לשכת המבקרים הפנימיים IIA ישראל (חל"צ)
IIA Israel - Institute of Internal Auditors

הכנס המקצועי השנתי של הביקורת הפנימית 2026

Internal Audit NextGen2.0

הרצאה בנושא:

מסגרת לניהול סיכוני AI

מרצה: ערן רז, שותף PwC Israel

אימייל: Eran.Raz@pwc.com

נייד: 054-6660260

חברות נמצאות בשלבים שונים במסע הבינה המלאכותית שלהן... היכן אתם?

אופטימלי	מנוהל	מוגדרת	בתכנון	יזום ראשוני	טרם החל
הבינה המלאכותית הוטמעה במרקם העסקי והאסטרטגי שלכם	ניהול AI הינו משולב ומנוהל כאסטרטגיה בחברה. השפעת הטרנספורמציה הייתה גבוהה בחברה.	הוגדרה מסגרת לניהול AI בחברה לרבות פרקטיקות לניהול ומסגרת בקרות כולל ביצוע פיילוט / או קיים ארגז חול לניהול פיילוטים	מאמצעי AI בתכנון בחברה עם הגדרת מפת דרכים, תכנון בקרות וממשל	קיימים מספר תהליכים ראשוניים בחברה לרבות מודעות לנושא ואוטומציה של מספר תהליכים	הנושא של AI טרם יושם בחברה. לא קיימות יוזמות או אימוץ כלים של צ'טים.

פרקטיקות אחראיות של AI יכולות להביא ערך

ארגונים המאמצים בינה מלאכותית בקנה מידה רחב שואפים להניע יתרון תחרותי, להפחית חיכוכים ולבנות אמון. יישום פרקטיקות בינה מלאכותית אחראיות יכול לסייע להניח את היסודות לאמון וגמישות ולאפשר ביצוע מהיר יותר, החלטות חכמות יותר וערך מתמשך.

שיפור המוכנות לדרישות השוק והרגולציה

יצירת המסגרת, התיעוד והבקרות הנדרשות לעמוד עם דרישות השוק וציפיות הרגולטור

שיפור יכולות הפיתוח והפיקוח

האפשרות ליצור מעגלי פיתוח מהירים יותר והקטנת חיכוכים פנימיים תוך מסגרת ניהול סיכונים מובנית ואחריות מוגברת

הגברת הביטחון בקרב בעלי העניין

בניית בטחון ושקיפות מול הדירקטוריון, רגולטורים, משקיעים ולקוחות

חיבור טוב יותר של AI עם יוזמות אסטרטגיות

העצמת צוותים למקד את מאמצי הבינה המלאכותית במה שחשוב ביותר באמצעות קליטה מובנית, בעלות ברורה וקבלת החלטות מותאמת.

מה מעכב את ה ROI של ה AI? אנו מזהים ארבעה פערים אסטרטגיים

4

מסגרת ממשל תאגידי

למרות השקעה משמעותית בבינה מלאכותית, ארגונים רבים מתקשים לקדם את ה Use Cases הנכונים בארגון מה שמוביל לדיווחים לא מדויקים ולהזדמנויות שהוחמצו

3

רגולציה מתפתחת

המסגרת הרגולטורית של בינה מלאכותית משתנה במהירות, ובאזורים רבים, כמו האיחוד האירופי, היא כבר קיימת

2

ניהול סיכונים

כניסה של סיכונים חדשים (תפעולים, דאטה,ציות וסיכונים ברמת הארגון) הקשורים לבינה מלאכותית וכתוצאה מכך נדרש להעריך ולהפחית את הסיכון כדי לאפשר שימוש בינה מלאכותית בצורה רחבה

1

ציפיות לא מותאמות בקרב בעלי העניין

ציפיות סותרות והבנה מוגבלת של בינה מלאכותית בין בעלי העניין מובילים לעיתים קרובות לפתרונות מוגבלים, עלויות גבוהות יותר וקושי לקבוע סדרי עדיפויות לטווח ארוך

A recent survey of enterprise AI leaders found that 80% have 51+ GenAI use cases in the proposal phase—but most have only a handful in production

01 ציפיות לא מותאמות בקרב בעלי העניין: האחריות על AI חוצה פונקציות בארגון

לבינה מלאכותית יש השפעה רב-תחומית ודורשת שיתוף פעולה והשתתפות כדי לספק שקיפות, נראות וביטחון לבעלי עניין

ארגונים מקימים מסגרות ממשל בינה מלאכותית כדי להקצות אחריות ברורה, לייעל את העבודה בקרב קווי הגנה השונים ולהבטיח שהשימוש בבינה מלאכותית תואם את אסטרטגיית הארגון ואת התיאבון לסיכון

אחריות חדשה ספציפית ל AI בשלושה קווי ההגנה

הנהלה ודירקטוריון

- ההנהלה קובעת את "כוכב הצפון" לאתיקת הבינה המלאכותית, קובעת את מדיניות הסיכון של הבינה המלאכותית ועקרונות ממשל תאגידי
- ההנהלה מעלה סוגיות מורכבות לוועדת דירקטוריון בנושא AI ומדווחת התאמה עם אסטרטגיית החברה וציפיות הרגולטורים

קו שני, ציות, משפטי, פרטיות, סייבר

- ניהול סיכונים בונה מדיניות ותקנים לפרקטיקות פיתוח מודלי AI ומתחזקת את מלאי המודלים
- אחראית על התחייבויות ציותיות הקשורות ל-NIST, ISO ורגולציית בינה המלאכותית של האיחוד האירופי, המחייבים מוסדות להקים מתודולוגיה ותהליך לקליטת סיכונים ודירוג בינה מלאכותית
- המדיניות תכלול גם קביעת רמת הסיכון והגורמים הנדרשים לתת הערכות סיכון (למשל, משפטי, סייבר, פרטיות) לסקירה נוספת בהתבסס על מאפייני סיכון ספציפיים

בעלי קו ראשון/בינה מלאכותית

- לזהות צרכים עסקיים, להצדיק מקרי בוחן ולשלב בקרות מאושרות ביישומים.
- מעקב יומיומי, תגובה לאירועים והפחתת סיכונים למודלי הבינה המלאכותית שלהם.

02 ניהול סיכונים במודלי AI - להבין את החשיפה שלכם

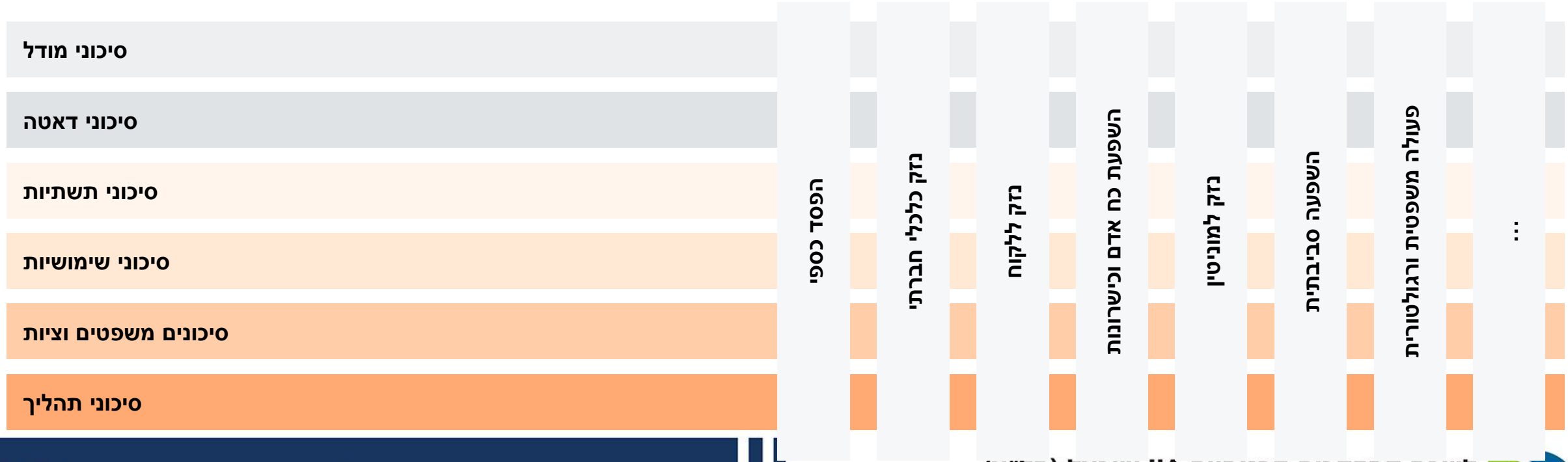
סיכונים יכולים להיווצר ממקורות מגוונים בעת הטמעת בינה מלאכותית. ביצוע הערכת סיכונים ודירוג ברמת השימוש—מיפוי הפרופיל הייחודי של כל יישום (מודל, דאטה, משפטי וכו') כדי להבין את הסיכון הטמון בו וליישם את הממשל, הבקרות, הבדיקות והניטור המתאימים; מה שמאפשר לבינה המלאכותית לפעול עם סיכון שנותר בתוך התיאבון לסיכון של הארגון

השפעת הסיכונים

איזו השפעה שלילית יכולה להיגרם מהסיכונים?

מקורות סיכונים

מה יכול להשתבש במודלי הבינה המלאכותית?



03 רגולציה מתפתחת: ניווט ברגולציה המתפתחת של בינה מלאכותית

הזרקור הרגולטורי על בינה מלאכותית מתגבר, עם ביקורת גוברת מצד רגולטורים אמריקאיים וגלובליים. מבחינות ה-SEC ועד לחששות לגבי מצגי שווא של בינה מלאכותית ושינויים בכללים מדינתיים, פדרליים ובינלאומיים, חברות חייבות להיות מוכנות לנווט באקוסיסטם של ציות שמתפתח במהירות

רגולציה בינלאומית

- צו נשיאותי של ממשל טראמפ מינואר 2025 מדבר על פיתוח תוכנית פעולה ל-AI
- חוק הבינה המלאכותית של האיחוד האירופי מציג לוח זמנים מובנה לאכיפה, החל מ-2025 עם סט התקנות הראשון בנוגע לממשל ושקיפות בבינה מלאכותית
- לפחות תריסר מדינות שוקלות כיום חקיקה דומה המבוססת על סיכון של חוק הבינה המלאכותית של האיחוד האירופי
- ישראל ?

AI-washing

- מקרים אחרונים עם Global Predictions, Inc. (מרץ 2024) ו-Delphia (USA) Inc. (מרץ 2024) מראים את נכונות ה-SEC לנקוט בפעולות אכיפה נגד חברות שמפרסמות הצהרות שקריות או מטעות לגבי יכולות הבינה המלאכותית שלהן
- תיקים אלו חשפו הפרות לכאורה כולל ייצוג שווא של יכולות הבינה המלאכותית, גילוי לא מספק של מגבלות הבינה המלאכותית, וכישלון ביישום ותחזוקה של מדיניות ונהלים כתובים מתאימים

מבחני ה-SEC

- סדרי העדיפויות לשנת 2025 כוללים התמקדות בבינה מלאכותית, ובפרט בשימוש או ייצוג שווא של מוסדות בבינה מלאכותית
- הרגולטורים יעריכו את דיוק הייצוגים ואת פיקוח החברות על שימוש בבינה מלאכותית, כולל בתפעול, מניעת פשעים פיננסיים, פונקציות מסחר ואסטרטגיות השקעה
- הם גם יבחנו את השימוש בבינה מלאכותית המסופקת על ידי צדדים שלישיים וכיצד חברות מונעות אובדן או שימוש לרעה בנתוני לקוחות

04 מהשקעה לאימפקט: גישור פער הביצוע במסגרת הממשל התאגידי

השקעה בבינה מלאכותית

- ההוצאה העולמית על בינה מלאכותית צפויה להכפיל את עצמה ל-631 מיליארד דולר עד 2028, מה שמעיד על מומנטום השקעה חזק
- למרות השקעות כבדות, רוב הארגונים מתקשים להרחיב את הבינה המלאכותית, כאשר מעט מאוד שימושים מגיעים לייצור
- זמני המתנה ארוכים והתקדמות מוגבלת נובעים לעיתים קרובות מהיעדר אוטומציה במחזור חיים של בינה מלאכותית ויישום מסגרת ממשל תאגידי חלש

שיפור התוצאות הללו מתחיל בארבעה שלבים בסיסיים

- מדיניות AI
- מסגרת הפיקוח
- הערכת סיכונים וקטלוג המודלים
- הכשרה ומודעות

למרות השקעה כבדה, רוב הארגונים מתקשים להתמודד עם ממשל תאגידי נאות, כאשר מעט מאוד Use Cases מגיעים לייצור

תהליכי ממשל קיימים אינם יעילים

*רוב הארגונים דורשים **6-18** חודשים כדי להביא פרויקט בינה מלאכותית גנרטיבית לייצור*

מקרי השימוש בבינה מלאכותית תקועים במחזור חיי הפיתוח

***80%** מהארגונים מציעים לפחות 51 מקרים של שימוש בבינה מלאכותית גנרטיבית, אך רובם הפכו לפועל רק קומץ*

זמנים ארוכים מובילים להתקדמות מוגבלת

*72% מהארגונים מחזיקים בפחות **מ-20** מקרי שימוש בבינה מלאכותית בייצור*

מסגרת ממשל תאגידי: ממדיניות לפרקטיקה

כדי להתמודד עם אתגרי התעשייה והרגולציות המשתנות — וליהנות מהיתרונות המלאים של הבינה המלאכותית — חברות נוקטות בארבעת הצעדים הבסיסיים הללו כדי למקם את עצמן להרחבת הבינה המלאכותית באחריות ובביטחון בסביבה משתנה במהירות

1

מדיניות בינה מלאכותית

הגדרה של עקרונות מנחים, שימוש מקובל וציפיות בסיסיות, על מנת לייצר בסיס לאימוץ עקבי ואחראי של בינה מלאכותית ברחבי הארגון

2

מסגרת פיקוח

מבנה ממשל תאגידי מוגדר היטב עם תפקידים ברורים, אחריות ומסגרות מבטיח שיוזמת ה AI תואמת את אסטרטגיית הארגון, סיבולת הסיכונים והתחייבויות רגולציה מתואמות.

3

הכשרה ומודעות

הענקת ידע לעובדים על סיכוני בינה מלאכותית מקדמת תרבות של שימוש אחראי ומבטיחה שמדיניות הממשל מובנת ומיושמת בפועל

4

הערכת וסיכונים וניהול קטלוג מודלים

תיעוד כל מקרי הבוחן בבינה מלאכותית, הן ל-AI שפותח פנימית והן עבור ספקים, וביצוע הערכת הסיכונים שלהם, מאפשרים לארגונים לזהות באופן יזום נזקים פוטנציאליים ולעמוד בציפיות רגולטוריות ופנימיות

ערכת הכלים של PwC לפרקטיקה מובילה ב-AI

ערכת הכלים המובילה של PwC לפרקטיקה RAI עוסקת בבניית אמון. אסטרטגיה וממשל נכונים, המבוססים על עקרונות מנחים מרכזיים ומדעי נתונים טובים, יכולים לסייע ללקוחות לספק באופן עקבי ואחראי **מהאסטרטגיה ועד הביצוע.**

להבטיח איכות באמצעות פרקטיקות מקובלות

תיקוף
הערכו את ביצועי המודל והמשיכו לשפר את האפיון והפיתוח כדי לשפר מדדים.

טיפול בבעיות מהותיות
זיהוי של הבעיה הקונקרטית לה מתאים פתרון והאם היא מצדיקה פתרון בינה מלאכותית/למידת מכונה.

ניטור
הטמיעו ניטור רציף לזיהוי סטיות בביצועים וסיכונים.

סטנדרטים
פעלו לפי תקני התעשייה והנהלים המובילים.

להבטיח תהליכי עבודה אחראיים

אבטחה וסייבר
לשפר את אבטחת הסייבר של המערכות.

פרטיות
פיתוח מערכות ששומרות על פרטיות הנתונים.

שקיפות במידע
אפשר קבלת החלטות שקופה במודל.

קיימות
לצמצם את ההשפעה הסביבתית השלילית.

עמידות
אפשר מערכות ביצועים גבוהות ואמינות.

הטיה והגינות
הגדר ומדוד הוגנות ובדיקת מערכות מול דרישות רגולטוריות.

מסגרת בקרה

ממשל תאגידי
יישום פיקוח על מערכות בשלושת קווי ההגנה.

עמידה בדרישות רגולטוריות
לעמוד ברגולציה, במדיניות ארגונית ובסטנדרטים בתעשייה.

ניהול סיכונים
יישום של תהליכי זיהוי, הערכה והפחתת סיכונים כדי להתמודד עם סיכונים ואירועי כשל ייחודיים לבינה מלאכותית.

בניית אסטרטגיה

אתיקה של נתונים ובינה מלאכותית
צריך לבחון את ההשלכות המוסריות של שימוש בנתונים ובבינה מלאכותית ולהגדיר אותם בערכי הארגון שלך.

מדיניות ורגולציה
בחינה של ציפיות הרגולטורים והמדיניות ציבורית כדי ליישר תהליכי עמידה ברגולציה.

בדיקת מודלי AI

מודלים של בינה מלאכותית נבדקים לעיתים קרובות באמצעות שילוב של תהליכי עבודה ישירים ועקיפים להבנת האיפיון, הנחות וביצועים

מודל	השוואה מול פרקטיקות מידול מובילות הקשורות לאיפיון ומפרט מודל, הצדקה לבחירת מודל, שיטות השוואת ביצועים וגישת יישום בחברה.
דאטה	השוואה מול פרקטיקות דאטה מובילות הקשורות לנתונים המשמשים למטרות בינה מלאכותית, כולל גישה להנדסת תכונות, שלמות, סבירות והתאמה של נתוני הטסטים, סבירות הקשרים בין פלטי המודל לנתוני הטסטים.
תשתית **בינה מלאכותית**	השוואה מול פרקטיקות תשתיות מובילות בבינה מלאכותית, כולל יציבות חיזוי, השהיה, זמינות ושיעורי גיבוי או כישלון של המודל.
מוכנות **לבדיקות**	השוואה מול פרקטיקות ותהליכים מובילים לבדיקות בינה מלאכותית הקשורים להרכב צוות, הגדרת קריטריוני בדיקות מבוססי סיכונים, כלי בדיקה ופיתוח תוכניות בדיקה.

יכולות הערכה רחבות יותר

הכנה לאישורי רו"ח

מוכנות לאישור נפרד של מערכות ומודלים של בינה מלאכותית.

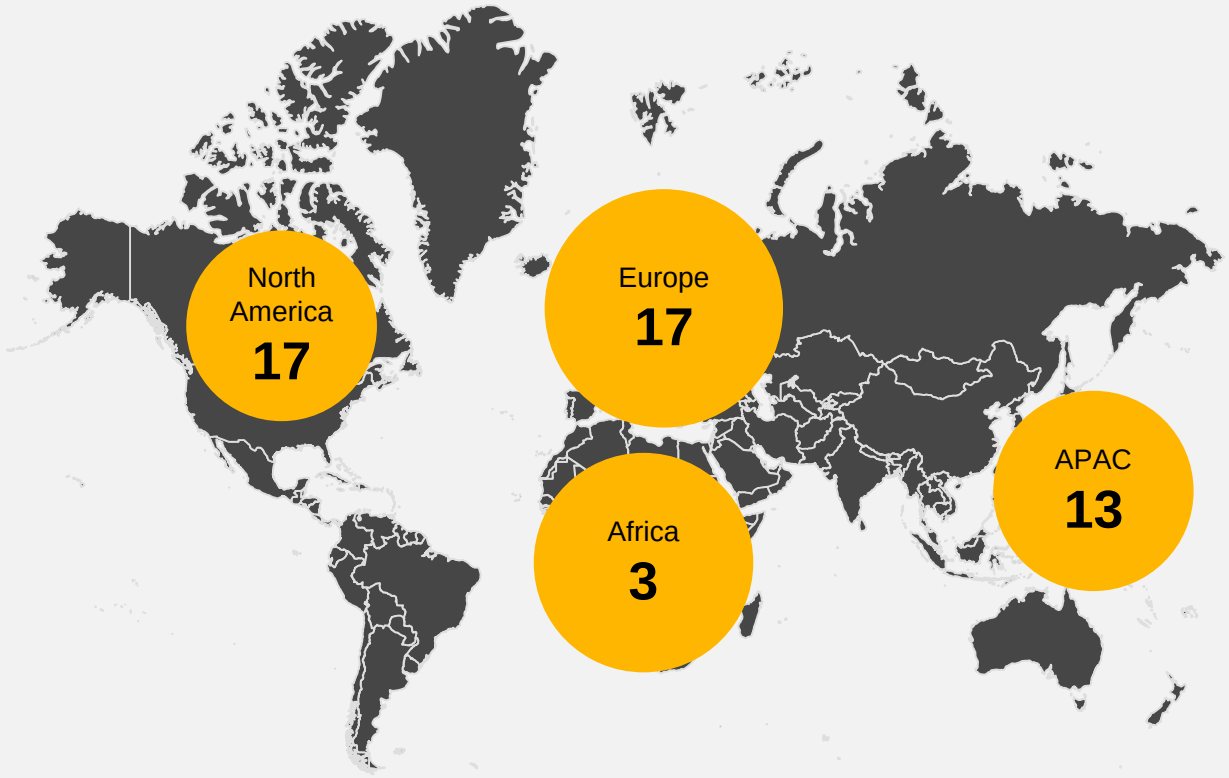
התאמה רגולטורית

הערכת פתרונות AI ו-AI גנרטיבי להתאמה עם פרקטיקות מובילות בתעשייה, דרישות רגולטוריות ותקני רגולציה.

שירותים נוספים

הערכת בקרות ההנהלה על איכות הנתונים, מקור ועיבוד. הערכת יעילות התכנון והתפעול של תפעול ובקרות IT הרלוונטיות ליישום מערכות בינה מלאכותית.

תובנות מסקר PwC בנושא ניהול סיכונים וביקורת פנים לשנת 2025



Interviewed a senior Risk representative (CRO, Head of ERM / CAO, COO of Risk) across 6 focus areas

Collected data via a standard survey questionnaire, informing an analytical dashboard

Identified priorities, trends and change agenda across the industry

רקע

המחקר הינו רב שנתי החל משנת 2018

החל משנת 2022 אנו רואים את הגידול בטרנספומציה דיגיטלית

השנה אנחנו רואים את הקפיצה הגדולה באימוץ טכנולוגיה מתקדמת

50
Stakeholder interviews across global and regional banks

15
GSIBs, 29 large internationally active banks amongst participant set

22
Banks participated in a Risk Survey feeding into our analytical dashboard

המסקנות העיקריות

- כקו ההגנה השלישי, הביקורת הפנימית צריכה להתמקד בהבנת מטרות החברה והשימושים שלה, כמו גם את תפיסת מנהל הסיכונים לגבי סיכונים שקיימים ותוכניות הפחתה.
- בנו ביקורת ייעודית למערכי הליבה - הנתונים - המשמשים לאימון, כיול והפעלת המערכת והמודלים.
- בנו תוכנית ביקורת ייעודית סביב המערכות והמודלים של AI.
- בנו תוכנית ביקורת ייעודית עבור הפלט של מודלי ה AI.

כיצד לבקר את הסיכון – דוגמאות

תחום	מה הביקורת צריכה לעשות בפועל	השלכות על הממשל הארגוני
מעבר לניהול סיכוני מודל (MRM) מסורתי	זיהוי פערים במסגרות הבקרה הקיימות על מודלים.	דחיפה להתפתחות של שכבות הבטחה (assurance) חדשות, ייעודיות לבינה מלאכותית.
ממשל ריכוזי של מקרי שימוש	בחינת מנגנוני ממשל לבחירה, פיתוח ויישום של מקרי שימוש כלל-ארגוניים; הבהרת המקרים בהם נדרש תיקוף (ולידציה) של המודל.	תמיכה בפיקוח על הטמעות בעלות השפעה גבוהה; הגדרת רמת ההבטחה הנדרשת.
ממשל רב-תחומי	השתתפות או השקפה על פורומי פיקוח משותפים על AI המשלבים נציגים מניהול סיכוני מודל, משפט, ציות, טכנולוגיה ודאטה, ופרטיות.	אספקת מבט מתואם על סיכונים מורכבים.
פיקוח על AI של צד שלישי	בחינת פתרונות AI המוטמע בפתרונות של ספקים ומיפוי תלויות.	מתן מענה לסיכונים הנמצאים מחוץ לשליטתה הישירה של החברה; שיפור השקיפות לגבי חשיפות סמויות.
התמודדות עם חשש מהסתמכות-יתר	הכשרת הצוות לשמירה על חשיבה ביקורתית ויכולת לאתגר.	מניעת "האצלת חשיבה" (cognitive offloading) והסתמכות עיוורת על תוצרי ה-AI.
חדשנות בתחום התיקוף (ולידציה)	בחינת מדדי פרוקסי, מודלים מאתגרים (challenge models) ומנגנוני ביקורת.	התקדמות לקראת יכולת פירוש (interpretability) ועקיבות (traceability).

לסיכום

ביקורת פנימית אמורה להיות חלק מרכזי בממשל תאגידי של מודלי AI כדי לאשר שמערכות AI תוכננו ויושמו בהתאם למטרות החברה.

אבל כדי ליצור תוכנית ביקורת מבוססת סיכונים ספציפית ל- AI יוצרת, הביקורת הפנימית חייבת לתכנן ולאמץ מתודולוגיות ביקורת חדשות, צורות חדשות של ביקורת וסט של מיומנויות חדשות. נהלי הביקורת צריכים לכלול אלמנטים של ביקורת ממשל תאגידי וביקורת טכנולוגיה כאחד.

נהלי הביקורת הקיימים יצטרכו להתעדכן על ידי ארבעה מאפיינים :

- גנרטיביות - פתיחות היישומים שלהם.
- ברבור שחור -הופעה פתאומית ובלתי צפויה של התנהגויות חדשות.
- חוסר במידע בסיסי - מידע שאינו קיים או לא מתכתב עם העולם האמיתי.
- העובדה שהמודלים נגישים רק דרך ממשקי אפליקציה ושפת תכנות יעודית.

תודה

ערן רז
שותף, PwC ישראל
eran.raz@pwc.com
נייד: 054-6660260