Секція 1

ДОСЛІДЖЕННЯ ПИТАНЬ КОНФІДЕНЦІЙНОСТІ ТА БЕЗПЕКИ В СИСТЕМАХ РОЗПІЗНАВАННЯ ТЕКСТУ

Щеглов А. О.
Національний аерокосмічний університет ім. М.Є. Жуковського «ХАІ»
Науковий керівник: Шостак А. В.

Актуальність. Інтенсивний розвиток комп'ютерних систем логічно призвів до широкого впровадження інноваційних методів обробки даних, які базуються на використанні штучного інтелекту. Ця тенденція особливо помітна у сфері оптичного розпізнавання тексту. Оптичне розпізнавання тексту (англ. optical character recognition, OCR) — це механічне або електронне переведення зображень рукописного, машинописного або друкованого тексту в послідовність кодів, що використовуються для представлення в текстовому редакторі [1]. Однак із збільшенням використання OCR також зростає занепокоєність щодо конфіденційності та безпеки.

Мета. Дослідження та аналіз питань безпеки в системах оптичного розпізнавання тексту.

Основні положення. Технологія оптичного розпізнавання символів (OCR) має можливість сканувати та аналізувати інформацію зображень документів, яка часто містить конфіденційну інформацію. Ці документи можуть включати медичні файли, фінансові звіти, юридичні договори та інші конфіденційні дані. Це породжує занепокоєння щодо можливого ризику витоку конфіденційної інформації, яка може потрапити до несанкціонованих осіб. Технологія OCR також може неочікувано отримувати особисті дані, такі як ім'я, адреса чи номери соціального страхування, що може викликати загрозу крадіжки особистої інформації. Відтак, важливо, бути уважним до цих конфіденційних питань та приймати заходи для захисту конфіденційної інформації під час використання технології OCR. Технологія оптичного розпізнавання символів викликає занепокоєння через потенційні проблеми не лише з конфіденційністю, але й з безпекою даних. Ця технологія передбачає зберігання та передачу даних, що створює ризики, які можуть використовувати кіберзлочинці. Якщо дані, отримані через OCR, не захищені належним чином в базі даних, несанкціоновані особи можуть мати до них доступ. Це може призвести до витоку даних, що, в свою чергу, може призвести до фінансових збитків, шкоди репутації або юридичних відповідальності. Отже, важливо вирішити ці проблеми безпеки та гарантувати захист конфіденційних даних під час усього процесу використання OCR.

Висновки. Роботу з OCR необхідно розпочати з проведення оцінки ризиків для виявлення потенційних загроз та вразливостей, пов'язаних із

використанням технології. Одним із ключових аспектів є обробка конфіденційних даних під час процесу розпізнавання тексту. Слід передбачити вразливість даних при передачі їх відкритими каналами зв'язку. Важливо встановити протоколи обробки даних, що захищають конфіденційну інформацію від несанкціонованого доступу, використання або розголошення. Це може включати універсальні методи кодування повідомлень, контроль доступу та процедури резервного копіювання, відновлення даних та їх безпечне знищення у разі припинення використання.

Список літератури

1. Оптичне розпізнавання символів. *Wikipedia*. URL: https://uk.wikipedia.org/wiki/Оптичне_розпізнавання_символів (дата звернення: 22.11.2023);
2. Konstantin Dergachov. Development of tools for information protection of optical text recognition systems. URL: https://www.researchgate.net/publication/361828225_Development_of_tools_for_information_protection_of_optical_text_recognition_systems (дата звернення: 22.11.2023).

Відомості про авторів

Щеглов Антон Олександрович, магістрант кафедри комп'ютерних систем, мереж і кібербезпеки, НАУ «ХАІ», a.shchehlov@student.csn.khai.edu
Шостак Анатолій Васильович, доцент кафедри комп'ютерних систем, мереж і кібербезпеки, НАУ «ХАІ», к.т.н., доцент, a.shostak@csn.khai.edu