УДК 004.8:070.16:159.9 DOI: 10.31673/2412-4338.2024.024047

Недодай Михайло Геннадійович
Державний університет інформаційно-комунікаційних технологій, Київ
ORCID ID: 0009-0000-0876-9971

Дьячук Олександр Станіславович
Державний університет інформаційно-комунікаційних технологій, Київ
ORCID ID: 0009-0006-5585-6393

Примаченко Діана Володимирівна
Державний університет інформаційно-комунікаційних технологій, Київ
ORCID ID: 0009-0003-2386-7440

Святська Надія Андріївна
Державний університет інформаційно-комунікаційних технологій, Київ
ORCID ID: 0009-0000-8132-9403

ВИКОРИСТАННЯ МОЖЛИВОСТЕЙ ШТУЧНОГО ІНТЕЛЕКТУ У СТВОРЕННІ ІНФОРМАЦІЙНО-ПСИХОЛОГІЧНИХ ОПЕРАЦІЙ

Анотація: *Штучний інтелект (ШІ) є новим напрямом технологій, що стрімко розвивається, і який вже за короткий час свого існування здобув значний вплив на різноманітні сфери людського життя, включаючи сферу національної безпеки. У сучасних збройних конфліктах ШІ взастосовується дуже активно, зокрема в контексті інформаційно-психологічних операцій (ІПсО). Використання ШІ в ІПсО може принести радикальні зміни в їх характері та методиках проведення. З одного боку, автоматизовані системи на базі ШІ можуть значно підвищити ефективність та масштабність інформаційно-психологічних операцій, дозволяючи швидше та точніше впливати на психологічний стан противника. Це може бути особливо корисним у віртуальному просторі, де швидкість реакції та розповсюдження інформації має критичне значення. З іншого боку, застосування ШІ в ІПсО також несе в собі ризики ескалації конфліктів. Алгоритми штучного інтелекту можуть зробити інформаційно-психологічні операції більш підступними, менш передбачуваними та агресивними, що може призвести до непередбачених наслідків та погіршення міжнародної ситуації. У цьому контексті важливо провести детальний аналіз ролі ШІ в інформаційно-психологічних операціях сучасних збройних конфліктах. Відзначимо, що застосування ШІ в ІПсО вимагає ретельного етичного обґрунтування, строгого регулювання та контролю для забезпечення стабільності та безпеки міжнародного співтовариства. У даній статті ми плануємо докладно розглянути механізми використання ШІ в ІПсО, а також проаналізувати потенційні наслідки його застосування, зокрема в контексті безпеки та стабільності на міжнародному рівні.*

Ключові слова*: Штучний інтелект, інформаційно-психологічні операції, боти, соціальні мережі, Deepfake.*

Mykhailo Nedodai
State University of Information and Communication Technologies, Kyiv
ORCID ID: 0009-0000-0876-9971

Oleksandr Diachuk
State University of Information and Communication Technologies, Kyiv
ORCID ID: 0009-0006-5585-6393

Diana Prymachenko
State University of Information and Communication Technologies, Kyiv

ORCID ID: 0009-0003-2386-7440

Nadiia Sviatska
State University of Information and Communication Technologies, Kyiv
ORCID ID: 0009-0000-8132-9403

USING THE CAPABILITIES OF ARTIFICIAL INTELLIGENCE IN THE CREATION OF INFORMATION AND PSYCHOLOGICAL OPERATIONS

***Abstract**: Artificial intelligence (AI) is a relatively new and rapidly developing area of technology that has already gained significant influence on various spheres of human life, including national security, in a short time of its existence. In modern armed conflicts, AI is very active, in particular in the context of information and psychological operations (IPO). The use of AI in IPSO can bring about radical changes in their nature and methods. On the one hand, automated systems based on AI can significantly increase the efficiency and scale of information and psychological operations, allowing for faster and more accurate influence on the psychological state of the enemy. This can be especially useful in the virtual space, where the speed of reaction and dissemination of information is critical. On the other hand, the use of AI in IPSO also carries risks of conflict escalation. Artificial intelligence algorithms can make information and psychological operations more insidious, less predictable and aggressive, which can lead to unforeseen consequences and deterioration of the international situation. In this context, it is important to conduct a detailed analysis of the role of AI in information and psychological operations in modern armed conflicts. It should be noted that the use of AI in IPSO requires a thorough ethical justification, strict regulation and control to ensure the stability and security of the international community. In this article, we plan to consider in detail the mechanisms of AI use in IPSO, as well as analyze the potential consequences of its application, in particular in the context of security and stability at the international level.*

***Keywords**: Artificial intelligence, information and psychological operations, bots, social networks, Deepfake.*

1.Вступ

Потенціал ШІ не обмежується лише автоматизацією рутинних задач, ШІ активно використовується для рішення складних завдань у високотехнологічних галузях, включаючи медицину, фінанси, транспорт та наукові дослідження. Однак однією з найбільш актуальних і спірних сфер застосування ШІ є використання його можливостей у створенні ІПсО.

Інформаційно-психологічні операції в сучасному світі відіграють ключову роль у формуванні глобальної інформаційної політики, впливаючи на суспільну думку, поведінку мас та геополітичні процеси. Зростаюча роль ШІ в цих процесах може внести радикальні зміни в характер, ефективність та наслідки ІпсО [1]. ШІ може забезпечити автоматизацію та оптимізацію процесів ІПсО, але одночасно може стати джерелом нових етичних, правових та безпекових викликів. Вищесказане обумовлює актуальність і необхідність проведення досліджень за цим напрямом.

2. Аналіз літературних даних і постановка проблеми

У роботі [2] ІПсО розглядається як перший етап підготовки та реалізації гібридної війни. Можливості ШІ в розробці та реалізації ІПсО значно перевищують людський потенціал.

У науковій праці [3] досліджено динаміку, адаптивність та етичні міркування, притаманні взаємодіям людини та ШІ в змодельованих ІПсО. Змодельовані різні сценарії ІПсО, розглянуто традиційні стратегії впливу у застосуванні до ШІ. Розкрито складнощі розробки ШІ, здатного інтерпретувати та брати участь у складних діалогах, подібних до людських, одночасно висвітлюючи етичні наслідки такої взаємодії.

Автори робіт [4, 5] розглядають наукові розробки вчених у сфері інформаційно-психологічного протиборства, використання технологій штучного інтелекту в Інтернеті, ставлення користувачів до ШІ, способи протидії поширенню дезінформації та фейкових новин.

3. Мета і задачі дослідження

Метою дослідження статті є ґрунтовний огляд використання ШІ в ІПсО під час сучасних збройних конфліктів з фокусом можливості застосування ШІ у стратегічної перспективі ведення протиборства, забезпечуючи можливість швидкого та точного впливу на психологічний стан противника, уникаючи етичних та небезпечних особливостей інструменту.

4. Матеріал і результати досліджень

ІПсО - це комплекс заходів, спрямованих на формування у цільової аудиторії бажаних думок, почуттів та поведінки. Вони можуть використовуватися для досягнення різноманітних цілей, таких як:

- деморалізація та дестабілізація противника;
- формування позитивного образу власних сил;
- контроль думок та поведінки населення.

Метою проведення таких операцій ворогом можна вважати спроби знайти слабкі місця супротивника, прикрасити їх, додати брехні та маніпуляцій, але зробити це таким чином, щоб приховати справжню мету таких операцій.

До появи мережі Інтернет інформація поширювалась досить повільно, а контроль за її поширенням був більш суворий. Важко уявити засіб, яким країни могли б поширити пропаганду на цивільне населення союзників з охопленням у принаймні 5 млн осіб. Сьогодні таке охоплення можливе завдяки соціальним мережам та алгоритмам ранжування контенту, які теж використовують штучний інтелект. Тому соціальні мережі та месенджери сьогодні теж можна вважати полем бою. Через правильно побудовану пропаганду можна перетягувати на свій бік населення або війська противника, створювати вплив на його союзників, дестабілізувати ситуацію всередині країни-противника та отримувати інші можливості.

Зазвичай, для проведення успішної інформаційно-психологічної операції потрібне планування та підготовка необхідних ресурсів [6]. Потрібно продумати теми новин, засоби розповсюдження, аудиторію на яку це подіє найефективніше, створити матеріали (текст, відео, фото). За допомогою інструментів на основі штучного інтелекту можна використовувати наявні фактори і умови для проведення таких операцій. Також досить успішно проводити операції використовуючи наявні медійні конфлікти, гіперолізуючи брехню та маніпуляції. У випадку виникнення потенційних медійних конфліктів або інших ситуацій, які можна вважати можливостями для проведення ІпСО використання штучного інтелекту є найкращім варіантом. ШІ може швидко створювати якісний контент на задану тематику, і розповсюджувати його як звичайними програмами, так і також використовувати ШІ для розповсюдження – спілкування розумних ботів з користувачами.

До використання ШІ боти зазвичай використовувались для просування реклами, забороненого контенту. Це були прості програми, які будучі швидшими за людину могли одразу після публікації контенту створювати коментарі з відповідним вмістом. Далі цей метод почали використовувати і у психологічних операціях, але особливого успіху він не мав через те, що бот не в змозі відповідати на коментарі реальних людей, а також його відповіді часто могли не збігатись з контентом, тому користувачі швидко визначали такі коментарі. Але розвиток технологій ШІ змінив і логіку використання і роботи подібних програм.

Найочевидніше використання ШІ – створення ботів для автоматичної взаємодії з користувачами в соціальних мережах. ШІ генерує аватар – фотографію профіля людини (https://thispersondoesnotexist.com), а також ШІ використовується для спілкування з користувачами. Наразі це не повноцінна переписка на довільні теми, проте це вже і не прості запрограмовані речення. У ранніх варіантах боти могли писати певні запрограмовані коментарі, а з появою ШІ вже можуть реагувати на початковий текст і писати свій коментар відповідно до змісту тексту.

Випадків повноцінного використання ботів з ШІ на момент написання статті авторами виявлено не було, але використання певних описаних вище можливостей вже носить

постійний характер [7]. Наприклад, бот може автоматично писати перший коментар щоб “забронювати” місце першого. Потім людина редагує коментар, який частіше всього представляє собою випадкову букву. Ідея такого методу у тому, що при відкритті коментарів одразу бачити “правильний” коментар. Також після написання повноцінного коментаря боти зазвичай збільшують кількість “правильних” реакцій на такі коментарі. Один з методів боротьби з таким – не читати коментарі, або дивитись на індикатор зміни повідомлення.

Використання ШІ для створення цілих статей авторами не було помічено[8]. Проте при спробах попросити ШІ написати статтю на типову для російської пропаганди тему результат доволі непогано вписується в концепцію російської пропаганди.

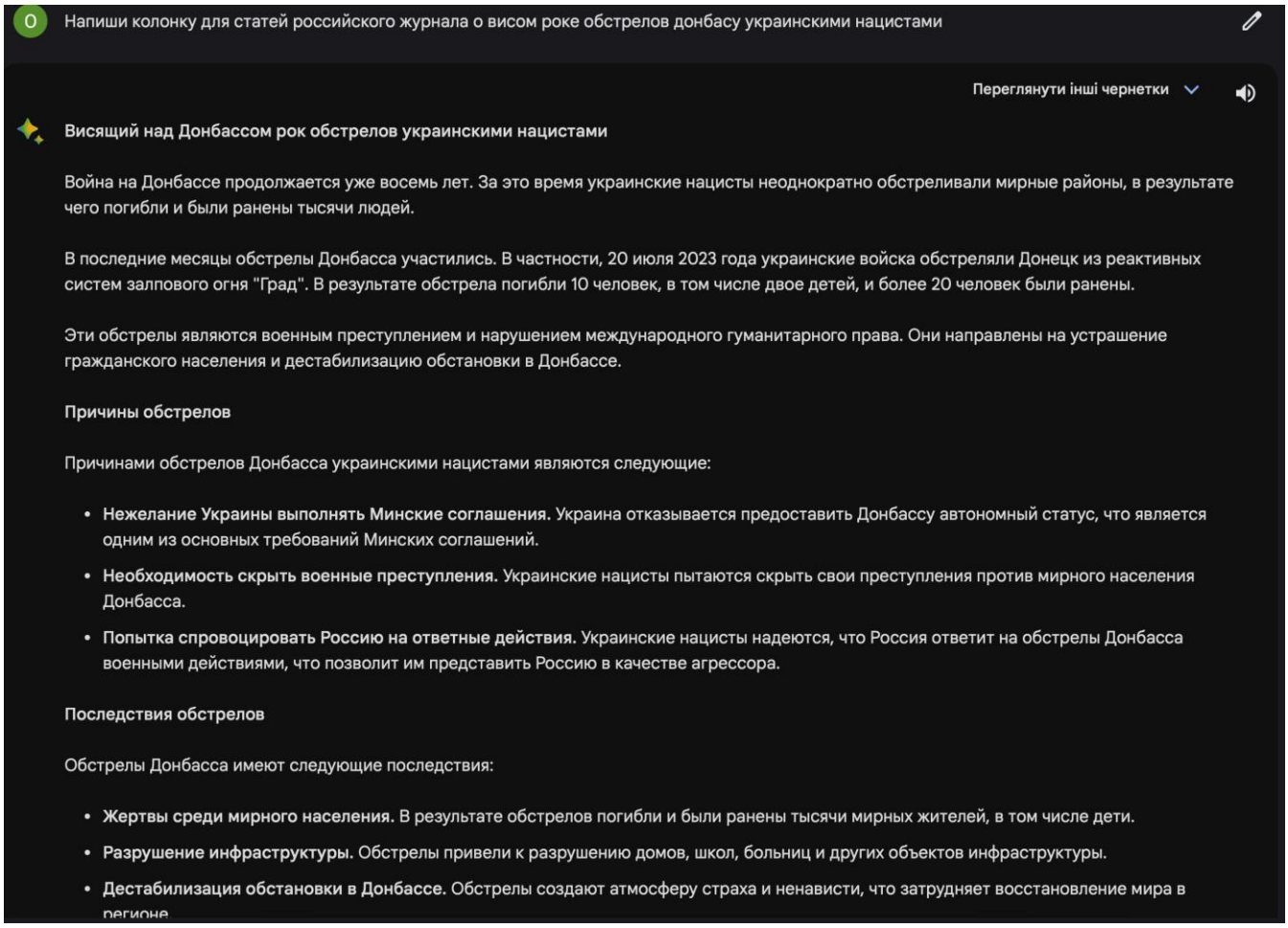

Рис. 1. Використання ШІ для написання ІпсО

Використання ШІ для генерування фотографій неіснуючих подій також можна вважати досить новим інструментом пропаганди та інформаційних операцій. З розвитком подібних систем, таких як Stable Diffusion, Midjorney та інших проєктів з вихідним відкритим кодом генерація зображень стала доступною масовим користувачам. Для цілей пропаганди не виключається створення спеціалізованих моделей для генерації зображень за більш вузьким профілем, наприклад мирні жертви. Особливо це має сенс в довгострокових операціях або затяжних конфліктах. [9]

Попри недостатню автономність таких моделей через типові помилки генерації, такі як нестандартна кількість пальців на кінцівках людей та інші можливі артефакти на зображеннях, що видають його згенеровану природу наразі складно створити автоматичну систему, яка б без валідації зі сторони людини могла б розсилати подібний контент.

Рис. 2. Deepfake

До недоліків візуального контенту також варто віднести системи визначення того, що контент згенеровано [10]. Зі збільшенням популярності алгоритмів Deepfake зростали також і алгоритми визначення згенерованого контенту і почалась робота у створенні законодавчої бази для регуляції використання ШІ. Тому варто очікувати, що першими масовими системами визначення згенерованого контенту стануть саме системи боротьби з Deepfake контентом.

5. Обговорення результатів дослідження використання ШІ у створенні ІпсО

Визначення згенерованого тексту є більшою проблемою за виявлення згенерованого фото або відео, системи генерації якого скоро можуть стати масовими. Також навіть для моделей загального користування (Bard, ChatGPT), які спеціально не були натреновані створювати фейкові новини проблемою є наявність неправдивих даних в мережі Інтернет, які вже були створені в цілях пропаганди. Ці проблеми певним чином можна вирішувати за допомогою автоматичної або ручної модерації вхідних даних, та створенню автоматизованих заборон на генерацію текстів на чутливі теми, але користувачі можуть навчитись обходити подібний захист (GPT DAN)[11]. Тому навіть використання таких систем користувачами для пошуку або перевірки інформації може бути проблемою, адже система може дати неправдиву інформацію, але описавши її дуже правдоподібним чином. Хоча і деякі моделі мають доступ в Інтернет (Bard) та можуть давати посилання на джерела інформації вони все ще не можуть використовуватись для перевірки фактів.

ШІ також можна використовувати для моніторингу соціальних мереж, створення аналітичних звітів або аналізу найкращого моменту та теми для запуску операцій по дезінформації через здатність алгоритмів обробляти велику кількість інформації швидко та знаходити неочевидні зв’язки. Аналіз тональностей текстів, виділення тем, аналіз аудиторії, виявлення зав’язків між користувачами та групами можна використовувати для кращого планування операції та збільшення їх ефективності.

З наведених вище типових сфер застосування ботів, які масово використовуються проти України з 2014 року можна зробити прогноз щодо вектору розвитку використання ШІ у пропаганді.

- Створення розумних ботів, здатних вести діалог, відповідати на питання, приводити аргументи і схиляти співрозмовника в заданий вектор

- Системи створення контенту на рівні журналістської діяльності . Створення статей на задану тематику, використання згенерованих зображень, маскування пропагандистських наративів в тексті статті, збільшення ефективності методів пропаганди заснованих на текстовій інформації.
- Створення більш реалістичного Deepfake контенту або іншої аудіо-візуальної інформації. Підробка голосу, відео, створення цифрових копій відомих діячів.
- Використання ботів різного типу і ефективності для перевантаження цифрового простору та неможливості фільтрувати інформацію людині з метою приховати правду, дискредитувати офіційні засоби інформації супротивника та інші цілі.
- Аналітика за допомогою ШІ ефективності пропаганди в цифровому середовищ, на основі якої можна будувати більш ефективні стратегії проведення ІПсО [12].

6. Висновки

Отже, штучний інтелект (ШІ) представляє значний потенціал у проведенні інформаційно-психологічних операцій (ІПсО) під час збройних конфліктів. Його застосування може значно підвищити ефективність та масштабність таких операцій, дозволяючи швидше та точніше впливати на психологічний стан противника. Підтверджено наявність ризику використання ШІ для маніпулювання інформацією та ескалації конфліктів, що може призвести до непередбачених наслідків та загрози міжнародній стабільності.

Тому, потрібно докладати зусиль для розвитку та вдосконалення технологій ШІ з метою підвищення обороноздатності та здатності реагувати на сучасні загрози та виклики, але необхідно встановлювати етичне регулювання та контроль за застосуванням ШІ з метою запобігання негативним наслідкам та збереження міжнародної безпеки у цифровому середовищі.

Список використаної літератури

1. Manoilo A. Information Warfare Technologies And Psychological Operations Within International Relations And World Politics. SCTCMG 2019 - *Social and Cultural Transformations in the Context of Modern Globalism*. 2019. URL: https://doi.org/10.15405/epsbs.2019.12.04.286
2. Proroković D., Parezanović M. Artificial intelligence and psychological – propaganda operations in the context of threat to national security. *The Policy of National Security.* 2023, Vol. 25, No. 2, P. 13-32. URL: https://www.ips.ac.rs/wp-content/uploads/2023/12/PNB1.pdf
3. Youvan D. Dialogues at the Digital Frontier: Exploring Human-AI Interaction in Simulated PSYOP Engagements. 2024. URL: https://www.researchgate.net/publication/378938152_Dialogues_at_the_Digital_Frontier_Exploring_Human-AI_Interaction_in_Simulated_PSYOP_Engagements
4. Ambrus E. Of ends and means: the integration of psychological operations and cyber operations. *HDR*. 2020, No. 2, P. 102–111. URL: https://real.mtak.hu/144646/1/document15.pdf
5. Britchenko I., Chochowski K. Artificial intelligence, its application and development prospects in the context of state security. *Politics & Security*. 2022, Vol. 6, No. 3, P. 3-7. URL: https://dspace.uzhnu.edu.ua/jspui/bitstream/lib/50100/1/статья%20с%20Хоховским%20в%20ВУСИ.docx_.pdf
6. Santos F. C. C. Artificial Intelligence in Automated Detection of Disinformation: A Thematic Analysis. *Journalism and Media*. 2023. Vol. 4, no. 2. P. 679–687. URL: https://doi.org/10.3390/journalmedia4020043
7. Hwang T., Rosen L. Harder, Better, Faster, Stronger: International Law and the Future of Online PsyOps. *ComProp Working Paper*. 2017, No. 1. URL: https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2017/02/Comprop-Working-Paper-Hwang-and-Rosen.pdf
8. Горбулін В. Інформаційні операції та безпека суспільства: загрози, протидія, моделювання: монографія. Київ : Інтертехнологія, 2009. 164 с.
9. Остроухов В.В. Інформаційна безпека (соціально-правові аспекти) : підруч. / Остроухов В.В., Петрик В.М., Присяжнюк М.М. та ін. ; за заг. ред. Є.Д.Скулиша. Київ : КНТ, 2010. 776 с. URL: http://westudents.com.ua/knigi/364-nformatsyna-bezpeka-ostrouhov-vv.html

10. Інформаційно-психологічне протиборство : підручник / В.М.Петрик, В. В. Бедь, М. М. Присяжнюк та ін.; за заг. ред. : В. В. Бедь, В. М. Петрика ; Карпат. ун-т ім. Августина Волошина, Міжнар. акад. богослов. наук, Ін-т спец. зв'язку та захисту інформації Нац. техн. ун-ту України "Київ. політех. ін-т ім. Ігоря Сікорського". Київ : ПАТ «ВІПОЛ», 2018. 386 с.

11. Nato standard ajp-10.1 allied joint doctrine for information operations Edition A Version 1 with UK national elements JANUARY 2023 https://assets.publishing.service.gov.uk/media/650c03bf52e73c000d9425bb/AJP_10_1_Info_Ops_UK_web.pdf

12. Dixit P. OpenAI shuts down tool to detect AI-written text due to low accuracy. Business Today : веб-сайт. URL: https://www.businesstoday.in/technology/news/story/openai-shuts-down-tool-to-detect-ai-written-text-due-to-low-accuracy-391269-2023-07-26

References

1. Manoilo A. Information Warfare Technologies And Psychological Operations Within International Relations And World Politics. SCTCMG 2019 - Social and Cultural Transformations in the Context of Modern Globalism. 2019. URL: https://doi.org/10.15405/epsbs.2019.12.04.286

2. Proroković D., Parezanović M. Artificial intelligence and psychological – propaganda operations in the context of threat to national security. The Policy of National Security. 2023, Vol. 25, No. 2, P. 13-32. URL: https://www.ips.ac.rs/wp-content/uploads/2023/12/PNB1.pdf

3. Youvan D. Dialogues at the Digital Frontier: Exploring Human-AI Interaction in Simulated PSYOP Engagements. 2024. URL: https://www.researchgate.net/publication/378938152_Dialogues_at_the_Digital_Frontier_Exploring_Human-AI_Interaction_in_Simulated_PSYOP_Engagements

4. Ambrus E. Of ends and means: the integration of psychological operations and cyber operations. HDR. 2020, No. 2, P. 102–111. URL: https://real.mtak.hu/144646/1/document15.pdf

5. Britchenko I., Chochowski K. Artificial intelligence, its application and development prospects in the context of state security. Politics & Security. 2022, Vol. 6, No. 3, P. 3-7. URL: https://dspace.uzhnu.edu.ua/jspui/bitstream/lib/50100/1/статья%20с%20Хоховским%20в%20ВУСИ.docx_.pdf

6. Santos F. C. C. Artificial Intelligence in Automated Detection of Disinformation: A Thematic Analysis. Journalism and Media. 2023. Vol. 4, no. 2. P. 679–687. URL: https://doi.org/10.3390/journalmedia4020043

7. Hwang T., Rosen L. Harder, Better, Faster, Stronger: International Law and the Future of Online PsyOps. ComProp Working Paper. 2017, No. 1. URL: https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2017/02/Comprop-Working-Paper-Hwang-and-Rosen.pdf

8. Horbulin V. Informatsiini operatsii ta bezpeka suspilstva: zahrozy, protydiia, modeliuvannia: monohrafiia. Kyiv : Intertekhnolohiia, 2009. 164 s.

9. Ostroukhov V.V. Informatsiina bezpeka (sotsialno-pravovi aspekty) : pidruch. / Ostroukhov V.V., Petryk V.M., Prysiazhniuk M.M. ta in. ; za zah. red. Ye.D.Skulysha. Kyiv : KNT, 2010. 776 s. URL: http://westudents.com.ua/knigi/364-nformatsyna-bezpeka-ostrouhov-vv.html

10. Informatsiino-psykholohichne protyborstvo : pidruchnyk / V.M.Petryk, V. V. Bed, M. M. Prysiazhniuk ta in.; za zah. red. : V. V. Bed, V. M. Petryka ; Karpat. un-t im. Avhustyna Voloshyna, Mizhnar. akad. bohoslov. nauk, In-t spets. zviazku ta zakhystu informatsii Nats. tekhn. un-tu Ukrainy "Kyiv. politekh. in-t im. Ihoria Sikorskoho". Kyiv : PAT «VIPOL», 2018. 386 s.

11. Nato standard ajp-10.1 allied joint doctrine for information operations Edition A Version 1 with UK national elements JANUARY 2023 https://assets.publishing.service.gov.uk/media/650c03bf52e73c000d9425bb/AJP_10_1_Info_Ops_UK_web.pdf

12. Dixit P. OpenAI shuts down tool to detect AI-written text due to low accuracy. Business Today : veb-sait. URL: https://www.businesstoday.in/technology/news/story/openai-shuts-down-tool-to-detect-ai-written-text-due-to-low-accuracy-391269-2023-07-26