Рішення
разової спеціалізованої вченої ради
про присудження ступеня доктора філософії

Здобувач ступеня доктора філософії **Неретін Олексій Сергійович**, 1983 року народження, громадянин України, освіта вища: у 2006 році закінчив Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут» і отримав повну вищу освіту за спеціальністю «Імпульсні теплові машини» та здобув кваліфікацію інженера-дослідника. Виконав акредитовану освітньо-наукову програму «Кібербезпека».

Разова спеціалізована вчена рада утворена наказом ректора Національного аерокосмічного університету «Харківський авіаційний інститут» Міністерства освіти і науки України, м. Харків, від «22» квітня 2026 року № 187 у складі (без змін):

голови разової спеціалізованої вченої ради – Лукіна Володимира Васильовича, доктора технічних наук, професора, завідувача кафедри інформаційно-комунікаційних технологій ім. О. О. Зеленського Національного аерокосмічного університету «Харківський авіаційний інститут»;

рецензентів – Брежнєва Євгена Віталійовича, доктора технічних наук, професора, професора кафедри кібербезпеки та інтелектуальних інформаційних технологій Національного аерокосмічного університету «Харківський авіаційний інститут»;

Тецького Артема Григоровича, кандидата технічних наук, доцента кафедри кібербезпеки та інтелектуальних інформаційних технологій Національного аерокосмічного університету «Харківський авіаційний інститут»;

офіційних опонентів – Яцківа Василя Васильовича, доктора технічних наук, професора, завідувача кафедри кібербезпеки Західноукраїнського національного університету;

Каштальян Антоніни Сергіївни, доктора технічних наук, доцента, професора кафедри комп'ютерної інженерії та інформаційних систем Хмельницького національного університету

на засіданні «23» червня 2026 року прийняла рішення про присудження ступеня доктора філософії з галузі знань 12 Інформаційні технології Неретіну Олексію Сергійовичу на підставі публічного захисту дисертації «Методи та засоби аналізу кібербезпеки і захисту великих мовних моделей від генерації забороненого контенту на локальних і хмарних серверах» за спеціальністю 125 Кібербезпека.

Дисертацію виконано в Національному аерокосмічному університеті «Харківський авіаційний інститут» Міністерства освіти і науки України, м. Харків.

Науковий керівник: Харченко Вячеслав Сергійович, член-кореспондент НАН України, доктор технічних наук, професор, завідувач кафедри кібербезпеки та інтелектуальних інформаційних технологій Національного аерокосмічного університету «Харківський авіаційний інститут».

Дисертацію подано у вигляді спеціально підготовленого рукопису, у якому відображено нові науково обґрунтовані результати проведених здобувачем досліджень, що виконують конкретне наукове завдання і мають вагоме значення для галузі знань 12 Інформаційні технології. Дисертація виконана державною мовою і відповідає встановленим МОН вимогам щодо оформлення дисертації. Обсяг основного тексту є достатнім для розкриття теми в межах галузі 12 Інформаційні технології за спеціальністю 125 Кібербезпека. Таким чином, у дисертації дотримано вимоги п. 6 Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу

вищої освіти, наукової установи про присудження ступеня доктора філософії, затвердженого постановою Кабінету Міністрів України від 12 січня 2022 року № 44 (зі змінами).

Здобувач має 11 наукових праць за темою дисертації, з яких 5 статей опубліковано в наукових фахових виданнях України, зокрема 1 – у виданні, внесеному до міжнародних наукометричних баз даних Scopus, 5 – у матеріалах національних та міжнародних наукових конференцій, 1 – розділ у колективній монографії.

Наукові праці, у яких висвітлено основні наукові результати дисертації:

1. Неретін О., Харченко В. Забезпечення кібербезпеки систем штучного інтелекту: аналіз вразливостей, атак і контрзаходів. *Вісник Національного університету "Львівська політехніка". Інформаційні системи та мережі*. 2022. Т. 12. С. 7–22. DOI: 10.23939/sisn2022.12.007.

2. Neretin O., Kharchenko V. A model of ensuring LLM cybersecurity. *Radioelectronic and Computer Systems*. 2025. Vol. 2025, no. 2. P. 201–215. DOI: 10.32620/reks.2025.2.13.

3. Neretin O., Kharchenko V. Information Technology for Assessing and Ensuring Cybersecurity of Large Language Models. *Security of Infocommunication5 Systems and Internet of Things*. 2025. Vol. 3, no. 2, paper 02020. P. 1–7. DOI: 10.31861/sisiot2025.2.02020.

4. Неретін О., Харченко В. Метод аналізу критичності вразливостей великих мовних моделей. *Measuring and computing devices in technological processes*. 2026. № 1. С. 443–450. DOI: 10.31891/2219-9365-2026-85-54.

5. Neretin O., Kharchenko V. IMECA method of risk-based assessment and ensuring cybersecurity of Large Language Models. *Herald of Advanced Information Technology*. 2026. Vol. 9, no. 1. P. 60–70. DOI: 10.15276/hait.09.2026.05.

У дискусії взяли участь (голова, рецензенти, офіційні опоненти, інші присутні) та висловили зауваження:

Рецензент Брежнєв Євген Віталійович:

1. Для більш глибокого оцінювання кібербезпеки ВММ могло бути використано математичні методи, методи теорії ігор (розділ 1.3.1, сторінка 39), які є більш точними ніж евристичні методи. Але автор обрав більш простий апарат теоретико-множинного опису (розділ 1.3.1, сторінка 39) та IMECA орієнтованого аналізу (розділ 1.3.2, сторінка 44), що є менш формалізованими та більш залежними від людського фактору.

2. В роботі не досить повно розкрито особливості генерації забороненого контенту на локальних та хмарних серверах. Не показано специфічні загрози та методи кіберзахисту ВММ можуть бути застосовані на локальних та хмарних серверах.

3. Робота обмежується суто статичними моделями та методами оцінювання безпеки ВММ систем без урахування часових характеристик різних типів кібератак. Вектори атак, вразливості, загрози та ризики постійно змінюються. Без урахування часу показники кібербезпеки ВВМ швидко стають неактуальними.

4. В роботі стверджується щодо результати можуть бути використані для системи моделей ВММ. Мова є про те, що багатоагентні системи (MAS) на основі ВММ дозволяють групам інтелектуальних агентів координувати та вирішувати складні завдання колективно у великих масштабах, переходячи від ізольованих моделей до підходів, орієнтованих на співпрацю. Вони можуть співпрацювати з точку зору кібербезпеки. Аспект взаємодії в роботі не врахований в повному обсязі.

Рецензент Тецький Артем Григорович:

1. В роботі використано визначення рівня тяжкості згідно до законодавства ЄС, але не розкрито питання стосовно можливості адаптації цього рівня під законодавства інших країн.

2. В роботі не надано аналізу ризиків щодо комбінованих атак на системи LLMs.

3. Автор обрав спрощений варіант вибору контрзаходів оскільки розглядає обмежену множину вразливостей та контрзаходів, а також варіантів їх покриття.

4. Не зважаючи на те, що в темі говориться про локальні та віддалені сервери, більша увага була приділена локальним серверам. Тож експериментального підтвердження результатів для віддалених серверів в роботі не надано.

Офіційний опонент Яцків Василь Васильович:

1. В дисертаційній роботі розглядаються питання захисту інформації на локальних та віддалених серверах. Зрозуміло, що там використовуються усталені методи захисту, однак, на мій погляд, доцільно було б надати рекомендації стосовно можливості використання криптографічних методів для захисту LLMs систем з урахуванням їх особливостей.

2. В роботі в явному вигляді не надається визначення та розрахунки для показника стійкості LLMs моделей до кібератак.

3. Автор не використовує моделі, які можуть оцінити зміни показників кібербезпеки в часі (з урахуванням часових рядів для кібератак).

Офіційний опонент Каштальян Антоніна Сергіївна:

1. В роботі не надано рекомендацій щодо використання LLMs моделей для підсилення захисту комп'ютерних систем і мереж.

2. З тексту дослідження залишається незрозумілим, який саме тип мовних моделей (відкриті або закриті) брався за основу. Потребує пояснення, чи аналізувався вплив архітектурного типу моделей на результати експериментів, зокрема в контексті генерації забороненого контенту.

3. В роботі розглянута вразливість статистично ймовірнісної генерації відповіді але, при цьому, моделі розроблені для множини вразливостей систем та моделей LLMs. Є певна суперечність в описі вразливостей, які враховуються.

4. У роботі зазначено, що методи оцінювання та забезпечення кібербезпеки мовних моделей можуть бути адаптовані до різних прикладних сфер, зокрема для безпілотних літальних апаратів. Разом з тим, цей напрям визначено у загальному вигляді (теоретично), без детального опису процесу цієї адаптації.

5. Експериментальна частина дослідження виконана виключно з використанням локально розгорнутих мовних моделей. У роботі не вистачає порівняльного аналізу використання розробленої методології оцінювання та забезпечення кібербезпеки моделей, розташованих на хмарних серверах. Відсутність інформації про специфіку експериментування у хмарах не дозволяє оцінити особливості функціонування запропонованих методів у контексті мережевих затримок, масштабування моделей та специфічних загроз безпеки хмарних моделей.

Результати відкритого голосування:

«За» 5 членів ради,

«Проти» 0 членів ради.

На підставі результатів відкритого голосування разова спеціалізована вчена рада присуджує Неретіну Олексію Сергійовичу ступінь доктора філософії з галузі знань 12 Інформаційні технології, за спеціальністю 125 Кібербезпека.

Відеозапис трансляції захисту дисертації додається.

Окрема думка члена разової ради не надходила.

Голова разової спеціалізованої вченої ради _____ Володимир ЛУКІН
(підпис)

Підпис голови разової спеціалізованої вченої ради Володимира ЛУКІНА засвідчую
Вчений секретар Національного аерокосмічного університету «Харківський авіаційний інститут» _____ Тетяна БОНДАРЄВА
(підпис)

М.П.