Анна ПОЛІЩУК

КІБЕРАТАКИ, ЗАГРОЗИ БЕЗПЕЦІ ТА ЗЛОВЖИВАННЯ CHATGPT

З моменту появи ChatGPT у 2022 році він став найпопулярнішою платформою зі штучним інтелектом, здатною генерувати тексти, коди, листи та сценарії. Проте будь-яка технологія має дві сторони: з одного боку — користь для освіти, науки й бізнесу, з іншого — ризик зловживань.

Останні роки показали, що хакери почали використовувати штучний інтелект не лише як помічника, а й як інструмент для підготовки кібератак, зокрема з участю ChatGPT.

За даними OpenAI, «закордонні загрозливі актори» вже використовують ChatGPT та інші інструменти ШІ для підкріплення старих способів атак:

- фішинг — створення правдоподібних листів для отримання конфіденційної інформації. ChatGPT генерує правдоподібні тексти, що важко відрізнити від легітимних повідомлень;

- шкідливе програмне забезпечення (malware) — автоматизоване написання кодів для атак;

- пропаганда та дезінформація — генерація переконливих текстів для впливу на аудиторію;

- автоматизація кіберзлочинів — швидке створення сценаріїв атак без необхідності глибоких технічних знань. Генеративний AI може створювати переконливі повідомлення для впливу на громадську думку.

Фіксація та публікація загроз дозволяє підвищити обізнаність про потенційні ризики та розробляти захисні механізми.

За даними OpenAI (компанії-розробника ChatGPT), "закордонні загрозливі актори" (foreign threat actors) вже використовують ChatGPT та інші інструменти ІІІ для «підкріплення старих способів атак»: фішинг, malware, пропаганда, автоматизація кіберзлочинів. Цей випадок розслідувався і оприлюднений як частина "public threat reporting".[5]

[5] https:// cybernews.com/security/openai-report-foreign-threat-actors-use-ai-tools-chatgpt-for-attacks/

Це означає, що інформацію про потенційні або реальні зловживання та загрози фіксують і повідомляють — як великі компанії безпеки Kaspersky, так і сама OpenAI + незалежні дослідники і аналітики.

Приклад загрози у 2025 році. Microsoft повідомила про новий тип загрози — вірус-троян під назвою PipeMagic[6], який поширювався як «підроблений десктоп-додаток ChatGPT». Цей «фейковий» додаток міг діяти як бекдор, викрадаючи дані або надаючи зловмисникам доступ до систем. Атака PipeMagic під виглядом ChatGPT Desktop App Microsoft у 2025 році повідомила, що вірус-бекдор PipeMagic розповсюджується через підроблений додаток ChatGPT. Постраждалі організації були в США, Європі та інших регіонах. PipeMagic дозволяє зловмисникам віддалено керувати комп'ютерами, встановлювати додаткове шкідливе ПЗ та викрадати дані.

PipeMagic є яскравим прикладом того, як популярність AI-сервісів використовується

[6] https://www.microsoft.com/en-us/security/blog/ 2025/08/18/dissecting-pipemagic-inside-the-architecture-of-a-modular-backdoor-framework/

кіберзлочинцями як прикриття для атак.

Кількість кібератак, де використовують ChatGPT або подібні інструменти ШІ, зростає щороку. Це свідчить, що популярність таких сервісів робить їх привабливою маскою для шахраїв та кіберзлочинців.

Поява нових методів автоматизації кібератак стає серйозною загрозою для бізнесу, освіти та державних структур.

Аналітики прогнозують подальше збільшення атак на 50–70% до 202С–2027 років, якщо не будуть впроваджені нові системи захисту[7].

Сервіси AI будуть дедалі більше використовуватися як інструмент соціальної інженерії та дезінформації.

ChatGPT та подібні ШІ-сервіси стають все більш привабливим інструментом для кіберзлочинців. Зростання кількості шкідливих файлів, фішингових атак і підроблених додатків свідчить про те, що популярність таких технологій використовується не тільки на користь

[7] https://www.ncsc.gov.uk/report/impact-ai-cyber-threat-now-2027?utm_source=chatgpt.com

людей, а й на шкоду.

Отже, популярність ChatGPT створює нові кіберзагрози, і їхнє своєчасне виявлення та запобігання, має ключова умова безпеки для користувачів і організацій.

Лише комбінований підхід, що включає технічні, освітні та організаційні заходи, дозволить зменшити ризики і забезпечити безпеку користувачів та організацій у світі, де AI стає все більш потужним інструментом.