

Кирило ЧЕРЕВКО,

кандидат юридичних наук, доцент, професор кафедри кримінального права та кримінології Навчально-наукового інституту №1 Харківського національного університету внутрішніх справ, м. Вінниця, Україна
ORCID: <https://orcid.org/0000-0002-3384-8388>
e-mail: cherevko.kyrylo@univd.edu.ua

DOI: <https://doi.org/10.32620/pls.2025.8.41>

ПРОБЛЕМИ КВАЛІФІКАЦІЇ ДІЯНЬ НЕПОВНОЛІТНІХ ПРИ ВЗАЄМОДІЇ ЗІ ШТУЧНИМ ІНТЕЛЕКТОМ

Анотація: у тезах проаналізовано кейс арешту 13-річного учня у Флориді за запит до ChatGPT, що інтерпретували як загрозу. Цей випадок загострює проблему кваліфікації діянь неповнолітніх в епоху тотального цифрового нагляду. Досліджено конфлікт між політикою «нульової толерантності» та фундаментальним принципом кримінального права – *mens rea* (суб'єктивне ставлення до діяння), який автоматизовані системи не здатні оцінити. З позицій фрустраційно-агресивної теорії (Доллард та ін.), імпульсивний «жарт» підлітка помилково кваліфікується як кримінальне правопорушення. Обґрунтовано необхідність адаптації алгоритмів ШІ та правових процедур для уникнення надмірної криміналізації молоді.

Ключові слова: *mens rea*, штучний інтелект (ШІ), цифровий нагляд, кваліфікація злочинів, фрустрація-агресія, неповнолітні, нульова толерантність.

У середній школі Southwestern у окрузі Волусія, штат Флорида заарештували 13-річного учня. Під час перебування в школі учень ввів у діалогове вікно ChatGPT, що працював на шкільному комп'ютері, запит: «як убити мого друга посеред уроку» (how to kill my friend in the middle of class). [1] Зіткнувшись із наслідками, учень негайно надав пояснення своїм діям. Він стверджував, що запит був задуманий як «жарт» або «тролінг», спрямований на друга, який його дратував. Ця заява про відсутність злого умислу є ключовим елементом справи, що лежить на перетині кримінально-правових та кримінологічних аспектів. Питання, чи може такий «жарт» бути кримінальним правопорушенням, є ключовим для юридичного аналізу.

Реакція правоохоронних органів була негайною та рішучою. Після отримання сигналу тривоги на місце прибули помічники шерифа. Вони провели допит учня, і, незважаючи на його пояснення про жартівливий характер запиту, інцидент було кваліфіковано як серйозну надзвичайну ситуацію на території кампусу. Як наслідок, учня було заарештовано та доставлено до окружної в'язниці. Такий підхід демонструє застосування політики «нульової толерантності», за якої будь-яке висловлювання, що може бути інтерпретоване як загроза, розглядається з максимальною серйозністю, незалежно від

контексту чи заявлених намірів.[2]

Сповіднення про загрозу було згенеровано не безпосередньо компанією OpenAI, а стороннім програмним забезпеченням для спостереження під назвою Gaggle [3], яке шкільний округ встановив для моніторингу всієї цифрової активності. Система Gaggle використовує алгоритми фільтрації за ключовими словами (наприклад, «убити»), які миттєво активують протокол безпеки та надсилають сповіщення правоохоронним органам.

Система автоматичного аналізу не враховує людські наміри, іронію чи гумор. Вона реагує виключно на наявність певних слів. Це призводить до механізму, який позбавляє висловлювання учнем певного контексту, перетворюючи підліткову фрустрацію безпосередньо на проблему для поліції, що у свою чергу призводить до негайного арешту та застосування політики «нульової толерантності».

Повсюдне впровадження систем на основі ШІ (таких як Gaggle, Securly) у шкільні цифрові екосистеми створює середовище тотального контролю. Школа трансформується з освітнього закладу на інструмент попереднього кримінального нагляду, рутинізуючи присутність правоохоронних органів у житті учнів. Ефективність цих систем недоведена незалежними дослідженнями. Водночас, вони генерують надзвичайно високий рівень хибних спрацьовувань. У деяких випадках майже дві

третини сповіщень були визнані шкільними посадовцями такими, що не становлять проблеми, включаючи хибні тривоги, викликані змістом домашніх завдань. Такі системи мають непропорційно негативний вплив на вразливі групи, зокрема ЛГБТК+ молодь, для якої зростає ризик ненавмисного «аутингу».

З точки зору кримінології, цей процес найкраще характеризується як «предиктивне (або прогностичне) поліціювання» (Predictive Policing), що функціонує в межах «цифрового паноптицизму» (Digital Panopticism) та призводить до «соціального сортування» (Social Sorting) та «розширення мереж контролю» (Net-Widening). Те, що ми називаємо «інструментом попереднього кримінального нагляду», є прямою ілюстрацією концепції, яку часто називають «кримінологією до-злочину» (Pre-Crime Criminology). Цей термін, хоч і походить з наукової фантастики (Ф. К. Дік, «Особлива думка» [4]), став академічним для опису практик, що зміщують фокус кримінальної юстиції з реакції на злочин, що вже стався, на превенцію злочину, що може статися. Американський правознавець Бернард Харкорт (Bernard Harcourt) у своїх працях, зокрема «Against Prediction»[5], критикує цю «актуарну юстицію». Він стверджує, що такі системи не стільки передбачають майбутнє, скільки формують його, посилюючи нагляд за тими, хто вже маркований як «група ризику». Системи на кшталт Gaggles аналізують дані (пошукові запити, листування) для виявлення «індикаторів ризику». Школа, таким чином, перестає бути виключно освітнім простором і перебирає на себе функції прогностичного поліцейського контролю.

«Середовище тотального контролю» – це класична концепція паноптицизму, розроблена французьким філософом Мішелем Фуко (Michel Foucault)[6]. Паноптикон – це модель ідеальної в'язниці, де центральна вежа дозволяє наглядачу бачити всіх в'язнів, але вони не знають, чи дивляться на них у конкретний момент. Сама можливість постійного нагляду змушує їх до самодисципліни та самоконтролю. Сучасні кримінологи, як-от Девід Лайон (David Lyon), говорять про «суспільство нагляду» (Surveillance Society) або «цифровий паноптицизм»[7]. На відміну від в'язниці Фуко, сучасний нагляд є децентралізованим, автоматизованим (ШІ) і майже невидимим. Учень, знаючи, що кожне його слово в шкільному акаунті може бути (і буде) проаналізоване системою Gaggles, починає самоцензуруватись. Він не просто боїться покарання за вчинок, але й за думку або запит (навіть невинний), що може бути хибно інтерпретований ШІ. Це має глибокий «охолоджуючий ефект» на свободу вираження поглядів.

Хоча описаний досвід стосується переважно США (де і функціонує Gaggles), це є критично важливим і для України, яка також активно

цифровізує освітній простір. У США доктрина «in loco parentis» (школа діє замість батьків) та прецеденти щодо обшуків у школах надають адміністрації широкі повноваження, які компанії на кшталт Gaggles використовують для продажу своїх послуг. Ця доктрина загального права США, яка визнає, що освітні установи (школи, а історично й університети) та їхні представники (вчителі, адміністратори) приймають на себе частину батьківських прав та обов'язків щодо учнів, які перебувають під їхнім наглядом. Історично ця доктрина надавала школам надзвичайно широкі повноваження, аналогічні батьківським, у питаннях дисципліни, контролю та нагляду. По суті, вступаючи до школи, дитина тимчасово переходила з-під повної влади батьків під повну владу школи. Оскільки це доктрина загального права, вона написана не в законі, а в судових прецедентах. Одним із класичних і найбільш цитованих американських прецедентів, який чітко сформулював цю доктрину, є рішення Верховного суду штату Вермонт у справі *Lander v. Seaver* (1859)[8]. У цій справі суддя підтримав право вчителя застосувати тілесне покарання до учня за образливі висловлювання на адресу вчителя, сказані після шкільних занять і поза територією школи. Суд обґрунтував це тим, що така поведінка підриває авторитет вчителя і дисципліну в школі, а повноваження вчителя, що діє in loco parentis, є достатніми для такого втручання. У другій половині ХХ століття в США відбувся кардинальний зсув. Верховний Суд США почав визнавати, що учні є особами за Конституцією і не скидають свої конституційні права біля шкільних воріт (*Tinker v. Des Moines*, 1969)[9]. Це суттєво обмежило доктрину in loco parentis. Також, ключовою є справа *New Jersey v. T.L.O.* (1985)[10]. У цій справі школа провела обшук особистих речей (сумки) учениці за підозрою у курінні, виявивши докази вживання наркотиків.

Натомість в Європейському Союзі такий рівень моніторингу наражався б на жорсткі обмеження Загального регламенту про захист даних (GDPR)[11], який вимагає чіткої правової підстави, пропорційності та мінімізації даних.

Впровадження подібних систем в Україні зіткнулося б із прямими правовими бар'єрами, так ст. 31 КУ гарантує таємницю листування, телефонних розмов та іншої кореспонденції. Винятки можливі лише за рішенням суду. Ст. 32 КУ забороняє збирання, зберігання, використання та поширення конфіденційної інформації про особу без її згоди (або згоди батьків, але навіть тут є межа).[12] Закон України «Про захист персональних даних» вимагає чітко визначеної, законної мети обробки даних. "Попередження потенційних загроз" є надто розмитою метою, що не відповідає принципу пропорційності.[13]

Постійний моніторинг створює стримуючий ефект, змушуючи учнів уникати обговорення суперечливих тем або звернення по допомогу до

психологів через цифрові канали, побоюючись неправильного тлумачення алгоритмом. Це призвело до федеральних судових позовів, у яких стверджується, що «масштабна програма моніторингу без наявності підозр» порушує права учнів, гарантовані Першою (свобода слова) та Четвертою поправками (захист від необґрунтованих обшуків) Конституції США. [14]

Політики безпеки OpenAI виявляють глибоку суперечність. У випадках погроз іншим, компанія залишає за собою право повідомити правоохоронні органи про безпосередню загрозу. Цей реактивний підхід спрямований насамперед на мінімізацію юридичної відповідальності для самої компанії. Ця вразливість лежить в основі гучного судового позову, поданого родиною 16-річного Адама Рейна, який наклав на себе руки після, за словами адвоката родини, місяців заохочення з боку ChatGPT. [15] У позові стверджується, що модель не лише не зупинила підлітка, а й надала йому інструкції щодо способів самогубства та запропонувала допомогу в написанні передсмертної записки. Це свідчить про катастрофічний збій захисних бар'єрів ШІ під час тривалої взаємодії.

Законодавство Флориди, зокрема Статут 836.10 [2], забороняє електронні погрози, але правозастосовча практика «нульової толерантності» фактично ігнорує суб'єктивне ставлення особи до своїх дій. Факт передачі повідомлення з погрозою, навіть якщо це «жарт», може бути кваліфікований як тяжке кримінальне правопорушення.

Основний юридичний виклик полягає у розмежуванні у діянні учень ввів фразу з погрозою до ChatGPT, взаємодіючи з нелюдським агентом, а не комунікуючи з іншою особою. Цю дію можна інтерпретувати як **виявлення умислу** або запис у «цифровому щоденнику», на відміну від використання ШІ для створення та надсилання погрози конкретній особі (де ШІ є інструментом для вчинення – об'єктивної сторони кримінального правопорушення). Чинне законодавство не містить понятійного апарату для розрізнення між використанням ШІ як приватного простору для думок та його використанням як інструменту кримінального правопорушення. Відсутність цього розмежування, у поєднанні з тотальним спостереженням, створює ідеальні умови для криміналізації підліткових уявних експериментів.

Наукові дані свідчать, що підлітки все частіше звертаються до ШІ для розваги та емоційної підтримки, оскільки боти пропонують соціальну взаємодію без конфліктів та засудження. Це особливо стосується молоді, яка відчуває соціальну ізоляцію. Одне з найбільш цитованих досліджень на цю тему виявило, що значна частина підлітків використовує ШІ-компаньйонів саме для соціальної взаємодії та емоційної підтримки. Дослідження показало, що 72%

підлітків у США (віком 13-17 років) користувалися ШІ-компаньйонами. Серед причин використання 14% відзначили «відсутність засудження», а 12% цінують можливість «ділитися тим, чим не можуть поділитися з друзями чи родиною». [16] Схожі результати показало дослідження «Rating Group» (Україна). Це релевантне українське дослідження, яке вивчало ментальне здоров'я підлітків та їхню взаємодію з ШІ. Кожен 10-й український підліток (10-18 років) часто питає поради у ШІ про свій емоційний стан. Дослідження прямо пов'язує це з якістю емоційного зв'язку з батьками. Серед тих, хто часто радиться з ШІ, лише 23% мають міцний зв'язок з батьками, порівняно з 42% серед тих, хто рідко це робить. Це підтверджує тезу, що ШІ заповнює нішу «підтримки та підказки, якої... не отримує від близьких» [17]. Надмірна довіра до ШІ спричиняє «когнітивне розвантаження», знижуючи навички критичного мислення та соціальну адаптивність. Підлітки, будучи вразливою аудиторією, схильні використовувати ШІ як інструмент подолання емоційного стресу.

Дія учня, який був «роздратований» другом, може бути науково переформульована як використання потужного технологічного інструменту в якості **дезадаптивного механізму подолання повсякденної соціальної фрустрації**. Ця гіпотеза, спочатку сформульована Джоном Доллардом (John Dollard) та Нілом Міллером (Neal Miller) у 1939 році, стверджує, що агресія є прямим наслідком фрустрації. Фрустрація – це блокування або перешкоджання досягненню мети (в нашому випадку, мети соціальної взаємодії, поваги чи спокою). [18] Це не обов'язково є доказом справжнього злочинного наміру, а скоріше проявом емоцій, які були перехоплені системою, нездатною зрозуміти психологічний стан чи контекст розвитку.

Розробники систем спостереження повинні перепроєктувати алгоритми з урахуванням контексту, а не лише ключових слів, та вбудувати обов'язкові «запобіжники» для верифікації кваліфікованим шкільним персоналом (психологами). Розробники ШІ повинні інвестувати у створення безпечних систем, здатних деескалувати шкідливі наміри та впроваджувати вікові обмеження.

Висновок: Використання цифрових технологій у Флориді спричинило юридичні труднощі: автоматизований нагляд разом із застарілим кримінальним законодавством ведуть до того, що підлітки можуть бути притягнуті до відповідальності за імпульсивні вчинки. Юристам слід контролювати, щоб технічний прогрес не порушував правові процедури та людський підхід до судових рішень. Для сучасної України, особливо в умовах війни та посттравматичних розладів, ця проблема є надзвичайно актуальною. Масовий рівень фрустрації, спричиненої війною, втратою житла, економічною нестабільністю та горем, створює значні криміногенні ризики. Розуміння

механізму фрустрації допомагає прогнозувати сплески насильницької злочинності та розробляти програми психологічної підтримки для запобігання агресивним проявам.

Бібліографічні посилання

1. Teen arrested in Florida over Chat GPT query [Відео]. Brut. 2025. 09 Oct. URL: <https://www.brut.media/us/videos/us/society/teen-arrested-in-florida-over-chat-gpt-query>

2. The 2025 Florida Statutes. Title XLVI. Crime. Chapter 836. Defamation; Libel; Threatening Letters and Similar Offenses. Section 836.10. The Florida Legislature. URL: http://www.leg.state.fl.us/Statutes/index.cfm?App_mode=Display_Statute&URL=0800-0899/0836/Sections/0836.10.html

3. Gaggle | K-12 Online Safety Management Software. Gaggle. URL: <https://www.gaggle.net/>

4. Дік Ф. Особлива думка. Харків : КСД, 2016. 224 с.

5. Harcourt B. E. Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age. Chicago : University of Chicago Press, 2007. 336 p. ISBN 9780226316130.

6. Foucault M. Surveiller et punir : Naissance de la prison. Paris : Gallimard, 1975. 360 p.

7. Lyon D. The Electronic Eye: The Rise of Surveillance Society. Minneapolis : University of Minnesota Press, 1994. 270 p.

8. Lander v. Seaver, 32 Vt. 114 (1859) : Decision of the Supreme Court of Vermont.

9. Tinker v. Des Moines Independent Community School District, 393 U.S. 503 (1969) : Decision of the U.S. Supreme Court.

10. New Jersey v. T.L.O., 469 U.S. 325 (1985) : Decision of the U.S. Supreme Court.

11. Про захист фізичних осіб у зв'язку з опрацюванням персональних даних і про вільний рух таких даних, та про скасування Директиви 95/46/ЄС (Загальний регламент про захист даних) : Регламент (ЄС) 2016/679 Парламенту Європи та Ради ЄС від 27 квітня 2016 року. Офіційний вісник Європейського Союзу. 04.05.2016. L 119. С. 1–88.

12. Конституція України : Закон України від 28.06.1996 р. № 254к/96-ВР. Відомості Верховної Ради України. 1996. № 30. Ст. 141.

13. Про захист персональних даних : Закон України від 01.06.2010 р. № 2297-VI. Відомості Верховної Ради України. 2010. № 34. Ст. 481.

14. Конституція США // Конституції країн світу. Київ : Парламентське вид-во, 2005. С. 112–130.

15. Teen killed himself after 'months of encouragement from ChatGPT', lawsuit claims. The Guardian : вебсайт. 2025. 27 August. URL: <https://www.theguardian.com/technology/2025/aug/27/chatgpt-scrutiny-family-teen-killed-himself-sue-open-ai>

16. How Are Teens Using AI Companions? Benton Institute for Broadband & Society : вебсайт. 2025. 24 July. URL: <https://www.benton.org/blog/how-are-teens-using-ai-companions>

17. Прес-реліз (Rating Group): Як часто українські підлітки радяться з ШІ про свій психологічний стан? URL: <https://ratinggroup.ua/news/family360-mh-2>

18. Dollard J., Miller N. E., Doob L. W., Mowrer O. H., Sears R. R. Frustration and Aggression. New Haven : Yale University Press, 1939. 209 p.

K. Cherevko

Problems of qualifying juvenile actions in interactions with artificial intelligence.

Abstract: The abstract analyzes the case of a 13-year-old student arrested in Florida for a ChatGPT query interpreted as a threat. This case highlights the problem of qualifying juvenile acts in the era of total digital surveillance. It examines the conflict between "zero tolerance" policies and the fundamental criminal law principle of mens rea (criminal intent), which automated systems cannot evaluate. From the perspective of the frustration-aggression theory (Dollard et al.), the teenager's impulsive "joke" is mistakenly qualified as a crime. The necessity of adapting AI algorithms and legal procedures to avoid the over-criminalization of youth is substantiated.

Keywords: mens rea, artificial intelligence (AI), digital surveillance, qualification of crimes, frustration-aggression, juveniles, zero tolerance.

Зразок для цитування:

Черевко К. Проблеми кваліфікації діянь неповнолітніх при взаємодії зі штучним інтелектом. Пропілеї права та безпеки, 2025. №8 С. 173-176. DOI: <https://doi.org/10.32620/pls.2025.8.41>.