PROCEEDINGS OF THE V INTERNATIONAL SCIENTIFIC AND THEORETICAL CONFERENCE

SCIENCE OF XXI CENTURY: DEVELOPMENT, MAIN THEORIES AND ACHIEVEMENTS

26.01.2024

HELSINKI
REPUBLIC OF FINLAND

SCIENTIA
COLLECTION OF SCIENTIFIC PAPERS

SCIENTIA
COLLECTION OF SCIENTIFIC PAPERS

with the proceedings of the

V International Scientific and Theoretical Conference

Science of XXI century: development, main theories and achievements

26.01.2024

Helsinki, Republic of Finland

Helsinki, 2024

UDC 082:001
S 40

https://doi.org/10.36074/scientia-26.01.2024

S 40 **Science of XXI century: development, main theories and achievements**: collection of scientific papers «SCIENTIA» with Proceedings of the V International Scientific and Theoretical Conference, January 26, 2024. Helsinki, Republic of Finland: International Center of Scientific Research.

ISBN 979-8-88955-774-6 (series)
DOI 10.36074/scientia-26.01.2024

Papers of participants of the V International Multidisciplinary Scientific and Theoretical Conference «Science of XXI century: development, main theories and achievements», held on January 26, 2024 in Helsinki are presented in the collection of scientific papers.

The conference is included in the Academic Research Index ReserchBib International catalog of scientific conferences and registered for holding on the territory of Ukraine in UKRISTEI (Certificate № 318 dated June 16th, 2023).

UDC 082:001

ISBN 979-8-88955-774-6

SECTION 11.
GENERAL MECHANICS AND MECHANICAL ENGINEERING

SECTION 12.
AUTOMATION AND APPLIANCES MAKING

SECTION 13.
ENERGY AND POWER ENGINEERING

SECTION 14.
ECOLOGY AND ENVIRONMENTAL PROTECTION TECHNOLOGIES

SECTION 15.
COMPUTER AND SOFTWARE ENGINEERING

Демидов Захар Георгійович
Старший науковий співробітник
науково-дослідної лабораторії
з проблем інформаційних технологій та протидії злочинності у кіберпросторі
Харківський національний університет внутрішніх справ, Україна

Коломійцев Сергій Олександрович
Науковий співробітник
науково-дослідної лабораторії
з проблем інформаційних технологій та протидії злочинності у кіберпросторі
Харківський національний університет внутрішніх справ, Україна

ПРОБЛЕМИ БЕЗПЕКИ НА 2024 РІК ВІД ШТУЧНОГО ІНТЕЛЕКТУ

Однією з ключових тем обговорення у 2023 році у світі технологій та науки став термін "штучний інтелект". Останнє щорічне глобальне дослідження McKinsey[1] щодо поточного стану штучного інтелекту підтверджує вибухове зростання генеративних інструментів штучного інтелекту (generative AI) . Менш ніж через рік після появи багатьох із цих інструментів одна третина респондентів цього опитування заявили, що їхні організації регулярно використовують штучний інтелект покоління принаймні в одній бізнес-функції. На тлі останніх досягнень ШІ піднявся з теми, яка була віднесена до технічних працівників, до сфери уваги керівників компаній: майже чверть опитаних керівників старших класів кажуть, що вони особисто використовують інструменти ШІ покоління для роботи. А більше чверті респондентів компаній, які використовують штучний інтелект, кажуть, що штучний інтелект вже є на порядку денному їхніх рад директорів. Більше того, 40 відсотків респондентів кажуть, що їхні організації загалом збільшать свої інвестиції в штучний інтелект завдяки прогресу в галузі штучного інтелекту.

Зазначені вище тенденції стрімко набирають сили, змушуючи задуматися про те, що чекає на світ науки та бізнесу попереду, наскільки це небезпечно та які є ризики.

Розвиток повнофункціонального штучного інтелекту в галузі кібербезпеки передбачає появу інтелектуального асистента, спеціально призначеного для професіоналів у цій галузі. Фахівці з тестування на проникнення та дослідники активно розробляють інструменти, засновані на генеративному штучному інтелекті (ГШІ), вносячи свій внесок у розвиток кібербезпечної спільноти. Ця тенденція продовжуватиметься і, ймовірно, призведе до створення нових інструментів, таких як асистент для фахівців з кібербезпеки на основі великих мовних моделей (LLM) або інших моделей машинного навчання. Цей асистент здатний виконувати різноманітні завдання в області тестування на проникнення, такі як пропозиція методів розвідки, ексфільтрації або підвищення привілеїв у ході атаки, а також частково автоматизувати подальші кроки мережі після первинного вторгнення. Отримавши контекст про виконані команди в тестовому середовищі, генеративний ШІ-бот зможе пропонувати рекомендації щодо дій. Він здатний аналізувати поточні результати та пропонувати наступні команди чи конкретні інструменти. Також він зможе виконувати запропоновані команди після схвалення користувачем. Такі рішення вже існують.

Шахраї все частіше вдаються до різних хитрощів, щоб обійти пильність своїх жертв та все частіше використовуватимуть нейронні мережі для створення зображень та відео. У наступному році ефективність подібних тактик може зрости. У сучасному цифровому світі існує безліч інструментів на базі штучного інтелекту, які дозволяють легко генерувати

реалістичні або візуально привабливі зображення, або навіть розробляти повноцінні лендінги. На жаль, цими інструментами можуть скористатися зловмисники для створення більш переконливого обману. В результаті спроби шахрайства можуть стати більш вдосконаленими, що призведе до збільшення числа атак та їхніх жертв.

Поширення різноманітних чат-ботів і великих мовних моделей, які спрощують працю представникам різних професій, викликає побоювання щодо конфіденційності та безпеки даних, на основі яких ці моделі навчаються. Це особливо актуально для великих корпорацій та інших організацій, які мають великі обсяги даних. Багато широко відомих попередньо навчених моделей LLM базуються на даних з відкритих джерел, що містять конфіденційну інформацію. При використанні корпоративних даних у цих моделях виникає ризик зловживання та невизначеності щодо збереження конфіденційності цих даних або їх використання для навчання моделі. У відповідь на ці побоювання очікуються нові тенденції на користь приватних великих мовних моделей (Private Large Language Models, PLLM), які навчаються на закритих даних, специфічних для конкретних організацій або галузей.

Крім захисту LLM, компанії усвідомлюють необхідність навчання своїх співробітників з питань безпечного використання популярних чат-ботів, таких як ChatGPT та Microsoft Copilot , а також інші інструменти, які використовують штучний інтелект. Це передбачає зростання попиту на спеціалізовані освітні курси, присвячені використанню штучного інтелекту.

У 2024 році очікується збільшення кількості ініціатив щодо регулювання штучного інтелекту. Ця активність спостерігатиметься на глобальному рівні і рухатиметься у двох основних напрямках. По-перше, прогнозується, що до цього процесу приєднається більше країн та міжнародних організацій, включаючи активну участь африканських та азіатських держав, незважаючи на те, що в цих регіонах ще не створено основ для внутрішнього регулювання ШІ. По-друге, країни та організації, які вже залучені до регулювання ШІ, будуть розширювати свою регулятивну базу, затверджуючи більш конкретні норми, що стосуються окремих аспектів , таких як створення навчальних наборів даних та використання персональних даних.

Недержавні компанії, особливо корпорації, відіграватимуть ключову роль у формуванні норм та практик, пов'язаних із штучним інтелектом. Маючи великий досвід у розробці та використанні ШІ, ці недержавні структури будуть робити неоціненний внесок в обговорення питань регулювання ШІ як на міжнародному, так і на національному рівні. Законодавці по всьому світу вже активно залучають бізнес та науковців у розробку стратегій регулювання у цій сфері, використовуючи їх великі знання та досвід.

Також, очікується збільшення кількості нормативних актів, що вимагають обов'язкового маркування контенту, створеного за участю штучного інтелекту. Постачальники цифрових сервісів впроваджуватимуть цю вимогу у свої політики та продовжуватимуть впроваджувати технології розпізнавання такого контенту. Розробники та дослідники братимуть активну участь у створенні методів маркування згенерованого ШІ-контенту, щоб полегшити його ідентифікацію та визначення джерела походження.

Організації, які вже впровадили можливості штучного інтелекту, першими дослідили потенціал штучного інтелекту покоління, а ті, хто бачить найбільшу користь від більш традиційних можливостей штучного інтелекту, вже випереджають інших у прийнятті інструментів штучного інтелекту покоління.

Також, корпорації продовжують отримувати прибуток у сферах бізнесу, у яких вони використовують ШІ, і планують збільшити інвестиції в наступні роки. Ми бачимо, що більшість респондентів повідомляють про збільшення доходу, пов’язаного зі штучним інтелектом, у межах кожної бізнес-функції за допомогою ШІ. Заглядаючи вперед, понад дві третини очікують, що їхні організації збільшать інвестиції в штучний інтелект протягом наступних трьох років.

З огляду на все вищевикладене вважається, що в найближчому майбутньому ландшафт загроз навряд чи суттєво зміниться. Незважаючи на активне освоєння нових технологій кіберзлочинцями, це малоймовірно призведе до радикальних змін в образі атак. У багатьох випадках технології досі недостатньо досконалі чи складні у застосуванні. В інших випадках автоматизація кібератак включає також автоматизацію тестів на проникнення, а більш ефективне створення шкідливого програмного забезпечення означає аналогічний виграш в ефективності у розробників антивірусного ПЗ, що компенсує ризики новими можливостями їх зниження.

Але все це ще раз підкреслює необхідність підвищення рівня обізнаності користувачів щодо кіберзагроз і використання надійних антивірусних програм, які блокують шахрайські листи та попереджують про підозрілі сайти.

Список використаних джерел:

1. The state of AI in 2023: Generative AI's breakout year (2023) https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year