טוקן פלקס, 2024. באינטרנט. [זמין].
- ודמוקרטיזו את הגישה לנתונים (17 קולוקוויום משקיעים – אוניברסיטה – תעשיית BIM-א. בוויק", שכחו מ <https://www.bim.bayern.de/wp-content/uploads/2023/06/Kolloquium-17-TUM-Bauprozessmanagement-und-Bay-Bauindustrie.pdf>. [2024]. באינטרנט. [זמין].
- ד. היל, ד. פולדזי, ס. פרר, מ. פרידמן, א. לוך ו-פ. פלאשקה, פתרון חידת הפרודוקטיביות בתעשיית הבנייה, <https://www.bcg.com/publications/2015/engineered-products-project-business-solving-construction-industrys-productivity-puzzle>. [2015].
- סקופ – סביבת נתוני פרויקטים ודימוי של מוצרים בנויים רב-תכליתיים עם דגש על מעטפת הבניין, 1 ינואר [2018]. באינטרנט. [זמין]. <https://www.ise.fraunhofer.de/de/forschungsprojekte/scope.html>. [2025].
- כדי לקדם סטנדרטים פתוחים לתוכן תלת-OpenUSD-אפל. קום, פיקסאר, אדובי, אפל ונבדיה מקימים ברית ל [2023]. באינטרנט. [זמין].

- [] <https://www.apple.com/newsroom/2023/08/pixar-adobe-apple-adsk-and-nvidia-form-alliance-for-openusd/>. [תאריך גישה: 2 במרץ 2025].
- [125] AECmag, באינטרנט. [זמין 25, ADSK, אסטרטגיית הנתונים הגרנולרית של, יולי 2024]
[] <https://aecmag.com/technology/autodesks-granular-data-strategy/>. [15: תאריך גישה: 2025 מרץ].
- [126] IFC אחרים מוכנים לוותר על CAD וספקי ADSK הוא דבר מהעבר או מדוע IFC: א. בויוקו, "עידן השינוי
[] ב-14 עובדות מפתח", 24 בנובמבר 2024. [באינטרנט. [זמין USD לטובת
<https://boikoartem.medium.com/the-age-of-change-ifc-is-a-thing-of-the-past-or-why-adsk-and-other-cad-vendors-are-willing-to-3f9a82ccd10a>. [23: תאריך גישה: 2025 בפברואר].
- [127] בבנייה, 7 באוגוסט ChatGPT ו-LLM הקרב על נתונים ויישום של | ENG BIM Cluster 2024, א. בויוקו
[] ChatGPT ו-LLM הקרב על נתונים ויישום של | ENG BIM Cluster 2024: [באינטרנט. [זמין 2024.
[בבנייה]. [תאריך גישה: 15 במרץ 2025]
- [128] <https://zeldman.com/2008/05/06/content-precedes-design/>: ג'פרי זלדמן מציג, 6 במאי 2008. [באינטרנט. [זמין
[] [תאריך גישה: 15 במרץ 2025].
- [129] בבנייה מונעת נתונים, 5 מרץ 2024. [באינטרנט. [זמין | ChatGPT ניתוח עם DWG, א. בויוקו
[] <https://www.kaggle.com/code/artemboiko/dwg-analyse-with-chatgpt-datadrivenconstruction>. [15: מרץ 2025]. [תאריך גישה: 2025 מרץ].
- [130] מכנסי, מדריך מכנסי להתמודדות בעידן הדיגיטלי והבינה המלאכותית, 2023. [באינטרנט. [זמין
[] <https://www.mckinsey.com/featured-insights/mckinsey-on-books/rewired>. [תאריך: 2025 מרץ 30: גישה].
- [131] פורבס, "סיפור נתונים: מיומנות מדע הנתונים החיונית שכולם צריכים", 31 מרץ 2016. [באינטרנט. [זמין
[] <https://www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-the-essential-data-science-skill-everyone-needs/>. [15: מרץ 2025]. [תאריך גישה: 2025 מרץ].
- [132] ג. 'ברטין, גרפיקה ועיבוד מידע גרפי, 8 בספטמבר 2011. [באינטרנט. [זמין
[] https://books.google.de/books/about/Graphics_and_Graphic_Information_Process.html?id=csqX_xnm4tcC&redir_esc=y. [15: מרץ 2025]. [תאריך גישה: 2025 מרץ].
- [133] וולס/ווילקס על חשיבה סטטיסטית, באינטרנט. [זמין, CauseWeb
[] <https://www.causeweb.org/cause/resources/library/r1266>. [15: מרץ 2025]. [תאריך גישה: 2025 מרץ].

- [134] :מגזין משרד, "איך המדע גילה את הבריאה", ינואר 1986. באינטרנט. [זמין]
] <https://www.ministrymagazine.org/archive/1986/01/how-science-discovered-creation>. [תאריך גישה: 15 מרץ 2025].
- [135] :המרה מונחית נתונים: האצה בקנה מידה עכשיו "23 מאי 2017. באינטרנט. [זמין], BCG
] <https://www.bcg.com/publications/2017/digital-transformation-transformation-data-driven-transformation>. [תאריך גישה: 15 מאי 2024].
- [136] :איך לבנות ארכיטקטורת נתונים כדי להניע חדשנות - היום ומחר, 3 ביוני 2020. באינטרנט. [זמין]
] <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/how-to-build-a-data-architecture-to-drive-innovation-today-and-tomorrow>. [תאריך גישה: 15 מרץ 2025].
- [137] :אוקספורד, "וודרו וילסון 1856–1924", באינטרנט. [זמין]
] <https://www.oxfordreference.com/display/10.1093/acref/9780191866692.001.0001/q-oro-ed6-00011630>. [תאריך גישה: 15 מרץ 2025].
- [138] :ממשקים, 2024. באינטרנט. [זמין]
] <https://datadrivenconstruction.io/index.php/convertors/>.
- [139] סיכום הפרס מהו הערך האמיתי של בינה מלאכותית עבור העסק שלך וכיצד תוכל להפיק תועלת, PWC, 1? [ינואר 2017]. באינטרנט. [זמין]
] <https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf>. [תאריך גישה: 18 פברואר 2025].
- [140] :זמין. באינטרנט. [2024, צנרת בבנייה, <https://datadrivenconstruction.io/index.php/pipeline-in-construction/>]
]
- [141] :ויקיפדיה, "אפאצ'י ניו", 1 ינואר 2025. באינטרנט. [זמין]
] https://de.wikipedia.org/wiki/Apache_NiFi. [תאריך גישה: 5 מרץ 2025].
- [142] :אוטו-מגיב: צור טיוטות תגובות להודעות דוא"ל נכנסות, 1 מאי 2024. באינטרנט. [זמין] n8n, AI גמיל
] <https://n8n.io/workflows/2271-gmail-ai-auto-responder-create-draft-replies-to-incoming-emails/>. [תאריך גישה: 15 מרץ 2025].
- [143] של זילו, גוגל שייטס וגיימיל, 1 במרץ 2025. באינטרנט [API אוטומציה של עסקאות יומיות בנדל"ן עם n8n, זמין]
] <https://n8n.io/workflows/3030-real-estate-daily-deals-automation-with-zillow-api-google-sheets-and-gmail/>. [תאריך גישה: 15 במרץ 2025].

- [144] B. T. O'Neill, 85% = שיעורי כישלון עבור פרויקטים של אנליטיקה, אינטליגנציה מלאכותית ונתוני ענק [ינואר 2025]. באינטרנט. [זמין 1, ikes!]
<https://designingforanalytics.com/resources/failure-rates-for-analytics-bi-iot-and-big-data-projects-85-yikes/>.
- [145] ג. 'ניימן, על שני ההיבטים השונים של השיטה הייצוגית: שיטת הדגימה השכבתית ושיטת הבחירה המכוונת. הוצאת אוניברסיטת אוקספורד, 1934.
- [146] T. J. S. a. J. S. 2025. ג'סי פרלה, בעיה שהביכה את מילטון פרידמן, כלכלה כמותית עם ג'וליה, 1 ינואר 2025. [באינטרנט. [זמין]
 תאריך]. https://julia.quantecon.org/dynamic_programming/wald_friedman.html. גישה: 1 מאי 2024.
- [147] ט. לנדסל-וולפר, חיזוי מצב הרוח של האומה בהווה, משמעות, 2012.
- [148] א. בויקו, "סן פרנסיסקו. מגזר הבנייה 1980-2019", [2024]. באינטרנט. [זמין]
<https://www.kaggle.com/search?q=San+Francisco.+Building+sector+1980-2019>.
- [149] פרויקטים, [2024]. באינטרנט. [זמין RVT IFC 5000 קבצי "Kaggle", א. בויקו
<https://www.kaggle.com/datasets/artemboiko/rvtifc-projects>.
- [150] CFMA, "זמין", <https://cfma.org/articles/preparing-for-the-future-with-connected-construction>. מוכנים לעתיד עם בנייה מחוברת] באינטרנט. [זמין]. תאריך גישה: 15 מרץ 2025.
- [151] נכשלו, [22 מאי 2017]. באינטרנט [IoT סיקו", סקר סיקו מגלה שכמעט שלושה רבעים מהפרויקטים של
 זמין: <https://newsroom.cisco.com/c/r/newsroom/en/us/a/y2017/m05/cisco-survey-reveals-close-to-three-fourths-of-iot-projects-are-failing.html>.
- [152] תנאים הנדרשים לשימור מאובנים של צמחים, [2024]. באינטרנט. [זמין]
<https://ucmp.berkeley.edu/IB181/VPL/Pres/PresTitle.html>.
- [153] על אג"ח, מיזוגים ורכישות, מיתון בארה"ב, בחירות: ריאיון מלא, [2023]. באינטרנט [BlackRock-פינק מ"
 זמין: <https://www.bloomberg.com/news/videos/2023-09-29/blackrock-s-fink-on-m-a-recession-election-full-intv-video>.
- [154] מפורסמים, [02 אוקטובר 2024]. באינטרנט. [זמין AI אסונות 12", CIO, <https://www.cio.com/article/190888/5-famous-analytics-and-ai-disasters.html>.
 [תאריך גישה: 15 מרץ 2025].

- [155] PublicAffairs, 2017. ג. קספרוב, חשיבה עמוקה.
]
- [156] <https://en.wikipedia.org/wiki/Kaggle>. [זמין 1, "Kaggle", ויקיפדיה. ינואר 2025].
] [תאריך גישה: 15 מרץ 2025].
- [157] Kaggle, "זמין". באינטרנט. [1 ינואר 2025].
] <https://www.kaggle.com/competitions/titanic/overview>. [תאריך גישה: 10 מרץ 2025].
- [158] ChatGPT: ש. יוהרי, יצירת. [2024].
] <https://sitn.hms.harvard.edu/flash/2023/the-making-of-chatgpt-from-data-to-dialogue/>.
- [159] פ. דומינגוס, כמה דברים מועילים לדעת על למידת מכונה, "2024", באינטרנט. [זמין].
] <https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf>.
- [160] ג. סאראמאגו, "ציטוט ציטוט", באינטרנט. [זמין].
] <https://www.goodreads.com/quotes/215253-chaos-is-merely-order-waiting-to-be-deciphered>. [תאריך גישה: 17 מרץ 2025].
- [161] NVIDIA, "NVIDIA NeMo Curator Classifier", 19 שפר את נתוני האימון שלך עם מודלים חדשים של NVIDIA, [זמין].
] <https://developer.nvidia.com/blog/enhance-your-training-data-with-new-nvidia-nemo-curator-classifier-models/>. [תאריך גישה: 25 מרץ 2025].
- [162] AI וכלי נתוני Cosmos Foundation מכריזה על שחרור משמעותי של מודלים עולמיים של NVIDIA, [זמין].
] <https://nvidianews.nvidia.com/news/nvidia-announces-major-release-of-cosmos-world-foundation-models-and-physical-ai-data-tools>. [תאריך גישה: 25 מרץ 2025].
- [163] NVIDIA, "NVIDIA Isaac Sim", [זמין].
] <https://developer.nvidia.com/isaac/sim>. [תאריך גישה: 25 מרץ 2025].
- [164] מ. רבעוני, מדוע אסטרטגיות דיגיטליות נכשלות, 25 ינואר 2018. באינטרנט. [זמין].
] <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/why-digital-strategies-fail>. [תאריך גישה: 15 מרץ 2025].
- [165] [באינטרנט]. 2006 נובמבר 17, פרידמן מילטון של האהובות מילותיו, פרי. י.ג. מ.
] <https://www.aei.org/carpe-diem/my-favorite-milton-friedman-quotes/>. [1: תאריך גישה: 2025 מרץ].

- [166] קשרים אחר החיפוש: קוונטום, פיזיקה, מידע, וילר. א. ג. 1990.
- [169] חלק 5: בלאקרוק – מאסטר כל הטכנולוגיות. כיצד תאגידים שולטים. BIM. א. בויוקו, לוחמת לוביסטים ופיתוח
[] <https://boikoartem.medium.com/lobbyist-wars-and-the-development-of-bim-d72ad0111a7d>. באינטרנט. [2024, בקוד פתוח, [זמין]
- [170] T. Krijnen ו-J. Beetz עבור מודלים של בניינים בפורמט בינארי, אינפורמטיקה SPARQL מנוע שאילתות, 2024, הגדסית מתקדמת, 2024.
- [171] מספר העסקים במגזר הבנייה בבריטניה בשנת 2021, לפי גודל העסק, [2024, באינטרנט. [זמין]
[] <https://www.statista.com/statistics/677151/uk-construction-businesses-by-size/>.
- [172] [באינטרנט. [זמין]. IFC&RVT, 2024, פרויקטים 5000
[] <https://www.kaggle.com/code/artemboiko/5000-projects-ifc-rvt-datadrivenconstruction-io>.
- [173] מ. פופובה, "זה מתוך בית: הפיזיקאי החלוץ ג'ון ארצ'בלד וילר על מידע, טבע המציאות ולמה אנו חיים ביקום
[] <https://www.themarginalian.org/2016/09/02/it-from-bit-wheeler/>. [2008, "משתתף", באינטרנט. [זמין]. [תאריך גישה: פברואר 2025]
- [174] [סרט. [גרמניה. BIM מלחמות לובינג על נתונים בבנייה | טכנולוגיה-פיאודליזם וההיסטוריה של העבר הנסתר של
[] ארטיים בויוקו, 2023.
- [175] שליפת מסמכים ונתונים אוטומטית מפרויקטים, 16 נובמבר | IFC ו-REVIT עם CHATGPT, א. בויוקו
[] תאריך גישה: https://www.youtube.com/watch?v=ASXolti_YPs&t. [באינטרנט. [זמין]. 2023. [מרץ 2025]
- [176] מ. וחברה, שלושה מנדטים חדשים לתפיסת הערך המלא של טרנספורמציה דיגיטלית, 22 ינואר 2022
[] <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/three-new-mandates-for-capturing-a-digital-transformations-full-value>. [באינטרנט. [זמין]. [תאריך גישה: 15 פברואר 2025]
- [177] בנייה בעולם הדיגיטלי, 1 במאי 2021. [באינטרנט. [זמין]. KPMG,
[] <https://assets.kpmg.com/content/dam/kpmg/ca/pdf/2021/05/construction-in-the-digital-age-report-en.pdf>. [2024, באפריל 5, [תאריך גישה]
- [178] זמן להכרעה, 17 מרץ 2023. [באינטרנט. [זמין]. LLP, KPMG, Cue Construction 4.0:
[] <https://kpmg.com/ca/en/home/insights/2023/05/cue-construction-make-or-break>

[תאריך גישה: 15 פברואר 2025]. time.html.

- [179 ינואר 2025]. באינטרנט [10 SaaS, א. ביזנס, סאטיה נדלה חושף כיצד סוכני בינה מלאכותית יפרו את מודלי
] זמין: <https://www.outlookbusiness.com/artificial-intelligence/microsoft-ceo-satya-nadella-reveals-how-ai-agents-will-disrupt-saas-models>. [תאריך גישה: 15 מרץ 2025].
- [180 פורבס", ניקוי נתוני ביג: משימת מדע הנתונים הכי צורכת זמן, הכי לא מהנה, סקר אומר "23 מרץ 2016
] <https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/>. [תאריך גישה: 15 מרץ 2025].
- [181 משרד החוץ, לענייני חבר העמים ופיתוח בריטניה, אסטרטגיית פיתוח דיגיטלי 2024 עד 2030, 18 במרץ
] <https://www.gov.uk/government/publications/digital-development-strategy-2024-to-2030/digital-development-strategy-2024-to-2030>. [תאריך גישה: 15 במרץ 2025].
- [182 חזון ואסטרטגיה בתעשיית עיצוב הבניינים, 7 בנובמבר 2003]. באינטרנט. [זמין
] <https://web.archive.org/web/20030711125527/http://usa.adsk.com/adsk/servlet/item?id=2255342&siteID=123112>. [תאריך גישה: 5 במרץ 2025].
- [183 מ. בוצ'ארוב", מודליזציה מידע "מרץ 2025]. באינטרנט. [זמין
] <https://www.litres.ru/book/mihail-evgenevich-bocharov/informacionnoe-modelirovanie-v-rossii-71780080/chitat-onlayn/?page=5>. [תאריך גישה: 15 מרץ 2025].
- [184 עיצוב משולב-דרך-ייצור: יתרונות והגיון, 2000]. באינטרנט. [זמין
] https://web.archive.org/web/20010615093351/http://www3.autodesk.com:80/adsk/files/734489_Benefits_of_MAI.pdf. [תאריך גישה: 25 מרץ 2025].
- [185 היא קמפיין שיווקי, 12 מרץ 2012]. באינטרנט. [זמין Open BIM תוכנית, CAD ספק
] <https://web.archive.org/web/20120827193840/http://www.graphisoft.com/openbi> m/. [תאריך גישה: 30 מרץ 2025].

אינדקס נושאים

337, 351, 356, 358, 361, 389, 413, 416, 448, 462, 466,
475, 476, 492, 494, 497.
243, 242, 240, בלאקבוקס
ביה"ס, 8
בוכה, 337, 320
בומ, 263, 257, 79, 77, 76
קופסת גבול, 480, 415, 414, 394, 393, 392, 373, 234,
ברפ, 284, 283, 276, 264, 263, 234, 142,

2'

CAD, 6, 14, 18, 24, 56, 57, 58, 63, 70, 71, 72, 73, 74, 75, 76,
77, 78, 79, 80, 84, 85, 95, 97, 111, 126, 137, 138, 139,
140, 141, 142, 144, 146, 147, 152, 153, 155, 156, 166,
172, 175, 183, 184, 186, 187, 190, 195, 196, 206, 208,
210, 211, 213, 214, 215, 216, 217, 218, 219, 220, 221,
222, 224, 227, 228, 232, 234, 237, 238, 239, 242, 243,
251, 252, 255, 256, 257, 258, 259, 260, 261, 262, 263,
264, 265, 266, 271, 272, 273, 274, 275, 276, 277, 278,
279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289,
291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301,
303, 308, 309, 324, 328, 337, 344, 351, 356, 358, 361,
367, 372, 376, 383, 389, 402, 403, 405, 413, 416, 431,
448, 462, 466, 475, 476, 480, 492, 494, 497
CAE, 16, 283
CAFM, 14, 24, 62, 84, 151, 153, 172, 233, 278, 326, 356,
387, 458, 473
קאמ, 16, 78
הוצאות הון, 14, 82
CDE, 84, 175, 388, 389, 390
צ'אטג'יפיטי, 103, 104, 109, 110, 124, 129, 221, 303, 307,
343, 349, 421
קלוד, 103, 104, 109, 129, 133, 166, 219, 221, 300, 329, 343,
349, 381, 429, 443
CO₂, 72
CO₂, 229, 234, 235, 236, 237, 238, 239, 243
קובי, 156, 292, 289
מרכז מציניות, 56, 168, 169, 170, 477
קופילוט, 116, 114, 459
CPIXML, 143, 272, 273, 276, 277, 279, 285, 296
סי.פי.אם, 17, 14, 62, 175, 166, 326, 473
CQMS, 14, 84, 177, 178, 462
CRM, 109, 369, 459
CRUD, 51, 459
CSG, 263
CSV, 61, 62, 88, 89, 120, 123, 128, 129, 130, 131, 135, 168,
268, 272, 280, 329, 333, 344, 354, 356, 373, 377, 378,

263, 234, 232, 215, 210, 191, 84, 73, 71, 14, 8, תלת מימד,
276, 277, 287, 298, 299, 302, 303, 306, 307, 337, 375,
393, 394, 448, 468, 480

4287, 237, 234, 229, 210, 199, 196, 172, 84, ד
תעשייה 43.4.0

5D, 84, 172, 196, 210, 237, 287, 492

6235, 234, 233, 232, 229, 196, 172, τ

7287, 234, 233, 232, T

8287, 234, 233, 232, 229, 196, 172, T

 $\approx$

בינה מלאכותית, 3, 52, 100, 102, 103, 106, 107, 116, 395
457, 459, 461, 477
איה, 289
מטרה, 3, 289
AMS, 14, 84, 151, 153
אפאצ'י איירפולוא, 311, 361, 362, 363, 364, 366, 367, 369
371, 399, 494
אפקה ניפי, 116, 311, 361, 367, 368, 371, 399, 482, 494
אפקה אורק, 62, 330, 378
אפאצ'י פארקט, 62, 67, 380, 381, 494
API, 54, 94, 95, 97, 109, 137, 138, 154, 168, 218, 219, 256,
257, 260, 271, 273, 294, 295, 296, 297, 300, 308, 326,
329, 342, 366, 369, 370, 488

ב

ב"דס, 259, 258,
נחוני ענק, 9, 67, 245, 383,
בימ, 2, 3, 4, 6, 14, 17, 24, 56, 58, 60, 70, 71, 72, 73,
74, 75, 76, 77, 78, 79, 80, 81, 84, 92, 137, 138, 139, 140,
141, 142, 144, 150, 154, 156, 166, 172, 183, 184, 186,
187, 190, 195, 196, 208, 210, 211, 213, 216, 217, 218,
219, 220, 221, 222, 227, 228, 237, 238, 239, 242, 243,
246, 250, 251, 252, 255, 256, 257, 258, 259, 260, 261,
262, 263, 266, 271, 272, 273, 275, 276, 277, 278, 279,
280, 282, 285, 287, 288, 289, 290, 291, 292, 293, 294,
295, 296, 297, 298, 299, 300, 301, 308, 309, 324, 328.

379, 380, 401, 407, 414, 474, 480

361, 362, 363, 365, 383, 384, 450, 475

ת

פ

דאָי, 414, 285, 284, 281, 280, 278, 277, 276,

דאג, 366, 365, 363, 362,

דש, 337, 336, 320,

ממשלת נתונים, 401, 400, 398, 396, 395, 373,

אגם נתונים, 389, 388, 387, 386, 385, 384, 376, 373, 214, 390, 400

אגם נתונים, 388, 387, 386, 373, 67,

מינימליזם נתוני, 401, 400, 396, 395, 373,

ביצת נתונים, 401, 397, 395, 373,

מחסן נתונים, 400, 383, 382,

נתונים-כשרות-כשירות, 487,

מונחה נתונים, 486, 484, 461, 460, 170, 50,

דאטה פריים, 131, 130, 129, 125, 123, 122, 121, 117, 67,

133, 134, 135, 137, 219, 220, 224, 237, 308, 328, 329,

330, 332, 333, 344, 345, 347, 348, 354, 365, 377, 381,

409, 414, 415, 475

דאטאופס, 475, 401, 400, 399, 398, 373, 170,

דיפסיק, 219, 166, 133, 129, 124, 110, 109, 107, 104, 103,

221, 300, 329, 343, 349, 381, 421, 429, 443, 494

דגן, 357, 227, 186, 140, 8,

דז"ח, 302, 287, 272, 227, 211, 186, 140, 97, 73, 71, 70, 8,

303, 304, 307, 357, 376, 497

דזה, 390, 389, 388, 387, 386, 384, 383, 382, 376, 373, 67,

דקס, 277, 73, 8,

"האות"

ECM, 58, 175

ECS, 142

אייר, 289,

eLOD, 289

הוראת אנגלית כשפה זרה, 385, 384,

אָיפּי, 198, 166, 14,

ERP, 2, 11, 12, 14, 17, 18, 24, 25, 58, 62, 109, 153, 166,

172, 175, 196, 198, 210, 232, 239, 240, 241, 242, 243,

244, 245, 246, 247, 249, 272, 277, 278, 279, 282, 326,

351, 356, 361, 369, 387, 388, 389, 390, 458, 462, 473,

475, 487, 492

ESG, 196, 235, 236, 238

ETL, 6, 1.1-8, 32, 49, 81, 113, 116, 119, 128, 188, 193, 219,

291, 311, 312, 317, 323, 324, 325, 326, 327, 328, 330,

331, 333, 338, 339, 340, 343, 344, 348, 349, 350, 351,

353, 354, 356, 361, 362, 363, 364, 365, 367, 371, 372,

381, 382, 383, 384, 385, 399, 409, 430, 475, 481, 494

אקסל, 132, 125, 123, 120, 111, 88, 85, 66, 65, 62, 61, 57,

154, 167, 187, 210, 223, 224, 226, 227, 228, 278, 291,

329, 333, 342, 351, 365, 376, 378, 459, 475

חילוך, 345, 330, 328, 326, 324, 323, 311, 193, 134, 128, 81,

נוצת, 378, 330, 123, 62,

FPDF, 339, 340, 341, 342, 343

ג

GDPR, 109

GIS, 58

GLTF, 143, 278

גוגל שיטס, 370, 368,

גרזק, 349, 343, 329, 300, 221, 219, 166, 133, 129, 104, 103,

381, 429, 443

ה

HDF5, 62, 67, 123, 329, 330, 378, 379, 380

HiPPo, 29, 37, 95, 424, 477, 484, 490

HTML, 123, 340, 365, 370

אנ

IDS, 289, 290, 291

IFC, 8, 73, 138, 142, 186, 227, 261, 262, 263, 264, 265, 266,

267, 268, 272, 273, 276, 277, 278, 279, 280, 284, 286,

292, 296, 302, 329, 357, 414, 415, 417, 497

IGES, 262, 263, 276

iLOD, 289

IoT, 10, 18, 67, 271, 367, 369, 405, 413, 417, 418, 419, 455,

460, 465, 482, 484

ISO 19650, 388

י

JavaScript, 320, 369, 378

JSON, 88, 89, 90, 92, 123, 128, 142, 269, 272, 280, 329,

330, 333, 378, 474, 480

Jupyter Notebook, 114, 115, 116, 130, 187, 224, 330, 346,

417, 425

ק

Kaggle, 115, 121, 130, 187, 224, 303, 307, 330, 346, 408,

415, 417, 425, 426, 430, 431, 433

k-NN, 392, 393, 442, 445, 446, 447, 448

KPI, 245, 311, 317, 318, 319, 320, 321, 324, 353, 372, 389,

478, 479

U

Uniclass, 154, 155, 156

285, 284, 281, 280, 279, 278, 277, 276, 255, 143, 142, דולר,
286, 292, 296

286, 292, 296

V

401, 400, 398, 373, וקטורופ

VR, 84, 271, 285

1

וויטבוקס, 240, 242, 243

X

XLSX, 8, 61, 62, 123, 128, 129, 231, 268, 276, 277, 280, 296,

299, 308, 351, 373, 377, 378, 379, 380, 401, 474, 480

XML, 61, 88, 89, 92, 128, 143, 269, 272, 277, 280, 291, 299,

329, 378, 414, 474

XLSX, 8, 61, 62, 123, 128, 129, 231, 268, 276, 277, 280, 296,

299, 308, 351, 373, 377, 378, 379, 380, 401, 474, 480

XML, 61, 88, 89, 92, 128, 143, 269, 272, 277, 280, 291, 299,