УДК 342.52

МАРУЩАК А.І., *доктор юридичних наук, професор,* ГО "Міжнародна академія інформації", професор НА СБ України.
ORCID: https://orcid.org/0000-0003-0069-3727.
ПЕТРОВ С.Г., *доктор юридичних наук,* провідний науковий співробітник Інституту спецзв'язку та захисту інформації НТУУ "КПІ імені Ігоря Сікорського".
ORCID: https://orcid.org/0000-0001-7786-4657.

ПРАВОВІ МЕХАНІЗМИ ЄС ТА США ДЛЯ ВИЯВЛЕННЯ І БЛОКУВАННЯ РОСІЙСЬКОЇ ДЕЗІНФОРМАЦІЇ У СОЦІАЛЬНИХ МЕРЕЖАХ

Анотація. *У статті здійснено аналіз правових механізмів ЄС та США задля виявлення і блокування російської дезінформації у соціальних мережах. Зроблено висновок, що ЄС посилює заходи правового реагування на протиправні дії щодо дезінформації, яка поширюється російськими державними органами. Зазначено, що положення щодо дезінформації, які містяться у Кодексі практики щодо дезінформації ЄС, Європейському акті про свободу медіа, Європейському законі про штучний інтелект мають бути відображені в українському законодавстві. Умови використання великих онлайн платформ розроблені на основі ліберального підходу щодо свободи слова та боротьби з поширенням дезінформації у США. Однак наведені приклади видалення дезінформаційного контенту приватними компаніями, розташованими в США, засвідчили необхідність розробки правових механізмів безпосереднього зв'язку із великими онлайн платформами задля протидії російській дезінформації.*

Ключові слова: *інформаційне право, дезінформація, законодавство ЄС, штучний інтелект, правове регулювання.*

Summary. *The article is devoted to the EU and US legal mechanisms for detecting and blocking russian disinformation in social networks. It was concluded that the EU is strengthening measures of legal response to illegal actions regarding disinformation spread by russian government. It is noted that the provisions on disinformation contained in the EU Code of Practice on Disinformation, the European Act on Media Freedom, and the European Act on Artificial Intelligence should be reflected in Ukrainian legislation. The terms of use of major online platforms are designed based on a liberal approach to freedom of speech and the fight against the spread of misinformation in the United States. However, the given examples of the removal of disinformation content by private companies located in the USA proved the need to develop legal mechanisms for direct communication with large online platforms to counter Russian disinformation.*

Keywords*: Information Law, Disinformation, EU Legislation, Artificial Intelligence, Legal Regulation.*

Постановка проблеми. За останнє десятиліття ІТ-технології розвивалися з неймовірною швидкістю. Зміни принесли переваги окремим особам і державам щодо економічного і політичного розвитку, однак вони також спричиняють масштабні маніпуляції інформацією. Зловмисники і держава-терорист систематично поширюють дезінформацію, часто через соціальні мережі, щоб дестабілізувати суспільства, втручатися в державне управління, тероризувати та радикалізувати населення.

Україна, маючи статус кандидата на вступ до ЄС, активно гармонізує своє законодавство із законодавством та правилами ЄС, зокрема щодо боротьби з дезінформацією. Необхідність правового регулювання щодо нейтралізації цієї загрози

© Марущак А.І., Петров С.Г., 2024

базується на останніх рішеннях ЄС. Наприклад, кампанії дезінформації/операцій впливу включено до списку нових загроз кібербезпеці, які мають вплив до 2030 року в ЄС [1].

Крім того, 21 травня 2024 року Рада ЄС схвалила два документи, що стосуються управління дезінформацією – Майбутнє цифрової політики ЄС [2] і Висновки Ради щодо демократичної стійкості: захист виборчих процесів від іноземного втручання [3]. Перший документ має на меті створити основу для розробки цифрової політики на наступні п'ять років, а дезінформація входить до списку шкідливих або незаконних явищ, з якими необхідно боротися. Запобігання зовнішньому втручанню у виборчі процеси є основними темами другого документа. Він також містить повний огляд законодавчих, підзаконних та інституційних інструментів, створених ЄС.

Однак складність міжнародного та внутрішнього законодавства щодо дезінформаційних операцій іноземних держав, насамперед росії, пов'язана як зі складністю правового регулювання кібероперацій, так і з міжнародно-правовим регулюванням свободи слова.

Негативні наслідки, які несе для українського суспільства російська дезінформація, вимагають аналізу теоретичних основ і практики розвитку правових механізмів ЄС та США щодо виявлення і блокування російської дезінформації у соціальних мережах.

Результати аналізу наукових публікацій свідчать про те, що подібні питання регулювання нових інформаційних технологій і протидії дезінформації були частково предметом досліджень українських учених, а саме Ю. Горбань, О. Олійник, В.Г. Пилипчука, В.І. Польового, О.Б. Розвадовського, Т.Ю. Ткачука, О.М. Юрченка та інших.

Зокрема, Ю. Горбань і О. Олійник розглядають питання медіаграмотності у контексті захисту інформаційного простору від дезінформації ворога у воєнний час [4]. Зарубіжними вченими досліджуються питання викликів використання штучного інтелекту (далі – ШІ) через створення соціальних ботів та поширення дезінформації в соціальних мережах [5]. Проблемам впливу законів про дезінформацію на поняття правди на цифрових платформах приділяли увагу П. Кавальєр [6], щодо дезінформування з використанням ШІ – Г. Спітале, Н. Біллер-Андорно, Ф. Германі [7], а також з точки зору проектування проти дезінформації Дж. Сміт [8].

Однак у цілому питання розкриття правових механізмів ЄС та США для виявлення і блокування російської дезінформації у соціальних мережах було предметом наукових досліджень лише фрагментарно.

Метою статті є розкриття науково-практичних підходів ЄС та США щодо впровадження правових механізмів виявлення і блокування російської дезінформації у соціальних мережах.

Виклад основного матеріалу. Аналіз практики ініціатив ЄС та провідних компаній світу щодо механізмів блокування і видалення дезінформації засвідчує, що одним із найперших прикладів врегулювання відповідних відносин в ЄС став Кодекс практики щодо дезінформації ЄС. Кодекс практики був спочатку підписаний Facebook, Google, а також Twitter, Mozilla, суб'єктами рекламної індустрії, Microsoft і TikTok [9]. І хоча Кодекс не мав нормативно-правового характеру, у ньому підписанти визнали "фундаментальне право на свободу вираження поглядів і на відкритий Інтернет" [9]. Особлива увага в Кодексі приділяється прецедентному праву Суду Європейського Союзу (СЄС) щодо пропорційності заходів, спрямованих на обмеження доступу та розповсюдження шкідливого контенту [9].

Деякі норми Європейського акту про свободу медіа (European Media Freedom Act, далі – EMFA) [10] також можуть бути гармонізовані з українським законодавством. Як

відзначає М. Гаміто, до введення в дію EMFA "свобода онлайн-медіа" регулювалася всередині держав-членів ЄС через загрози дезінформації, без урахування внутрішнього європейського ринку [11]. З точки зору України, особливо цікавим буде надання Європейській Раді з медіа-послуг (далі – Рада) повноважень брати участь у діалозі з великими онлайн-платформами, щоб контролювати дотримання ініціатив саморегулювання, спрямованих на захист користувачів від шкідливого контенту, зокрема протидії зовнішньому маніпулятивному впливу. EMFA також підкреслює "недостатність інструментів для регулятивної співпраці між національними регуляторними органами чи органами" [10].

Посилаючись на Кодекс практики ЄС щодо дезінформації, EMFA зобов'язує Раду організувати структурований діалог між великими онлайн-платформами, представниками постачальників медіа-послуг та представниками громадянського суспільства. Це має на меті сприяти доступу до різноманітних пропозицій незалежних медіа, у тому числі щодо процесів модерації контенту великих онлайн-платформ, і моніторингу дотримання ініціатив саморегулювання для захисту користувачів від шкідливого контенту, включаючи дезінформацію [10]. Імплементація цих норм в українське законодавство має життєво важливе значення для протидії російській дезінформації на великих онлайн-платформах.

Іншим важливим документом є Європейський закон про штучний інтелект (далі – Акт про ШІ) [12], який містить кілька положень щодо дезінформації, що мають бути відображені в українському законодавстві. Акт про ШІ безпосередньо стосується системних ризиків, створених моделями ШІ загального призначення. Ці ризики включають ризики від сприяння дезінформації. Юридично значущим є визначення "глибинних фейків" (від англ. – *deepfakes*) у Акті про ШІ як зображення, аудіо- чи відеоконтент, створений або оброблений ШІ, який нагадує існуючих осіб, об'єкти, місця чи інші об'єкти чи події та може здаватися людині автентичним або правдивим [12, с. 110].

Акт про ШІ визначає зобов'язання [12, с. 120], покладені на постачальників і розробників певних систем ШІ щодо розкриття штучності створення результатів цих систем, зокрема щодо зобов'язань постачальників великих онлайн-платформ або дуже великих онлайн-пошукових систем щодо виявлення розповсюдження маніпулятивного контенту, зокрема дезінформації. Крім того, Акт про ШІ передбачає, що розповсюджувачі, які використовують систему ШІ для створення "глибинних фейків", також повинні розкривати, що вміст було штучно створено [12, с. 134]. Дотримання цього зобов'язання не слід тлумачити як вказівку на те, що використання системи або її вихідних даних перешкоджає праву на свободу вираження поглядів і праву на свободу мистецтва та науки. Вимога щодо маркування вмісту, створеного системами ШІ, не шкодить зобов'язанню статті 16 (6) Регламенту (ЄС) 2022/2065 – постачальники послуг хостингу повинні обробляти будь-які повідомлення, які вони отримують відповідно до механізмів, що дозволяють будь-якій фізичній чи юридичній особі сповіщати про наявність у їхніх послугах елементів інформації, які фізична чи юридична особа вважає незаконним вмістом [12, с. 136].

Ці норми не будуть обов'язковими для України та діяльності великих онлайн-платформ у нашій країні, поки Україна не стане членом ЄС. Таким чином, норми мають бути відображені в українських нормативних актах щодо цифрових послуг та ШІ.

Варто відзначити і Заяву Високого представника від імені ЄС щодо триваючої гібридної діяльності росії проти ЄС та його держав-членів (далі – Заява), яка була оприлюднена 8 жовтня 2024 року Прес-службою Генерального секретаріату Ради ЄС

[13]. У заяві повідомлено про виявлену зростаючу кількість дій проти ЄС та його держав-членів, включаючи інформаційні маніпуляції та кампанії втручання. Ці зловмисні дії є частиною широкої скоординованої гібридної кампанії, спрямованої росією на розділ суспільства, дестабілізацію та послаблення ЄС та його держав-членів, а також на підрив підтримки України та її здатності захищатися.

З огляду на предмет даної роботи, звернемо увагу, що у Заяві зазначено, що згідно з новою законодавчою базою, ЄС може переслідувати тих, хто відповідальний та реалізує, підтримує або отримує вигоду від дестабілізуючих дій росії в усьому світі, а також їхніх спільників і прихильників [13].

Умови використання великих онлайн-платформ розроблені на основі правової системи країни походження, переважно США. Варто звернути увагу на нормативні акти США, щоб з'ясувати принципи, які використовують такі платформи під час боротьби з дезінформацією.

Законодавство США щодо дезінформації почало зароджуватися у формі Закону про підзвітність *deepfakes*, запровадженого в червні 2019 року для боротьби з поширенням такого виду дезінформації [14]. Крім того, США запровадили відповідне правове регулювання Законом про алгоритмічну звітність [15].

Проте Конституція США та прецедентне право не були особливо послідовними у застосуванні та тлумаченні правових можливостей щодо обмежень свободи слова. Перша поправка до Конституції США (щодо свободи слова) поширюється лише на закони, прийняті Конгресом, а також на місцеві, державні чи федеральні урядові установи, але не на дії приватних великих онлайн-платформ. Таким чином, відповідальність таких компаній щодо свободи слова та боротьби з поширенням дезінформації визначається переважно корпоративною політикою. Правовий підхід США досі базується на саморегулюванні.

Законодавство США про боротьбу з дезінформацією розвиває свободу слова та свободу підприємницької діяльності великих онлайн-платформ без належного врахування інтересів національної безпеки. Наприклад, Закон про заборону Ради з управління дезінформацією 2023 року припинив діяльність Ради з управління дезінформацією Міністерства внутрішньої безпеки США [16]. Крім того, Закон про захист свободи слова 2023 року забороняє федеральним службовцям і підрядникам керувати онлайн-платформами або цензурувати будь-які виступи, які захищені Першою поправкою до Конституції США [17].

Таким чином, пропозиції щодо законодавчих шляхів боротьби з дезінформацією на великих онлайн-платформах наражаються на суспільну дискусію щодо доцільності такого регулювання. Складність проблеми зумовлена також розвитком можливостей для дезінформації, викликаних технологіями, насамперед згаданими вище "глибинними фейками", створеними за допомогою ШІ. Прогалина у законодавчих і нормативних вимогах США щодо відповідальності великих онлайн-платформ впливає на національні інтереси інших демократичних держав.

Через відсутність двосторонніх інструментів урядової співпраці між США та Україною, українському уряду необхідно розробити правові механізми для прямого і безпосереднього зв'язку з великими онлайн-платформами, розташованими в США, задля протидії російській дезінформації. У цьому контексті варто звернути увагу і на принципи саморегулювання ШІ, добровільно прийняті та запроваджені в 2023 році декількома розробниками ШІ в США [18].

Звернемо увагу і на нещодавні приклади видалення дезінформаційного контенту приватними компаніями, розташованими в США. Так, Meta (Facebook) у серпні 2024

року оприлюднила Звіт про підривні загрози [19]. У ньому зокрема повідомляється про видалення 43 облікових записів Facebook і 85 сторінок за порушення політики щодо координації неавтентичної поведінки мережі, яка створена в росії і націлена, насамперед, на Україну та Молдову, а також українців, які проживають у Європі.

За операціями впливу в росії продовжують з'являтися нові комерційні структури. Це кампанії, керовані підрядниками (а не спеціальними службами, як було у минулому), які виконують низькоякісні, великі за обсягом роботи із залученням мережі тролів [19].

Окремі виявлені Meta підривні кампанії виглядають як невдалі спроби "тривимірних шахів", які підривають Україну, вдаючи, що підтримують її. З одного боку вони припускали, здавалося б проукраїнську позицію, закликаючи отримати більше зброї з Європи, але також висвітлювали українські бойові втрати, закликали до суворішого покарання українців, які ухиляються від призову, та вимагали проведення виборів Президента України в 2024 році. Ця операція була пов'язана з російською війною проти України і об'єднувала онлайн зусилля зі створення фіктивних новин і громадських організацій в Інтернеті щодо реального життя у Франції, Німеччині та Польщі. У Facebook ця операція почалася зі створення сторінки для фіктивної організації під назвою Український європейський фронт, який був підтриманий у Telegram. Сторінка була виявлена та вимкнена автоматизованими системами Meta.

Перша сторінка про Український європейський фронт з'явилася на польському вебсайті новин, а потім його підхопила невелика кількість преси переважно в Україні. У статті описано, що ця організація працює, щоб "запобігти тому, що Європа та увесь цивілізований світ забули про Україну". Операція включала новини щодо графіті біля Консульства України у Варшаві із закликами до виборів в Україні; трюк із виборчою урною у Варшаві, що пропонує вибори Президента України та заявляючи про проведення подібних заходів в інших містах Європи. У додатках Meta особи, які стояли за цією операцією, використовували підроблені облікові записи – деякі з них були виявлені та видалені нашими автоматизованими системами Meta.

Новини про Український європейський фронт без належної верифікації тиражували численні українські видання і інформагентства, наприклад вітчизняна агенція IPnews [20].

Інший приклад стосується компанії OpenAI і її ШІ чатбота ChatGPT, який використовується для створення і поширення фейкового контенту. OpenAI, яка розпочала кампанію нейтралізації спроб розповсюдження підробленого контенту вебсайтів і платформ соцмереж, у жовтні 2024 року оприлюднила звіт щодо операцій впливу і кібератак [21]. У звіті OpenAI зазначається, що деякі фейкові персони зосереджені на політиці в Європі чи США, зокрема поширюючи повідомлення про невтручання в "міжнародні проблеми, включно з вторгненням росії в Україну". Виявлені також операції, створені за допомогою ChatGPT, англійською, французькою та російською мовами з використанням сайту //euronewstop.co.uk із критикою України та її підтримки Великобританією. Французькою мовою багато коментарів і статей присвячені критиці України та Франції.

Висновки.

Підсумовуючи викладене, зазначимо, що ЄС посилює заходи правового реагування на протиправні дії щодо дезінформації, яка поширюється російськими державними органами. Аналіз Кодексу практики щодо дезінформації ЄС, Європейського акту про свободу медіа, Європейського акту про штучний інтелект дали підстави для висновку, що відповідні положення щодо дезінформації мають бути відображені в українському законодавстві.

Умови використання великих онлайн-платформ розроблені на основі ліберального підходу щодо свободи слова та боротьби з поширенням дезінформації визначаються переважно корпоративною політикою великих онлайн-платформ.

Приклади видалення дезінформаційного контенту приватними компаніями, розташованими в США, свідчать про необхідність посиленої уваги українського уряду до розробки правових механізмів безпосереднього зв'язку із великими онлайн-платформами, розташованими в США, задля протидії російській дезінформації.

Перспективами подальших наукових пошуків визначаємо питання наділення державних органів функціями співробітництва з великими онлайн-платформами задля протидії російській дезінформації.

Використана література

1. Skills shortage and unpatched systems soar to high-ranking 2030 cyber threats. March 27, 2024. URL: https://www.enisa.europa.eu/news/skills-shortage-and-unpatched-systems-soar-to-high-ranking-2030-cyber-threats

2. The Future of EU Digital Policy, May 21, 2024 URL: https://data.consilium.europa.eu/doc/document/ST-9957-2024-INIT/en/pdf

3. Council conclusions on democratic resilience: safeguarding electoral processes from foreign interference. May 21, 2024. URL: https://data.consilium.europa.eu/doc/document/ST-10119-2024-INIT/en/pdf

4. Horban Y, Oliinyk O. Media Literacy as a Factor in Protecting the Information Space from Enemy Disinformation in Time of War. Ukr Inf Space. 2024 Mar 29; (13):194-205.

5. Hajli N, Saeed U, Tajvidi M, Shirazi F. Social Bots and the Spread of Disinformation in Social Media: The Challenges of Artificial Intelligence. Br J Manag. 2022 Jul; 33(3):1238-53.

6. Cavaliere P. The truth in fake news: How disinformation laws are reframing the concepts of truth and accuracy on digital platforms. In: European Convention on Human Rights Law Review [Internet]. Brill Nijhoff; 2022 p. 481-523. URL: https://brill.com/view/journals/eclr/3/4/article-p481_005.xml

7. Spitale G, Biller-Andorno N, Germani F. AI model GPT-3 (dis)informs us better than humans. Sci Adv. 2023 Jun 30; 9(26).

8. Smith J. Designing Against Misinformation [Internet]. Design at Meta. 2017 [cited 2024 Nov 2]. URL: https://medium.com/designatmeta/designing-against-misinformation-e5846b3aa1e2

9. European Commission [Internet]. 2024. The Code of Practice on Disinformation. URL: https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation

10. European Media Freedom Act [Internet]. Apr 11, 2024. URL: http://data.europa.eu/eli/reg/2024/1083/oj/eng

11. Gamito M.C. The European Media Freedom Act (EMFA) as meta-regulation. Comput. Law Secur. Rev. 2023; 48:105799.

12. European Parliament. Artificial Intelligence Act [Internet]. Jun 13, 2024. URL: http://data.europa.eu/eli/reg/2024/1689/oj/eng

13. Hybrid threats/Russia: Statement by the High Representative on behalf of the EU on Russia's continued hybrid activity against the EU and its Member States. URL: https://www.consilium.europa.eu/en/press/press-releases/2024/10/08/hybrid-threatsrussia-statement-by-the-high-representative-on-behalf-of-the-eu-on-russia-s-continued-hybrid-activity-against-the-eu-and-its-member-states

14. Rep. Clarke YD [D N 9]. Deepfakes Accountability Act [Internet]. Jun 28, 2019. URL: https://www.congress.gov/bill/116th-congress/house-bill/3230/text

15. Sen. Wyden R [D O. Algorithmic Accountability Act of 2022 [Internet]. Feb 3, 2022. URL: https://www.congress.gov/bill/117th-congress/senate-bill/3572/text

16. Rep. Bice SI [R O 5]. Disinformation Governance Board Prohibition Act [Internet]. 2023. URL: https://www.congress.gov/bill/118th-congress/house-bill/4514/text

17. Sen. Paul R [R K. Free Speech Protection Act [Internet]. 2023. URL: https://www.congress.gov/bill/118th-congress/senate-bill/2425/text

18. House TW. The White House. 2023. Fact Sheet: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI. URL: https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai

19. Adversarial Threat Report. URL: https://files.elnashra.com/ elnashra/documents/2034261_1724410551.pdf

20. Український європейський фронт: українці за кордоном об'єдналися заради боротьби з Путіним. URL: https://ipne.ws/vojna-v-ukraine/politika/Ukrainskiy-ievropeyskiy-front-ukra

21. Influence and cyber operations: an update. October 2024. URL: https://cdn.openai.com/threat-intelligence-reports/influence-and-cyber-operations-an-update_October-2024.pdf?trk=public_post_comment-text

* * *