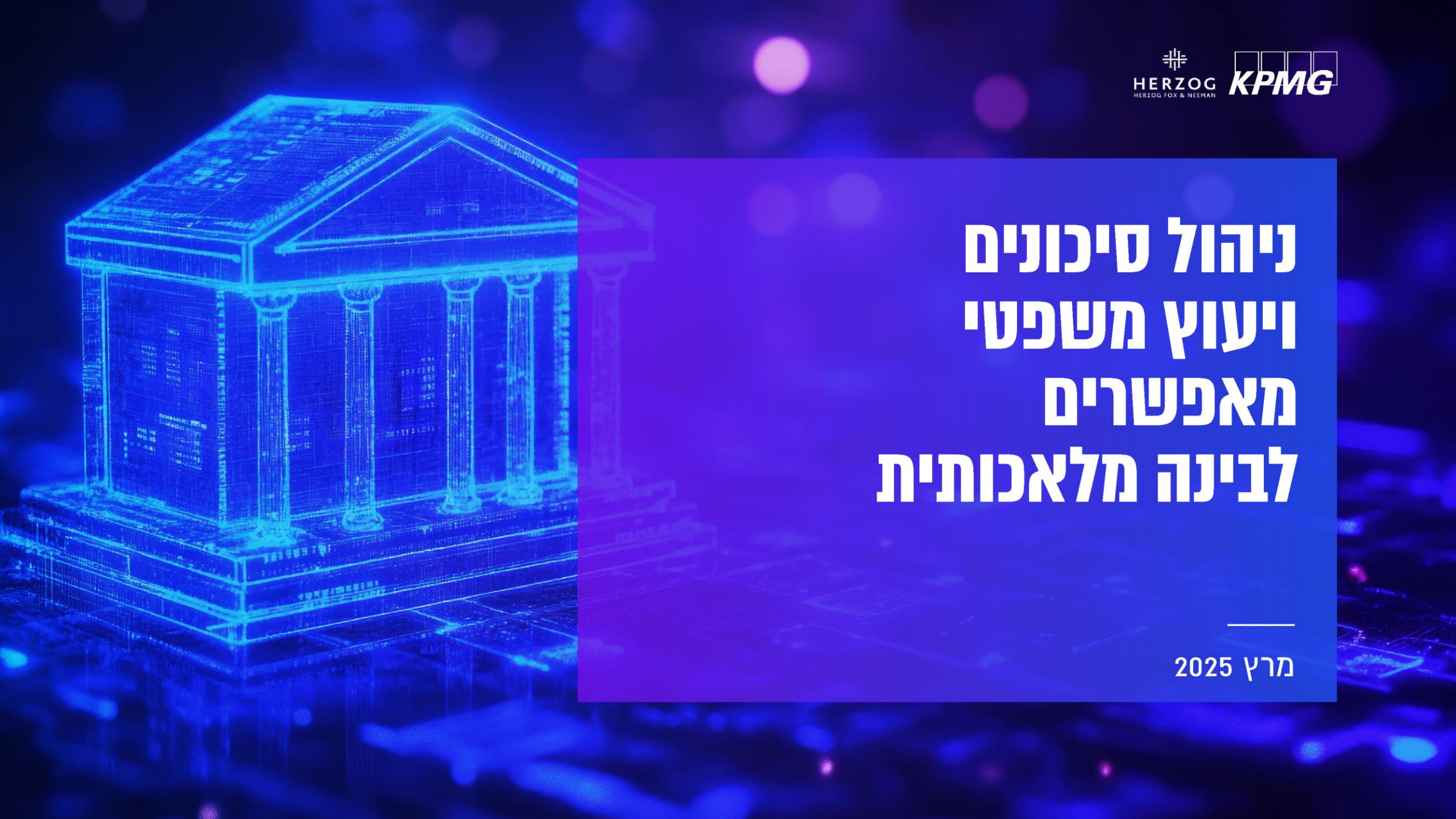


ניהול סיכונים ויעוץ משפטי מאפשרים לבינה מלאכותית

מרץ 2025

סיכוני בינה מלאכותית במוסדות פיננסים

ההשפעות של הבינה מלאכותית יוצרת Generative AI עבור הסקטור הפיננסי הינן עצומות.

היכולת להשתמש במודלים יוצרים אשר עשויים להחליף ו/או לייעל פעילות אנושית מאפשרת לבנקים לספק שירותים חדשניים ללקוחותיהם, לייעל תהליכים תפעוליים ולשפר את אפקטיביות ניהול הסיכונים.

"צריך לאזן בין ההבטחה העצומה שגלומה בבינה המלאכותית לבין הסיכונים המאוד רציניים שכרוכים בה. תצטרך להיות רגולציה על התחום כדי שלא תתפתח סופר-אינטליגנציה בלתי מרוסנת."

Sam Altman

CEO OpenAI

HERZOG
HERZOG FOX & NEEMAN

KPMG

01

האתגרים והתועלת ביישום מערכות AI

אתגרים ייחודיים למוסדות פיננסיים

- חובות עפ"י דין ופסיקה
- דרישות רגולטוריות
- הוגנות (Conduct)
- ניהול סיכונים

התועלת עבור המוסד הפיננסי והלקוחות

מרבית היתרונות המובהקים שבהטמעת מערכות בינה מלאכותית במוסדות פיננסיים מתבססים לא רק על **חיסכון במשאבים אנושיים** אלא גם **במתן כלים לביצוע פעולות שהגורם האנושי בארגון מתקשה לבצע בשגרה.**

היתרונות עבור הלקוחות מתבטאים **בשיפור חוויית ההתקשרות** מול המוסד הפיננסי יחד עם **חיסכון בזמן, טיפול בתקלות, קבלת תובנות מותאמות אישית ועוד.**

מוסד פיננסי

- **חסכון בעלויות וכוח אדם** ויצירת Scalability.
- **מניעת טעויות** כגון בהוראות להעברת כספים.
- **טיפול בתהליכים הדורשים ניתוח מידע רב** מניעת הונאות, ניטור פעילות, ניטור שוק ועוד.
- **טיפול בתהליכים מורכבים** כגון תשלומים חוצי גבולות.
- **שיפור חווית לקוח.**

לקוחות

- **קבלת החלטות ביחס ללקוח** ללא הטיה אנושית.
- **יכולת לתקשר מול הבנק** באמצעים דיגיטליים נגישים.
- **סיוע בתהליכים** כגון ביצוע תשלומים ותשלום חשבונות.
- **קבלת תובנות** למשל, בעולמות התשלומים (ניהול הוצאות, ניתוח תשלומים ועסקאות).
- **ייעוץ פיננסי** (היכולת לשלב גם תובנות מהתנהלותו הפיננסית של הלקוח) ועד לחינוך פיננסי.
- **זמינות מתמשכת** שלא בשעות פעילות נציגי הבנק.

יישום ארגוני של מערכות בינה מלאכותית

הטמעת מערכות בינה מלאכותית מעוררת סוגיות רבות לטיפול הקשורות זו בזו בקשר הדוק.

לשם יישום נאות יש צורך בשיתוף פעולה בין הגורמים השונים ובכלל זה הגורמים הטכנולוגיים, מומחי התהליכים והגורמים המשפטיים, כדי להבטיח שתהליכי ההטמעה בארגון עומדים בדרישות הארגון ותאבון הסיכון שלו, בממשל תאגידי נאות, וכמובן בהוראות הדין.

מסגרת היישום

- יצירת ממשל תאגידי נאות לניהול תהליכי יישום בינה מלאכותית ונהלים מתאימים.
- התאמת האסטרטגיה לדרישות הדין והרגולציה הרלבנטית המשפיעות על הפעילות (בארץ ו/או בחו"ל, לפי העניין).
- טיפול בסוגיות אבטחת מידע, פרטיות, הגנה על קניין רוחני.
- יצירת מערכת חוזית מתאימה תוך התייחסות לשיקולים הרלבנטיים בהסכמים.

שיטת היישום

הקמת מסגרת כוללת

הקמת מסגרת כוללת מ- Day One תאפשר Scalability יחד עם מימוש החזון הארגוני לשימוש במערכות בינה מלאכותית ותקל על תהליכי ניהול סיכונים וציות אך גישה זו תלווה בצורך להשקיע משאבים וזמן לצרכי פיתוח וקסטומיזציה.

טיפול אד הוק

יאפשר יישום והטמעה של פתרונות בצורה גמישה ומהירה יחסית ללא השקעת משאבים מאסיבית אך גישה זו תקשה על שיפור יכולת ה- Scalability ועלולה ליצור הטמעת פתרונות בצורה לא עקבית ואחידה ברחבי הארגון.

HERZOG
HERZOG FOX & NEEMAN

KPMG

02

עקרונות לשימוש אחראי במערכות בינה מלאכותית

עקרונות לשימוש אחראי במערכות בינה מלאכותית

איכות המידע | Data Integrity

הגדרת מנגנונים שיבטיחו כי המידע שמופק מהמערכת (Output) והנתונים שמאמנים את המודל (Input) הינם נכונים ברמה סבירה של בטחון.

הוגנות

שמירה על מודל הוגן וערכי שתוצאותיו אינם חורגים מנורמות מקובלות וממדיניות הארגון לרבות היבטים של אפליה והוגנות מול לקוחות.

יכולת הסבר | Explainability

שמירה על היכולת לתעד ולהבין את תהליך קבלת ההחלטות של מערכת הבינה המלאכותית.

אחריות | Accountability

שמירה על מנגנונים המובילים ללקיחת אחריות מלאה על השלכות מוצרי ה-AI.
הגדרת גורמים אנושיים בחברה שהם יהיו בעלי האחריות להשפעות המערכת.

עקרונות לשימוש אחראי במערכות בינה מלאכותית

Reliability

שמירה על אמינות המערכת ובקרה על כך שהתוצאות עומדות בסטנדרטים שהוגדרו מראש לאורך זמן.
הגדרת סט בקרות שוטפות לניטור וזיהוי חריגים בזמן אמת.

אבטחת מידע והגנת סייבר

שמירה על אבטחת המערכת מפני תקיפה או מפני מתן גישה לגורמים לא רצויים. ככל ומדובר במיקור חוץ, יש לוודא כי לא מועבר מידע לא נדרש וכי רמת אבטחת המידע של הספק מספקת ביחס למידע המועבר.

הגנת הפרטיות

שמירה על ציות לחוקי הפרטיות בכל נושאי שמירת המידע האישי של לקוחות והשימוש בו.
ככל ומבוצע שימוש בספק ענן, יש לתת את הדעת לעניין מיקום הספק ביחס לחוק המקומי והגלובלי.

Safety

יישום נהלים שימנעו ממערכת הבינה המלאכותית להשפיע על הסביבה, הנכסים או העובדים בחברה.

אסדרה בתחום הבינה מלאכותית

HERZOG FOX & NEEMAN

KPMG

European Commission

חוק הבינה המלאכותית של האיחוד האירופי (The AI Act) נכנס לתוקף ב-1 באוגוסט 2024. עם זאת, יישום הוראות החוק מתבצע באופן הדרגתי.

החוק מקטלג את השימושים הפוטנציאלים בטכנולוגיות בינה מלאכותית לפי שלוש רמות סיכון:

https://artificialintelligenceact.eu

רמות סיכון עפ"י ה- AI ACT

מערכות בינה מלאכותית המעלות סיכון בלתי מתקבל
מערכות המציבות סיכון מהותי לבטיחות, פרנסה או שעלולות לפגוע בחוקי יסוד של אזרחים כגון הזכות לפרטיות.

מערכות בינה מלאכותית בסיכון גבוה
מערכות שבכוחן להשפיע על תשתיות קריטיות, חינוך, תוצאות בחירות ועוד, על מערכות אלו יהיו הגבלות ורגולציה.

מערכות בינה מלאכותית בסיכון נמוך
שלא נאסרו לשימוש או שהוגדרו במערכות בסיכון גבוה, אליהן כרגע לא הופנו מגבלות רגולטורית ו\או שתהיה עליהן מגבלה רגולטורית.

שימושים בגופים פיננסיים

שימוש בבינה מלאכותית לצורך דירוג אשראי של לקוח נכלל תחת הקטגוריה המסוכנת של High-Risk.
מערכות אלה מותרות לשימוש אך יסווגו כמערכות בסיכון גבוה אשר יעמדו תחת דרישות רישום מראש, דיווח, הגדרת השימוש המדויק של המערכת, האפיון שלה, צורת הפיקוח האנושי על התוצרים שלה ועוד (ניתן לראות בהרחבה את דרישות הדיווח המינימליות בנספח 4 לחוק).

הגדרת מערכת AI בהתאם לחוק
"מערכת מבוססת מכונה שתוכננה לפעול ברמות שונות של אוטונומיה ויכולה, למטרות מפורשות או מרומזות, לייצר תפוקות כגון תחזיות, המלצות או החלטות המשפיעות על סביבות פיזיות או וירטואליות".

KPMG

אסדרה בתחום הבינה מלאכותית

שימושים בבינה מלאכותית שיוגדרו תחת קטגוריית High-Risk יידרשו בעיקר לנושאים הבאים:

- **חובת תיעוד אופן יצירת הטכנולוגיה טרם שיווקה**
- **חובות בנוגע למידע שמשמש לאימון המודלים**
- **חובת שמירת מסמכים המאפשרים פיקוח ובקרה על הטכנולוגיה**
- **חובות בנוגע לאפשרות להסביר את תוצאות המודל (Explainability)**
- **נדרשת רמת דיוק נאותה**
  בקרות לבחינת דיוק תוצאות המודל
- **ניהול מאגרי המידע אליהם למערכת יש גישה על מנת לוודא שאינה פועלת באופן מוטה ועל מידע שלם**
- **תיעוד ארכיטקטורת המערכת והלוגיקה שבבסיסה**
- **דיווח לגורמים רשמיים מקומיים בגין כל אירוע חריג שיתכן והתרחש בעקבות שימוש במערכת**
- **נדרש לפתח יכולות למניעה של שימוש בלתי-מורשה וניצול של חולשות במודלים**
  לדוגמא, תוקף שיודע כיצד המודל ועל ובכך יגרום מצב של הערכת יתר של דירוג האשראי שלו
- **תהליכי אישור מראש של רשות חדשה שתוקם בעקבות החוק**

אסדרה בתחום הבינה מלאכותית

הרגולטורים בישראל פועלים לאסדרה של התחום

בנובמבר 2024, הצוות הבין-משרדי לבחינת שימושי בינה מלאכותית במגזר הפיננסי, המורכב מנציגי משרד המשפטים, משרד האוצר, רשות ניירות ערך, רשות שוק ההון, הפיקוח על הבנקים ורשות התחרות, פרסם דוח ביניים להערות הציבור.
הדוח סוקר את הפוטנציאל, האתגרים והסיכונים הכרוכים בשימוש בטכנולוגיית בינה מלאכותית בסקטור הפיננסי, ומציע המלצות ראשוניות לעיצוב הרגולציה בתחום זה.

מדובר במסמך שצפוי להיות בסיס לאסדרה עתידית של הרגולטורים הפיננסיים.

עקרונות המדיניות העיקריים המוצעים במסמך

- **שימוש בכלי רגולציה "רכה"** במקום חקיקת מסגרת רוחבית.
- **אימוץ עקרונות אתיים** בדומה למקובל בעולם.
- **הסברתיות** הצורך בשקיפות והבנה של תהליכי קבלת ההחלטות במערכות בינה מלאכותית במיוחד בהחלטות בעלות השפעה משמעותית על הלקוח.
- **מעורבות אנושית** הבטחת פיקוח ובקרה אנושיים על החלטות המתקבלות באמצעות בינה מלאכותית במיוחד במקרים שבהם ההשפעה על הפרט מהותית.

רשות התחרות
רשות שוק ההון, ביטוח וחיסכון
בנק ישראל BANK OF ISRAEL
רשות ניירות ערך
Israel Securities Authority
سلطة الأوراق المالية
שרת החדשנות, המדע והטכנולוגיה

HERZOG
HERZOG FOX & NEEMAN

KPMG

03

סיכונים עיקריים

סיכוני בינה מלאכותית | חוסר הוגנות, פגיעה במוניטין וחריגה מרגולציה

להלן מספר דוגמאות לאירועי כשל שאירעו בעקבות מודל בינה מלאכותית מוטה (Biased) שכלל אפליה ו/או התנהגות לא הולמת

אמזון מערכת לגיוס עובדים

בשנת 2018 השתמשה אמזון במערכת בינה מלאכותית לצורך סריקת קורות חיים ומיון מועמדים רלוונטיים. המערכת נמצאה מפלה נגד נשים. אמזון הפסיקו את השימוש במערכת מיידית.

Apple Card

ב-2019 השיקו אפל וגולדמן זאקס כרטיס אשראי במיזם משותף. החברות התהדרו ביכולתן להעריך את סיכון האשראי של הלקוח באופן דיגיטלי בזמן אמת באמצעות אלגוריתם של בינה מלאכותית. נגד החברות התנהלה חקירה של הרשויות על אפליית נשים של המודל והמיזם ספק הרבה ביקורת ציבורית ופגיעה במוניטין.

Microsoft's Tay Chatbot

מיקרוסופט אימנו צ'אט-בוט על המידע הקיים בטוויטר ושחררו את הצ'אט-בוט לאויר. לאחר 24 שעות הוחלט להשבית את המיזם לאור תגובות בלתי הולמות גזעניות וסקסיסטיות שיצר המודל.
תגובות המודל רק שיקפו את הנתונים עליהם הוא התאמן, אך כמובן שמיקרוסופט התנצלה והודתה בטעות.

מערכת לזיהוי פנים של IBM

בעקבות ביקורת לגבי יעילות המערכת על אנשים בצבע עור כהה, החליטה IBM לנטוש את המערכת לזיהוי פנים שבנתה במשך שנים. מחקרים הוכיחו כי המערכת עובדת בכ-30% פחות יעילות בזיהוי פנים של אנשים שחורים, והיעילות פוחתת עוד יותר כאשר מדובר באישה.

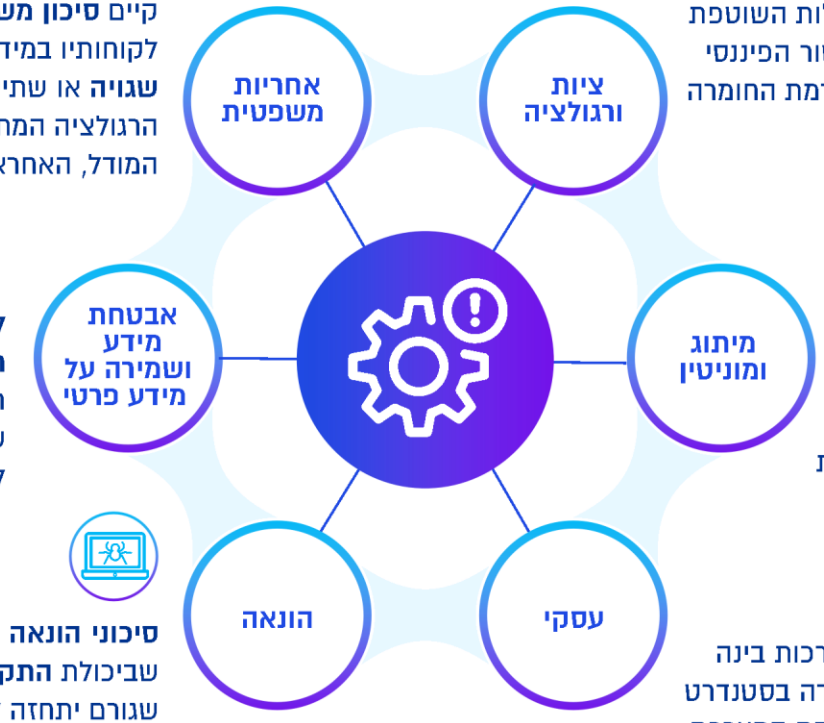
ההטיה במודל ככל הנראה נובעת ממאגרי המידע עליהם אומן המודל וחוסר הייצוג של אוכלוסיות מסוימות במאגרים אלו.

אפליה באלגוריתם של Uber

כנגד ענקית התחבורה השיתופית Uber קיימות מספר חקירות לגבי נאותות האלגוריתמים של החברה בנושאים הבאים:

- תשלום פרסונלי שונה לנהגים בהתבסס על קריטריונים לא אתים כגון מגדר וגזע.
- התארכות זמני המתנה או פערים במחירים ללקוחות על בסיס גזע או מגדר.

סיכונים עיקריים | שימוש ב-AI במוסדות פיננסיים



שמירה על כללי ציות ורגולציה הם נדבך קריטי בפעילות השוטפת של כל מוסד פיננסי. נכון להיום קיימת **אי ודאות** בסקטור הפיננסי לגבי עתיד הרגולציה על מערכות בינה מלאכותית ועל רמת החומרה הפוטנציאלית שברגולציה לכשתפורסם.

אמון ציבור הלקוחות בבנק היא אבן פינה שבלעדיה אף מוסד פיננסי לא יכול להתקיים. ללא פיקוח, ניטור ובקרה אחר תוצאות מודל ה-LLM, קיים סיכון פגיעה במוניטין ככל והמודל יפעל בניגוד למסגרת שהוגדרה לו. סיכון זה מקבל משנה תוקף נוכח מקרים שארעו לאחרונה בהם צ'אט מבוסס LLM החזיר תשובות פוגעניות ו/או המנוגדות למדיניות שהוגדרה.

סיכונים עסקיים שעלולים להתממש תוך שימוש במערכות בינה מלאכותית הם הסיכון ש**באובדן לקוחות** עקב אי עמידה בסטנדרט השירות המצופה, סיכונים תפעוליים הקשורים בהשבתת המערכת במידה של תקלה או מתקפת סייבר, אובדן יעילות שבממשק העבודה מול הלקוחות במידה ויידרש גיבוי כוח אדם שלא יסופק ע"י הבנק.

קיים **סיכון משמעותי** בכל הקשור לאחריות משפטית של הבנק אל מול לקוחותיו במידה שמערכת הבינה המלאכותית תייצר **המלצה פיננסית שגויה** או שתייצר **המלצה שתוליך שולל או אף תפלה לרעה את הלקוח**. הרגולציה המתגבשת בתחום קובעת כי על אף שההמלצה ניתנת על ידי המודל, האחראי הבלעדי לנזק הינו הבנק.

ליקויים הקשורים **באבטחת מידע ושמירה על נתונים רגישים של לקוחות** הם הסיכון העיקרי. הסיכון לדלף מידע רגיש (בגין לקוחות או מסחרי של הבנק) גדל כאשר מבוצע שימוש במערכות של ספק צד שלישי אשר מקבל נתונים לגבי לקוחות הבנק.

סיכוני הונאה שבשימוש במערכת בינה מלאכותית כוללים את הסיכון שביכולת **התקפת סייבר** תושג שליטה על המערכת ודרכה יודלף מידע, או שגורם יתחזה למערכת על מנת לנסות לקבל מידע רגיש מלקוחות ועד מצב בו עובד מלמד בזדון את המערכת לגשת ולהדליף למידע שאמור להיות חסום בפניה. בנוסף מתקפות כמו "Adversarial attacks" ו-"Jailbreaking" מבוססות על זיהוי וניצול חולשות המודל על ידי התוקף.

שימוש בספקי מערכות AI במוסדות פיננסים

ביצוע בדיקות נאותות מקיפות הכוללות ביקורות אבטחת מידע יחד עם עיגון חובות הספק ברמה החוזיות מסייעות בהפחתת סיכונים הן בעולמות הציות, הן בעולמות אבטחת המידע ותפעול המערכות השותף.

בנוסף, כפי שניתן לראות בחוק האירופי בנושא מערכות בינה מלאכותית (The AI Act), ישנה מגמה בה הרגולטור מחיל חובות **על ספקי מערכות הבינה המלאכותית עצמם** ולא רק על המשתמשים בה. כך שבעתיד ניתן יהיה לצפות לנדבך נוסף בבחינת הספקים שיצטרכו להוכיח את העמידה ברגולציה שתחול עליהם.

Microsoft Copilot

OpenAI

Google AI

שימוש בספקים חיצוניים לצורך הטמעת מערכות בינה מלאכותית בארגון דורשת בדיקת נאותות מעמיקה על הספקים הפוטנציאליים כאשר בחינה זו חיונית לשמירה על אבטחת הארגון, עמידה ברגולציות והאמינות תפעולית.

כתמונת מראה לרגולציה החלה על המוסדות פיננסיים על הספקים חלה האחריות להתאים את המוצרים שלהם לדרישות רגולטוריות ודרישות מרכזיות נוספות של המוסדות הפיננסיים כגון: שקיפות האלגוריתמים (Transparency), הפחתת הטיות (Bias Mitigation) והתאמה למערכות קיימות.

HERZOG
HERZOG FOX & NEEMAN

KPMG

04

נספח
הרחבה על סיכונים ודוגמאות לשימושים נוספים בגופים פיננסים

אסדרה נוכחית רלוונטית לבנקים

אמנם, כאמור, טרם הוגדרה רגולציה ייעודית לתחום הבינה המלאכותית המתפתח בבנקים. יחד עם זאת הרגולציה הנוכחית החלה על הבנקים עשויה לספק מסגרת מסוימת לניהול הסיכון. **להלן רגולציה אשר עשויה להיות רלוונטית עבור יישומים שונים בתחום ה-AI:**

שם ההוראה	נושא	רלוונטיות לשימושי בינה מלאכותית
הוראת ניהול בנקאי תקין A359	מיקור חוץ	חלק ניכר מיישומי מערכות הבינה המלאכותית מבוצעים באמצעות כלים של חברות חיצוניות אשר עונות להגדרת מיקור חוץ, על כן יש לבחון יישום של הבקרות והמגבלות כפי שהן שמפורטות ההוראה.
הוראת ניהול בנקאי תקין 361	ניהול הגנת הסייבר	הטמעת מערכת בינה מלאכותית בארגון במידה ותעשה ע"י ספק חיצוני עלולה לחשוף את הארגון לסיכוני סייבר עקב יצירת הממשק בין המערכת החיצונית למערכות הבנק ועקב הצורך בהעברת המידע.
הוראת ניהול בנקאי תקין 362	מחשוב ענן	מערכות בינה מלאכותית מתקדמות כדוגמת GPT של חברת OpenAI הינם מוצרי "ענן" ושימוש במערכות אלו מצריך הוצאת מידע מחצרות הבנק לסביבת הענן של הספק.
הוראת ניהול בנקאי תקין 363	ניהול סיכוני סייבר בשרשרת אספקה	במידה והארגון משתמש בשירותי מערכות בינה מלאכותית המסופקים ע"י חברה חיצונית יש לבחון כי התקשות זו אינה חושפת את הארגון לסיכוני אבטחת סייבר חדשים.
הוראת ניהול בנקאי תקין 310	ניהול סיכונים	שימוש לראשונה במוצר מבוסס בינה מלאכותית יצריך, בדרך כלל, אישור מוצר חדש בהתאם לסעיף 16 לנב"ת.
הוראת ניהול בנקאי תקין 350	ניהול סיכון תפעולי	על פי רוב, הטמעת מערכות חדשניות מורכבות כגון מערכות בינה מלאכותית בארגון תדרוש מיסוד נהלים, תפקידים, אחריות וניהול סיכונים מרמת הדירקטוריון מטה עפ"י העקרונות המנחים שבהוראה.
תקנות הגנת הפרטיות העברת מידע אל מאגרי מידע שמחוץ לגבולות ישראל	שמירה על פרטיות	כיוון שספקי מערכות בינה מלאכותית רבים הינן חברות בינלאומיות אשר מחזיקות את מארגי המידע שלהן במדינות אחרות יש לוודא כי המידע עובר ונשמר בצורה תקינה עפ"י ההוראה.
מכתב פיקוח על הבנקים בינה מלאכותית במודלים בתחום איסור הלבנת הון	מודלים וכלים טכנולוגיים בתחום איסור הלבנת הון ומימון טרור – דגשים לניהול סיכוני מודל, שימוש וחדשנות	חלק מהשימושים בבינה מלאכותית מבוססים על מודלים של למידת מכונה (ML) אשר עשויים לענות להגדרת "מודל" לעניין נב"ת 310 ועל כן יושטו עליהם ההנחיות של הפיקוח על הבנקים לניהול סיכוני מודל (תיקוף המודל, בחינת ההנחות, תיעוד ועוד).
הוראת ניהול בנקאי תקין 369	ניהול סיכוני מודל	האמור בהוראה חל גם על מודלים הכוללים או המתבססים על בינה מלאכותית. הוראה זו מתארת את ההיבטים העיקריים של ניהול אפקטיבי של סיכוני מודלים לרבות מנגנוני ממשל תאגידי ובקרה, כגון פיקוח מצד הדירקטוריון וההנהלה הבכירה, מדיניות ונהלים, בקרות וציות וכן תמריצים ומבנה ארגוני מתאימים.

סיכוני ציות ורגולציה

סיכון	פירוט הסיכון
דלף מידע רגיש	בשימוש ישיר מול הלקוח- היעדר בקרות על הוצאת מידע רגיש על ידי הלקוח - הלקוח כותב טקסט חופשי וטקסט זה מועבר למודל ה-LLM.
מתן מידע שגוי	ככל והמודל יפיק מידע שגוי הבנק עשוי להיות חשוף לסיכוני Conduct ולחוסר הוגנות מול לקוחותיו. סיכון זה מקבל משנה תוקף כאשר המודל יספק מידע שגוי שעל פיו הלקוח יבצע החלטה פיננסית.
מענה שגוי בנושאים רגולטוריים	הסיכון שלקוח יבקש ממערכת הבינה המלאכותית ו/או מבנקאי המסתמך על מערכת זו לספק פירוט על הוראה או מגבלה רגולטורית כלשהי וזו תוצג ללקוח בצורה שגויה. לדוגמה, "מהי הריבית המתואמת" של ההלוואה? או "מהי מסגרת זמנית"?
אפיון האחריות הספציפית	ישנו קושי באפיון תפקידה של ישות ספציפית שתחת אחריותה יהיו תחומי הפיקוח והתפעול של מערכות הבינה המלאכותית בארגון לרבות בהיבטים שאינם טכניים, דהיינו פוטנציאל נזק שהמערכת תגרום לארגון או ללקוחותיו. בנוסף, כיוון שהמערכת נועדה לעבוד בצורה עצמאית אל מול מערכות הארגון, לרוב ע"י שימוש בספק חיצוני, קים קושי באפיון גורם ספציפי שעליו תחול האחריות במידה והמערכת תפעל בצורה שאינה חוקית או שתגרום נזק לחברה או ללקוח. ביזור האחריות יוצר מצב בו אין גורם מתומרץ שישמור על תקינות המערכת.
אי יכולת הפקת מסמכי תיעוד ומתן הסברים לגבי המודל (XAI)	הסיכון שהארגון לא יוכל להתחקות אחר ה"קופסא השחורה" ולספק מידע ללקוח ו\או לרגולטור על הסיבות שהובילו לתוצרי המודל. לדוגמה תשובה לא הולמת ללקוח, חישוב לא נכון של פרמטרים דינאמיים ועוד.
מתן תשובות מבוססת דעה קדומה ו/או אפליה	הסיכון שהמודל יפעל בהתבסס על דעות קדומות ומפלות כנגד מגזר או פלח אוכלוסייה. למרות שהמודל לא יקבל מידע אישי על הלקוח עדיין תתכן אפשרות שבשאלת הלקוח עשוי להיות תוכן רגיש שעשוי לגרור הסקת מסקנות שאינה הולמת מהמודל.
פעילות תחת אי ודאות רגולטורית	פעילות הקשורה לענף טכנולוגי חדשני טרם פרסום רגולציה ברורה ומקיפה תמיד תכלול רמת סיכון מינימלי כלשהו. קיים סיכון שפעילות בתחום תהווה הפרה של הדין הקיים אשר יקבל פרשנות בדיעבד (לדוגמא, הגנת הפרטיות).

סיכונים משפטיים

סיכון	פירוט הסיכון
עדכון תנאי שימוש	קיימת חשיפה משפטית שהארגון יבצע שימוש במערכות מבוססות מודלי LLM שאינו מוגדר בהסכם התקשרות מולם ואשר חורג מהתנאים המותרים בהסכם ו\או ממדיניות הספק.
עלויות משפטיות	קיים סיכון בעליית תשלומי פיצויים או שכר טרחה מקצועי עבור סיכונים שהארגון לא היה חשוף אליהם ערב השקת המערכת.
זכויות יוצרים	קיים סיכון משפטי בהקשרי זכויות יוצרים בכל הקשור לחומרים שנוצרו ע"י מערכות בינה מלאכותית יוצרת. הבנק עשוי לערוך שימוש במודלים אשר אומנו על מידע אשר לגביו קיימות זכויות יוצרים ועל כן קיימת חשיפה משפטית לגבי שימוש בתוצרי המודל.

סיכוני אבטחת מידע וסייבר

סיכון	פירוט הסיכון
חשיפת עצמאית של מידע לקוחות	בשימוש של מודל חיצוני ישירות מול הלקוח- הסיכון שלקוח יבחר להכניס לשאילתה\לצ'אט במערכת מבוססת בינה מלאכותית מידע רגיש שיעבור אל הספק החיצוני.
חשיפת מידע של לקוחות ע"י עובדי הבנק	במצב של שימוש במערכת בינה מלאכותית בסביבה חיצונית לבנק- סיכון כי עובדי הבנק יזינו למערכת הבינה מלאכותית, נתונים רגישים בגין לקוחות הבנק ו/או מידע מסחרי של הבנק ו/או קוד תוכנה אשר עשויים לחשוף את הבנק לדלף מידע רגיש ו/או אירוע סייבר
העברת נתוני הבנק או לקוחותיו לגופים אחרים	במצב של שימוש במערכת בינה מלאכותית בסביבה חיצונית לבנק- הסיכון שהמערכת תספק נתונים של הארגון ללקוח אחר המשתמש במוצר שלהם, לדוגמא, על ידי אימון המודל בשאלות המכילות נתונים פנימיים של שהארגון.
מתקפת סייבר על ספק המודל	סיכון כי מערכת הבינה המלאכותית תופעל על ידי תוקף עוין והאחרקון ינצל את הממשקים שיש למערכת אל מול הבנק ו/או לקוחותיו לצורך מימוש אירועי סייבר. לדוגמא, בשימוש של צ'אט-בוט מול לקוח, פיתוי הלקוח לביצוע פעולות פיננסיות לטובת התוקף (מקרה פרטי של הנדסה חברתית).
Jailbreaking and Adversarial attacks	מתקפות כמו "Adversarial attacks" ו-"Jailbreaking" מבוססות על זיהוי וניצול חולשות המודל על ידי התוקף. התוקפים לומדים את המודל ובאמצעות Prompts זדוניים מצליחים לגבור על מגבלות המודל. לדוגמא, בשימוש של צ'אט בוט, שבירה של מגבלות בנוגע להוגנות מול לקוחות. דוגמא נוספת, בשימוש של מודל LLM הסורק מסמכים להסקת תובנות פיננסיות, זיהוי החולשות של המודל והוספת פסקאות העשויות לשבש את תוצאות המודל (ייפוי סנטימנט בסקירת דוחות כספיים, הפחתת סיכוני הלבנת הון ועוד).
זיהוי ספציפי של לקוחות הבנק	הסיכון שספק תוכנת הבינה המלאכותית לארגון יוכל להסיק מסקנות על פעילות הלקוחות עקב שאלות שנשאלות ע"י הלקוחות תוך כדי שישתמשו בשירות.

סיכוני מוניטין

סיכון	פירוט הסיכון
מחסור בתקשורת עם גורם אנושי	הסיכון שישנו פלח לקוחות פוטנציאלים או קיימים אשר מעדיף שההתקשרות מול הארגון תעשה ע"י נציג אנושי ושהטמעת מערכת התקשרות מבוססת בינה מלאכותית תגרום להם לעזוב או לבחור בשירות של בנק אחר.
היעדר שקיפות בנוגע לתצורת הפעילות מול ספק AI וסיכון לגבי דעת הציבור על הבנק ו\או התחום	הסיכון לפגיעה במוניטין הארגון בעקבות תצורת העבודה עם ספק מערכת הבינה המלאכותית במסגרתה לא קיימת שקיפות לגבי הצורה בה נשמר המידע הרגיש שלהם, מועבר לגוף חיצוני ואיזה שימוש מתבצע בו. כמו כן קיים סיכון לגבי עצם עבודה עם ספקי שירותי בינה מלאכותית שונים במידה ודעת הציבור על התחום תשתנה בעקבות סיכונים חדשים או קיימים שיתממשו בנוגע לבינה המלאכותית. לדוגמא, התבטאויות או האשמות כנגד מודלים כגון GPT, Bard, Llama אשר הבנק בחר לעבוד איתם עשוי להשפיע על דעת הציבור על הבנק.
פגיעה בזכויות הלקוח בכתב תנאי השירות	הסיכון שהלקוח יידרש לחתום על כתב וויתור מבלי להבין בצורה ראויה כי הוא מאשר פגיעה כלשהי בזכויותיו כפי שהן מגיעות לו, מצב זה עלול לגרום נזק מוניטין במידה וייוודע לתקשורת.
אפליה	הסיכון שיוטמעו במודל, בזדון או שלא בזדון, הנחות בסיס שיפלו פלח אוכלוסייה ספציפי אשר יחשפו את הארגון לסיכון מוניטין.
אמון	הסיכון לפגיעה באמון הלקוחות בעקבות כשלים של המודל שיתוקשרו במדיה.
תגובה אנושית מדי	הסיכון שבהיקשרות יתר של לקוח אל –"פרסונה" שתוצג ע"י מערכת הבינה המלאכותית תוך כדי התקשרות עם הלקוח כנציגת הארגון, מצב שעלול ליצור רושם שהארגון מנצל את סוג התקשורת הזה על מנת ליצר רווח אל מול הלקוחות.
יצירת תוכן פוגעני	הסיכון כי במהלך קמפיין שיווק מבוסס מערכת בינה מלאכותית יוצרת יוצג ללקוח תוכן ויזואלי שימצא כפוגעני.

סיכוני הונאה

סיכון	פירוט הסיכון
ניצול מודל הבינה המלאכותית	הסיכון שגורם זר ינסה לשתול מידע בצורה כזו שתבלבל את מודל קבלת ההחלטות של הבינה המלאכותית על מנת להפיק מידע ספציפי שברצונו לגלות. (מקרה פרטי של Jailbreaking/Adversarial attack)
הונאה בהתחזות | צד לקוח	הסיכון שגורם זר שאינו מורשה ישיג גישה ל-"שיחה" או הזנת שאילתות אל מערכת הבינה המלאכותית של הארגון (במידה ונועדה ליצירת קשר עם הלקוחות) ויצלח לשכנע אותה כי הוא לקוח אמיתי של הארגון ובכך יקבל גישה אל מידע של אותו לקוח אליו התחזה או שיוכל לבצע פעולות בשמו.
הונאה בהתחזות | צד הבנק	הסיכון שגורם זר יוכל לתקשר עם הלקוח תוך כדי התחזות לאחת ממערכות הבינה המלאכותית של הבנק עקב חולשת אבטחה בהן. (כגון הפרצה ShellTorch שהתגלתה ע"י חברת Oligo באוקטובר 2023 המאפשרת הרצת קוד זדוני מרחוק אל מערכת הבינה המלאכותית TorchServe)

סיכונים עסקיים

סיכון	פירוט הסיכון
סיכונים תפעוליים- כח אדם	הסיכון שככל ומערכת בינה מלאכותית עליה מסתמך הארון תקרוס לא יהיה מספיק כוח אדם בנמצא על מנת לגבות את כל הפעולות שנעשו עד כה בעזרת המערכת. סיכון זה ילך ויגבר ככל והשימוש במערכות מסוג זה יהווה תחליף לכוח האדם בקו הראשון של השירות.
ביצועים נמוכים	הסיכון שחווית המשתמש של הלקוח אל מול מערכות בינה מלאכותית שנועדו לצורך זה לא תהיה מספקת בעיני הלקוחות, מצב שעלול לגרום ללקוחות לא לרצות להשתמש בשירות זה ואף לעזוב את הארגון במקרי הקיצון.
סיכוני הקצאת משאבים	הסיכון כי הארגון יקצה משאבים מעטים מידי שלא יאפשרו פיתוח ותפעול של המערכת הבינה המלאכותית בסטנדרטים שהוגדרו. סיכון משאבים זה יכול להיות תקף הן עבור הקצאת הון והן עבור הקצאות שעות פיתוח או רכישת כלי מחשוב נוספים, מה שעלול להוביל לצורך בגיוס הון נוסף או לקצץ במשאבים אחרים.
קושי בגיוס עובדים מקצועיים	הסיכון שעקב הגידול המהיר בשוק הטכנולוגי סביב מערכות בינה מלאכותית הביקוש לעובדים מקצועיים בתחום יעלה, מה שעלול להקשות על גיוס אנשי מקצוע בתחום.
מכירת הספק החיצוני לגוף אחר	הסיכון כי ספק מערכת הבינה המלאכותית תימכר בעתיד לגוף שיבחר שלא ימשיך את ההתקשרות עם הארגון או שיבחר לא לעמוד בתנאים שהוסכמו בין הספק לארגון בעת ההתקשרות המקורית.
פיתוח תלות בכלי	הסיכון שהארגון יפתח תלות במערכת הבינה המלאכותית לשם מתן שירות ללקוחות. במצב בו המערכת קורסת או שהספק יחליט להעלות את דמי השימוש לרמה לא סבירה הארגון עלול למצוא עצמו ללא מערך גיבוי מספק.
סיכוני תדמית	במידה וגישת השוק אל מערכות בינה מלאכותית תהפוך לשלילית עם הזמן יתכן ולקוחות חדשים יעדיפו שלא להתקשר עם ארגון שמכיל או מתבסס על מערכת כזו.

סיכונים עסקיים המשך

סיכון	פירוט הסיכון
שגיאות מודל בשיווק	הסיכון שמערכת הבינה המלאכותית תבצע "טרגטינג" שגוי בצורה עקבית ומתמשכת ותציג פרסום שאינו רלוונטי לפלח האוכלוסייה הנחבר.
הסתמכות "עיוורת"	הסיכון כי עובדי הבנק יסתמכו במידה רבה מידי על מערכות בינה מלאכותית עד כדי שחיקת מקצועיות נותני השירות המשתמשים במערכות אלו.