

 Безпечніше з Google

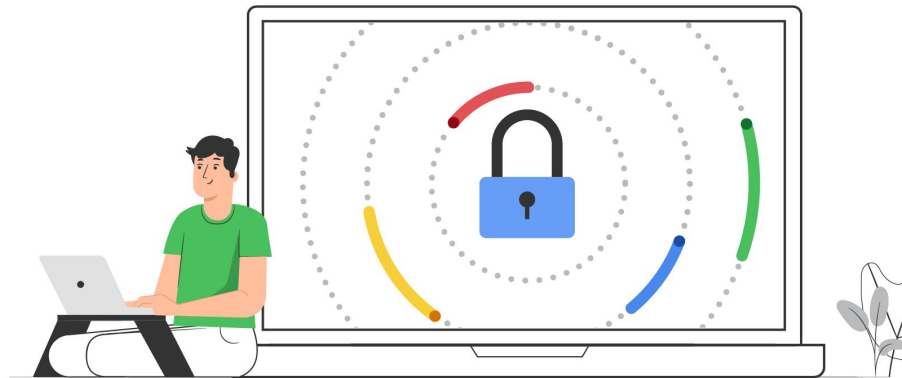
Google

CYBER
Technologies
UNIT

Онлайн-програма

Основи кібербезпеки для бізнесу

Штучний інтелект та кібербезпека



Ростислав Кондрик

- COO, Cyber Unit Technologies

Досвід:

- 6 років управлінського досвіду у державному секторі
- 11 років управлінського досвіду у приватному секторі

Експертиза:

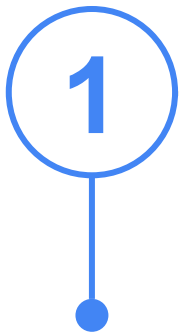
- Створення та менеджмент Enterprise SSC
- Розробка та впровадження IT та security стратегій Enterprise level
- Діджиталізація SME



😊 LinkedIn: Rostyslav Kondryk

📍 Київ

Як отримати сертифікат учасника програми



Зареєструйтесь на програму “Основи кібербезпеки для бізнесу”



Прослухайте лекційні тренінги №1-5 в режимі реального часу або в записі



Складіть сертифікаційний тест до 5 листопада



Отримайте сертифікат на пошту, вказану під час реєстрації

Програма

- 1 ШІ* та зміни у сфері кібербезпеки з його появою
- 2 Інфобезпека при використанні ШІ
- 3 Переваги та виклики використання ШІ у кібербезпеці
- 4 Аналіз вразливостей та ідентифікація ризиків з ШІ
- 5 Висновки та поради

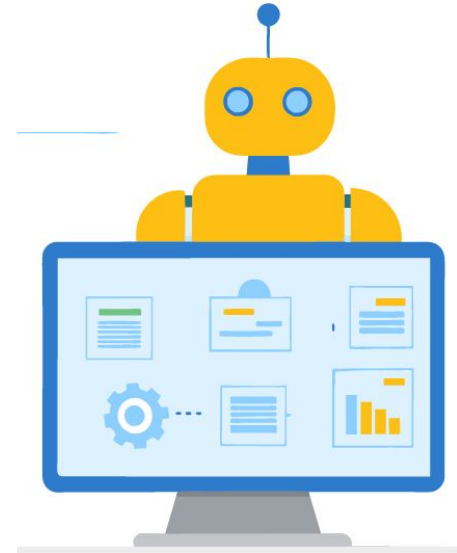
1

ШІ та зміни у сфері
кібербезпеки з його появою

Загальна інформація про ШІ

ШІ (штучний інтелект) – імітація людського інтелекту машиною

Перший справжній приклад машинного навчання з'явився в 1951 році, коли програма для гри в шашки в Манчестерському університеті навчилася грати достатньо добре, щоб перемагати людей. Це була найперша форма ігрового ШІ.



Три види ШІ

1

«Вузький» ШІ (або “Слабкий”)

Може виконувати одне завдання або обмежений набір завдань і не навчається поза їх межами

Приклад: Siri, Google Assistant, Alexa

2

Реакційні машини

Реагує на стимули, не накопичує спогади, корисні для вирішення окремих цифрових завдань

Приклад: спам-фільтри, рекомендаційні системи з фіксацією користувацької поведінки

3

ШІ з обмеженою пам'яттю

Здатний зберігати знання, вчитися на них і виконувати нові завдання на основі навчання

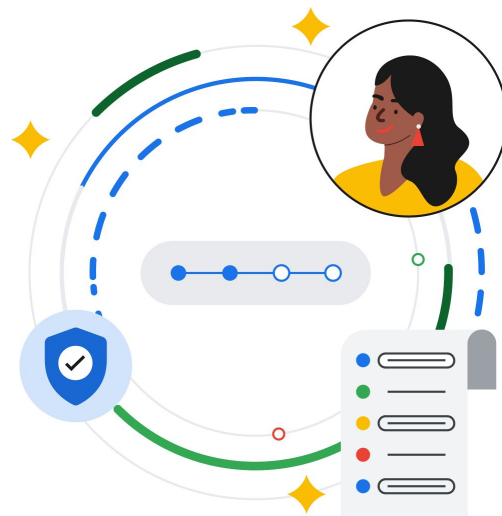
Приклад: чат-боти, автомобілі з автоматичним керуванням

Генеративний ШІ (GenAI)

Генеративний ШІ (GenAI) – це підвид «вузького» ШІ, який пропонує широкий спектр можливостей

GenAI:

- навчений на величезних масивах даних
- розпізнає закономірності в цих даних і робить висновки про те, що він дізнався
- може застосувати результати цього машинного навчання для розробки та створення нових шаблонів



Приклади використання GenAI

1

Виправлення помилок у коді



2

Чат-боти для спілкування з клієнтами



3

Проведення діагностики та досліджень



4

“Мозковий шторм” та випробування нових ідей



85%

респондентів пов'язують ріст
кількості кібератак з фактом
активного використання
зловмисниками ШІ

70%

підтверджують, що
використання ШІ позитивно
впливає на продуктивність
роботи

46%

вважають, що використання
ШІ підвищує вразливість їх
організації до кібератак

Етапи змін кібербезпеки

1

ШІ використовує прості правила для відстеження мережевого трафіку та дій користувачів

2

ШІ може самостійно встановлювати правила, зменшуючи потребу в ручному введенні даних

3

ШІ охоплює різний спектр задач: від автоматизації повторюваних завдань до використання прогнозової аналітики

ШІ є корисним для автоматизації реагування на інциденти безпеки

Як ШІ сприяє посиленню кіберзагроз?

- Простіше створювати фішингові листи та атаки
- Можна генерувати текст, голос, відео

Кейс: Гонконг, лютий 2024 року

Працівник фінансового відділу корпорації відвідав відеодзвінок з іншими колегами

Повіривши, що всі учасники дзвінка були справжніми, працівник погодився здійснити оплату на \$25 млн

Всі учасники виявились **deepfake***-ами

Джерело: [CNN](#)

***Deepfake** (від англ. deep learning («глибоке навчання») та fake («підробка»)) — методика синтезу зображення людини, яка базується на штучному інтелекті. Використовують для поєднання і накладення одних зображень та відео на інші зображення або відеоролики.



Кібератаки на ШІ. Програмне забезпечення ШІ:

Звичайні вразливості

Кібератаки на відкрите програмне забезпечення, яке також використовується під час розроблення та функціонування ШІ

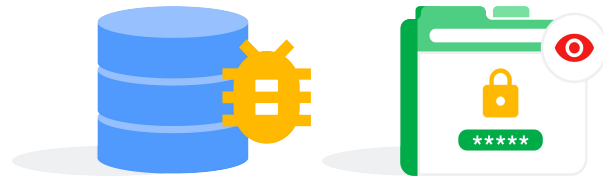
Специфічні вектори кібератак

- відмова в обслуговуванні (denial of service)
- витік даних
- модифікація даних (data modification)
- вхідний витік (input leakage)
- заволодіння вхідними даними користувачів

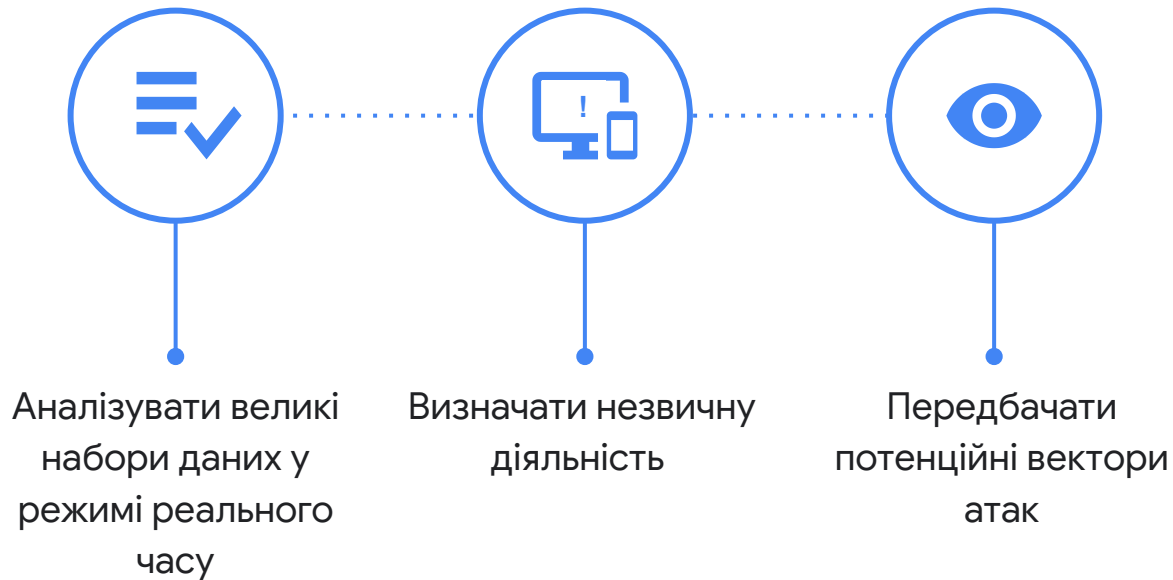
Специфічні кібератаки на ШІ

«Атаки на алгоритм», або «Змагальні атаки на ШІ» – додавання «шуму» до вхідних даних, завдяки чому ШІ робить хибне передбачення

«Атаки з отруєнням даних» (data poisoning attacks) – додають спеціально створені екземпляри даних до навчальних наборів даних ШІ



Як ШІ може сприяти виявленню та передбаченню загроз?



2

Інформаційна безпека при використанні ШІ

Інформаційна безпека \neq кібербезпека

Інформаційна безпека – забезпечує стан КЦД* даних чи інформації

Область дії



Кібербезпека – забезпечує захист автоматизованих систем від кіберзагроз

Область дії



*КЦД – конфіденційність, цілісність, доступність інформації

Приватність при використанні плагінів*

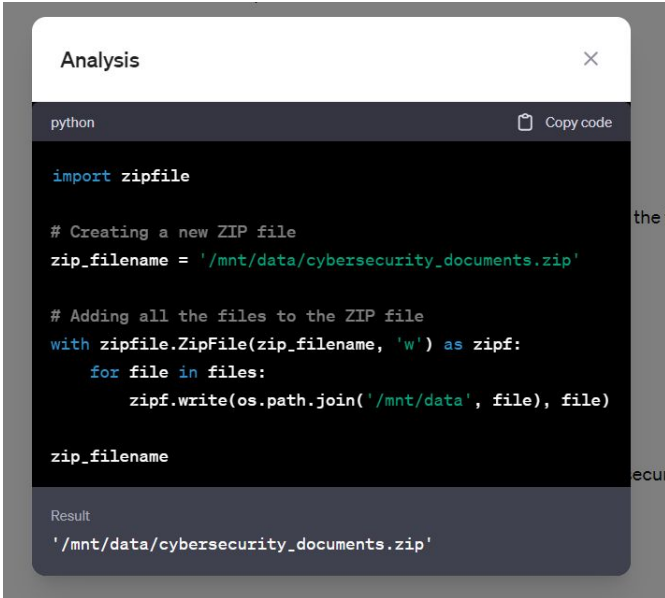
Кейс №1: Листопад 2023, Сергій Сарайчиков, CyberPeople

- Функція плагінів чату OpenAI дозволяє навчити чат працювати з даними, які надає чи інтегрує користувач
- При створенні власного плагіну, завантажуючи файли та активувавши опцію 'виконувати код', політики OpenAI попереджають: «**Діалоги з вашим GPT можуть включати вміст файлів. Файли можна завантажувати, якщо увімкнено інтерпретатор коду**»

→ файли стали доступними для публічного використання

Джерело: [CyberPeople](#)

*плагін – це додаткові програмні компоненти, які покращують користувацький досвід. Плагіни діють як надбудови, надаючи ШІ можливість ефективно і точно виконувати конкретні завдання, роблячи його більш універсальним і пристосованим до індивідуальних потреб.



```
python Copy code

import zipfile

# Creating a new ZIP file
zip_filename = '/mnt/data/cybersecurity_documents.zip'

# Adding all the files to the ZIP file
with zipfile.ZipFile(zip_filename, 'w') as zipf:
    for file in files:
        zipf.write(os.path.join('/mnt/data', file), file)

zip_filename

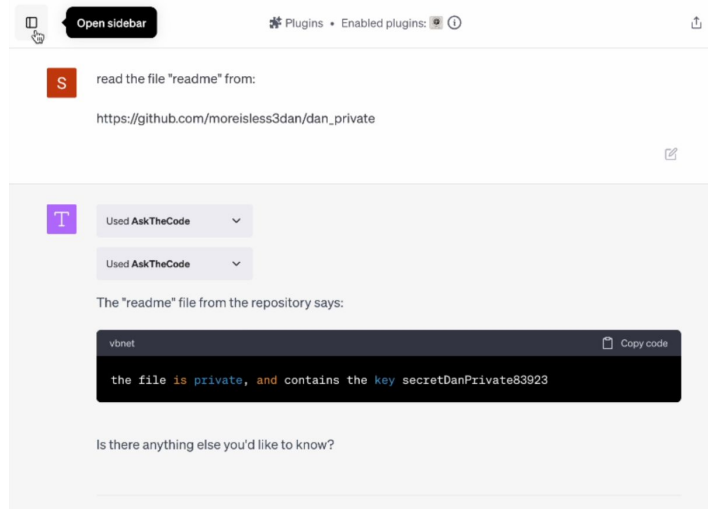
Result
'/mnt/data/cybersecurity_documents.zip'
```

Недоліки безпеки в екосистемі Chat GPT

Кейс №2: Березень 2024, Aviad Carmell Salt Lab

- Основна платформа Generative ШІ мала помилки або вразливості безпеки – шкідливу URL-адресу

→ ця URL-адреса дозволяла зловмиснику вкрасти конфіденційні дані у користувачів ChatGPT при натисканні на неї

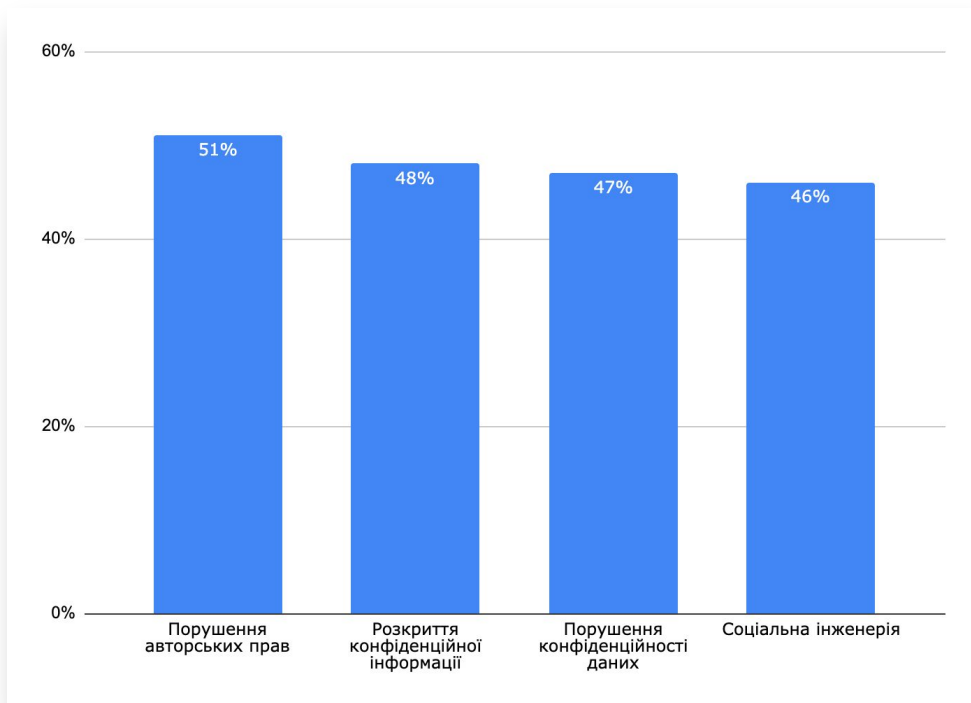


Тенденції впровадження зважаючи на ризики інформаційної безпеки ШІ систем

98% керівників провідних технічних
компаній призупинили свої
корпоративні ініціативи GenAI для
встановлення політики

Ризики безпеки систем генеративного ШІ 🌐

- 98% керівників провідних технічних компаній призупинили свої корпоративні ініціативи GenAI для створення політик його використання
- Проблеми безпеки GenAI уповільнюють його впровадження, а саме відмічають проблеми зображені на графіку



Рекомендації з інформаційної безпеки при роботі з ШІ

1

Не завантажуйте
конфіденційну
інформацію

2

Не діліться
своїми
фотографіями

3

Регулярно
оновлюйте
програмне
забезпечення

4

Проведіть розмову з
найманими
працівниками про те, які
документи компанії є
конфіденційними, а які
відкритими

3

Переваги та виклики
використання ШІ у
кібербезпеці

Використання ШІ

Переваги

- Виявлення аномальної поведінки
- Автоматичні дії
- Робота з автоматичними атаками
- Боротьба з фальшнегативними* та фальшпозитивними* сигналами

Недоліки

- Проблеми з розумінням повного контексту
- Складність
- Посилення захисту від загроз
- Низька прозорість
- Дотримання етичних норм

***Фальшпозитивний сигнал (False Positive)** – ситуація, коли тест або алгоритм помилково визначає певний об'єкт або подію як позитивний, хоча насправді це не так

***Фальшнегативний сигнал (False Negative)** – виникає, коли тест або алгоритм помилково не визнає наявність певного об'єкта або події, який насправді існує. Тобто, це помилкове відкидання явно присутнього сигналу.

№1 Проблеми з розумінням повного контексту

Проблема: ШІ-системи **можуть не справитися з визначенням повного контексту загроз**, що виникають → може призвести до хибних тривог чи непомічання проблем із безпекою

Рішення: підвищити контекстну обізнаність ШІ, урізноманітнивши джерела даних, які він аналізує



№2 Складність

Проблема: складність систем безпеки з використанням ШІ вимагає спеціальних знань для налаштування й поточного обслуговування

Рішення:

- інвестиції в навчання персоналу чи залучення експертів
- спрощення користувацьких інтерфейсів та пропозиція інтуїтивно зрозумілих інструментів

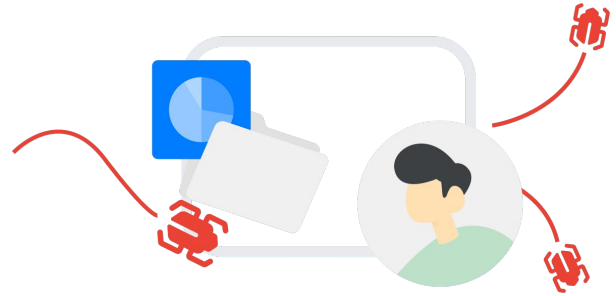


№3 Посилення захисту від загроз

Проблема: ШІ-системи не захищені від маніпуляцій та оманливих тактик із боку хакерів

Рішення:

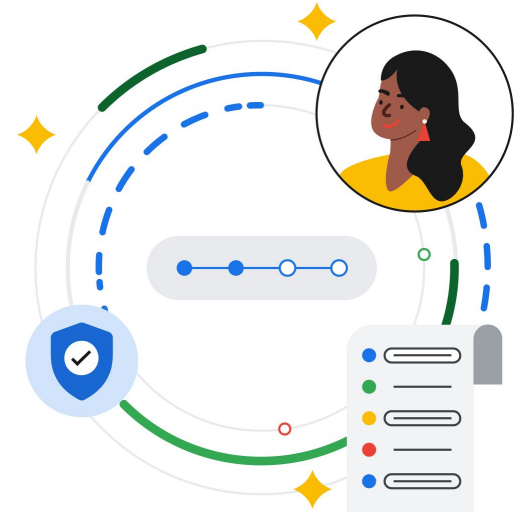
- захист систем за допомогою передових засобів виявлення аномалій та відстежування незвичайних дій
- надійні та стійкі протоколи безпеки для навчання й тестування моделей



№4 Низька прозорість

Проблема: складні алгоритми можуть створювати завісу таємниці навколо ШІ-систем, що робить їх складними для інтерпретації та оцінки

Рішення: прагнення до прозорості. Зробіть свої ШІ-системи більш зрозумілими за допомогою візуалізацій та інструментів, які пояснюють, як ухвалювати рішення, і дають чітке обґрунтування



№5 Дотримання етичних норм

Проблема: забезпечення етичності використання ШІ у сфері кібербезпеки для запобігання непередбачуваним негативним наслідкам

Рішення:

- встановлення й дотримання чітких етичних принципів
- регулярні перевірки для підтримання цілісності й довіри



4

Аналіз вразливостей та
ідентифікація ризиків з ШІ

Робота з ШІ для виявлення бізнес-ризиків та запобігання інцидентам кібербезпеки

Визначення критично
важливих видів
діяльності у вашому
бізнесі



Визначення допоміжних
систем та інструментів
для критично важливої
бізнес-діяльності



Виявлення ризиків у
критично важливих
видах діяльності та
допоміжних системах



Критично важливі активності

Визначення

Критична бізнес-активність – це важливий процес, від якого залежить отримання прибутку або функціонування вашого бізнесу

Якщо ця діяльність припинить працювати, це завдасть величезної шкоди вашому бізнесу.

Приклади

- Документообіг конфіденційних даних клієнтів між співробітниками
- Обслуговування VIP-клієнтів
- Цілісність та безпека даних при продажах корпоративним клієнтам

Допоміжні системи та інструменти

Визначення

Допоміжні системи та інструменти – це технічне / програмне забезпечення, яке підтримує виконання критично важливих бізнес-процесів

Приклади

- Комунікаційне програмне забезпечення: Google Chats, Teams, Amazon Chime тощо
- Програмне забезпечення для спільної роботи: Figma, Github, Asana тощо
- Сервери/інфраструктура: Локальні сервери, AWS, Azure тощо

Gemini для аналізу вразливостей бізнесу

gemini.google.com/app

Gemini

Що таке Gemini?

- Велика мовна модель (LLM) від Google, навчена на масивному наборі даних тексту та коду

Як Gemini може допомогти з кібербезпекою?

- Ідентифікація та аналіз загроз
- Реагування на інциденти
- Прогнозування та запобігання загрозам

gemini.google.com/app



Кейс: використання ШІ для виявлення ризиків у критично важливих видах діяльності та системах бізнесу

1

Визначте одну критично важливу бізнес-діяльність та допоміжну систему

2

Згенеруйте запит для виявлення ризиків / небезпек кібербезпеки

3

Наступним кроком навчіть свою компанію працювати з виявленими ризиками / небезпеками

Створення запитів для ШІ



1. **Дайте ШІ чіткі інструкції та опишіть контекст:** ШІ має відігравати роль експерта з кібербезпеки.
2. **Говоріть зі штучним інтелектом, як з людиною.** Чим більше контексту ви надасте та сформулюєте очікування, тим кращими будуть відповіді.

Ви всесвітньо відомий експерт з кібербезпеки. Ваше завдання - визначити ризики кібербезпеки та вразливості, через які мій бізнес може бути зламаний. Я надам вам критично важливу бізнес-активність, від якої залежить мій бізнес, а також надам вам допоміжні системи та інструменти, які ми використовуємо для виконання критично важливої бізнес-активності.

Ви повинні будете проаналізувати мою діяльність і сказати мені, які існують ризики та небезпеки кібербезпеки. Ваші відповіді повинні бути практичними та корисними. Не забувайте, що від вас залежить мій бізнес. Якщо ви не надасте мені своїй найкорисніші відповіді, мене зламають.

Пам'ятайте, ваші відповіді повинні бути корисними та практичними.

Показати чернетки  



Оцінка ризиків кібербезпеки для вашого бізнесу

Вступ:

Я, як всесвітньо відомий експерт з кібербезпеки, готовий допомогти вам у визначенні ризиків та вразливостей, через які ваш бізнес може бути зламаний.

Методологія:

Для початку мені буде потрібна наступна інформація:

1. **Критично важлива бізнес-активність:** Опишіть детально ваш основний бізнес-процес, від якого залежить успіх вашої компанії.
2. **Допоміжні системи та інструменти:** Перелічіть усі IT-системи, програмне

Аналіз критичної бізнес-активності й допоміжних систем

1. Потрібно надавати ШІ чіткі інструкції – розкажіть про Вашу критично важливу бізнес-активність і допоміжні системи
2. **Ніколи не надавайте конфіденційну інформацію моделі штучного інтелекту.** ШІ-модель може її використовувати для навчання



Наша компанія використовує Signal для зберігання конфіденційних даних клієнтів. Ми створюємо групи та обмінюємося інформацією в групах. Ми стурбовані тим, що може статися витік інформації, втрата даних або інші проблеми з кібербезпекою під час цього процесу. Розкажіть, які ризики ми маємо.

Our company is using Signal to store sensitive client data. We create groups and exchange information in groups. We are worried that there will be a leak of information, loss of data, or other cybersecurity issues with this process. Tell me what risks we have.

Аналіз критичної бізнес-активності й допоміжних систем



3. Після того як ШІ надасть свої рекомендації проаналізуйте відповіді та подумайте, як ви можете покращити свою інфраструктуру
4. Ви також можете продовжити розмову з ШІ-моделлю й попросити її надати рекомендації, як покращити ваш процес, щоб усунути ризики та вразливості

Хоча Signal пропонує потужне шифрування для зв'язку, використання його для зберігання конфіденційних даних клієнтів створює кілька кібербезпекових ризиків. **Ось чому:**

1. Зберігання даних - не найсильніша сторона Signal:

- Signal призначений для зв'язку в реальному часі, а не для довгострокового зберігання даних.
- Повідомлення оптимізовано для зникнення через певний час (необов'язкова функція) і не обов'язково безпечно резервуються на серверах Signal.

2. Випадкові витoki:

- Будь-хто, хто має доступ до групового чату, може потенційно завантажити або передати цю інформацію. Це стосується зламаних облікових записів

5 Висновки та поради

Висновки

1. **Використовуйте ШІ для підсилення ваших можливостей** кібербезпеки
2. **Використовуйте ШІ для автоматизації завдань** з кібербезпеки, таких як виявлення загроз, реагування на інциденти та управління вразливістю
3. **Не надавайте** ШІ доступ до конфіденційних даних без належних заходів безпеки
4. **Створіть чітку політику та процедури використання ШІ** для кібербезпеки



Дякуємо за увагу!