

УДК 343.346.8+001.103

**РАДУТНИЙ О.Е.**, доктор філософії (Ph.D.) в галузі права, доцент,  
доцент кафедри кримінального права Національного юридичного  
університету імені Ярослава Мудрого.  
ORCID: <http://orcid.org/0000-0002-6521-3977>.

## ВЕЛИКІ ДАНІ: КОРЕЛЯЦІЇ ТА ПРИЧИННІСТЬ (КРИМІНАЛЬНО-ПРАВОВИЙ АСПЕКТ)

DOI: ...

**Анотація:** На розвиток попередніх досліджень феномену Великих Даних та його впливу на сферу права в статті розглянуто переваги та недоліки використання кореляцій, які здатні конкурувати з встановленням причинності під час прийняття юридично значущих рішень. Звернуто увагу на той факт, що кореляції на підставі Великих Даних допомагають аналізувати об'єкти і явища не через з'ясування фундаментальних принципів їх внутрішньої побудови, але через корисні статистичні закономірності, які можуть бути не пов'язані з причинністю, але виявлення яких є цілком достатнім у значній кількості випадків. Доведено, що такі кореляції можуть вплинути на рішення, які приймає людина, доповнити аргументацію для обґрунтування рішень, які приймає людина, або суперечити рішенням, які приймає людина на підставі своїх знань та досвіду. Кореляції можуть бути корисними в тих галузях права, де статистичний аналіз є ефективним, їх розрахунок здійснюється на підставі математичних та статистичних методів з більшою швидкістю та ефективністю, а так само з меншими витратами, ніж встановлення причинно-наслідкових зв'язків. Автором доведено, що оскільки кореляції не є беззаперечним доказом причинно-наслідкового зв'язку, то вони повинні використовуватися в поєднанні з іншими типами доказів та відповідною юридичною аргументацією. Кореляції не повинні мати переважну силу над іншими доказами або аргументами. Звернуто увагу на небезпеку "диктатури" Великих Даних, коли даним або кореляціям може надаватися більше змісту та значення, ніж всьому іншому, коли захоплення кореляціями на підставі непрозорих алгоритмів здатне привести до критичних зловживань. Запропоновано передбачити запобіжники захисту від застосування висновків на підставі кореляцій для всіх осіб, які можуть через алгоритмічні та інші недоліки зазнати утисків, шкоди або іншого порушення прав та свобод, а так само встановити ієрархію між причинністю та кореляцією за критерієм точності у наступному вигляді: 1) кореляція (менша точність, одна з передумов встановлення причинності); 2) причинність (більша точність, яка може спиратися на кореляцію). Проголошено тезу про необхідність бути готовими до наступу епохи Великих Даних на фундаментальний принцип презумпції невинуватості у часи, коли світогляд, що базується на підґрунті з'ясування причини, може повільно поступатися світогляду, який спирається на Великі Дані та кореляції.

**Ключові слова:** Великі Дані, інформація, кореляція, причинність, алгоритм, штучний інтелект, персональні дані, інженерія оперативного управління, передбачене правосуддя, технології прогнозування, правосуддя на основі даних, предикативна аналітика, пеналізація, диктатура, чорна скринька, постійна пам'ять, кримінальне право, кримінальне правопорушення

**Summary:** Following up on previous studies of the Big Data phenomenon and its impact on the field of law, the article examines the advantages and disadvantages of using correlations which can compete with establishing causality in making legally significant decisions. The author draws attention to the fact that correlations based on Big Data help to analyse objects and phenomena not by clarifying the fundamental principles of their internal structure, but by revealing useful statistical patterns which may not be related to causality, but which are quite sufficient in a significant number of

*cases. It has been shown that such correlations can influence decisions made by a person, supplement arguments to justify decisions made by a person, or contradict decisions made by a person based on his or her knowledge and experience. Correlations can be useful in those areas of law where statistical analysis is effective, and their calculation is carried out on the basis of mathematical and statistical methods with greater speed and efficiency, as well as with lower costs than establishing causation. The author argues that since correlations are not conclusive evidence of causation, they should be used in conjunction with other types of evidence and appropriate legal argumentation. Correlations should not take precedence over other evidence or arguments. Attention is drawn to the danger of the dictatorship of Big Data, when data or correlations may be given more meaning and significance than anything else, when the fascination with correlations based on non-transparent algorithms can lead to critical abuses. The author proposes to provide safeguards to protect against the use of conclusions based on correlations for all persons who may, due to algorithmic and other shortcomings, be subjected to harassment, harm or other violation of their rights and freedoms, and to establish a hierarchy between causation and correlation based on the accuracy criterion as follows 1) correlation (lower accuracy, one of the prerequisites for establishing causation); 2) causation (higher accuracy, which may be based on correlation). The author argues that it is necessary to be prepared for the onset of the Big Data era on the fundamental principle of presumption of innocence at a time when the worldview based on the basis of finding out the cause may slowly give way to the worldview based on Big Data and correlations.*

**Keywords:** *Big Data, information, correlation, causality, algorithm, artificial intelligence, personal data, prompt engineering, predictive justice, predictive technologies, data-driven justice, predictive analytics, penalisation, dictatorship, black box, permanent memory, criminal law, criminal offence.*

**Постановка проблеми.** У попередньому дослідженні [37, с. 88-105] було розглянуто такі істотні властивості Великих Даних (Big Data), як безпрецедентно значний обсяг інформації в структурованому та(або) невпорядкованому (неструктурованому, швидкому, неточному) вигляді, який перевищує всі раніше доступні для одержання, зберігання та обробки і до якого не можуть бути застосовані попередні стандартні методи. Акцентовано увагу на ідеології зберігання абсолютно всіх даних без жодної втрати, незалежно від того, чи є в них потреба у певний проміжок часу, а так само вказано на прихований потенціал цінності будь-якої інформації, що може бути виявлений після її повторної обробки під новий запит. Оскільки поточний обсяг Великих Даних може бути доволі плінним і постійно змінюватися, то абсолютна точність може відходити на другий план. Але недоліки від такої втрати (на рівні одного споживача, найменшої одиниці вимірювання тощо) обертаються здобутками на макрорівні. Тому мистецтво роботи з Великими Данними включає у себе формулювання максимально ефективного запиту (який фактично дорівнює prompt engineering, тобто інженерії оперативного управління, використання навичок якої є витребуваним у взаємодії з алгоритмами штучного інтелекту на кшталт ChatGPT, Bing тощо), розробку унікальних методів його обробки, в тому числі, відповідне програмне забезпечення, яке здатне певним чином структурувати інформацію та відшукати в ній змістовні відповіді.

Не менш цікавий і важливий аспект феномену Великих Даних полягає у тому, що вони виступають підґрунтям для кореляцій, які здатні конкурувати з встановленням причинності під час прийняття юридично значущих рішень. Кореляція (correlation) є позначкою статистичного взаємозв'язку між двома випадковими змінними або двовимірними даними [30, с. 56]. Сильна кореляція вказує, що за зміну одних значень інші значення теж зміняться, завдяки чому вона допомагає аналізувати об'єкти не через з'ясування тих фундаментальних принципів, які утворюють їх підґрунтя, але через корисні статистичні закономірності, які можуть бути не пов'язані з причинністю [7].

Тож, у значній кількості випадків під час виявлення та використання кореляцій може бути відсутня відповідь на питання “чому”, але впевнене прогнозування ймовірності може виявитися корисним для з’ясування того, “що” відбудеться далі [14]. Великі Дані сприяють переходу від причинності до кореляцій, але в галузі права слід більш обережно підходити до використання останніх, ніж в інших сферах, зокрема, політики, економіки, екології тощо.

Поява феномену Великих Даних і значне поширення їх використання остаточно оформлює і підтверджує назву сучасного суспільства як інформаційного. Обсяг даних буде збільшуватися так само, як і нові способи їх обробки, тому необхідно приділяти увагу не тільки апаратному та програмному забезпеченню процесу обслуговування Великих Даних, але зосередитися на його наслідках, в тому числі тих, які мають правове значення.

**Результати аналізу наукових публікацій.** Дослідження феномену кореляцій на підставі Великих Даних та його впливу на сферу права неможливе без врахування внесків таких науковців і практиків, як Viktor Mayer-Schönberger, Kenneth Cukier, Katrina Yu, David Collins, Michelle Goodwin, Daniel Katz, Michael J. Bommarito, William H. Simon, David Restrepo Amariles, Catherine Locker, Daniel Cates, Jeffrey Hatherly, Arvind Narayanan, Cynthia Dwork, Paul Ohm, О.А. Баранов, Д. Білик, С.М. Брайчевський, В.М. Брижко, О. Винницький, М.В. Карчевський, О.В. Костенко, І. Марков, В. Песоцький, В.Г. Пилипчук, Н.А. Савінова та багатьох інших. Втім, ступінь розробки теми дозволяє проводити наступні повноцінні дослідження без втрати актуальності та новизни.

Так, на думку Річарда Берка, кримінолога з Університету Пенсільванії (Richard Berk, University of Pennsylvania) [4], засновані на кореляціях прогностичні моделі можуть бути корисними інструментами для виявлення осіб, які мають високий ризик вчинення кримінальних правопорушень. Його моделі використовують дані про попередню кримінальну історію, соціальні фактори та інші змінні для прогнозування ймовірності рецидиву і застосовуються у декількох юрисдикціях США для прийняття рішень про обрання запобіжного заходу, винесення вироку та умовно-дострокове звільнення. Між тим, Р. Берк разом з колегами визнає [5], що використання прогностичної аналітики (predictive analytics) має певні обмеження, які стосуються точності, справедливості та конфіденційності, виступає за продовження досліджень і оцінки таких інструментів з метою відповідального та етичного використання.

Бернард Гаркорт, професор права і політології Колумбійського університету (Bernard Harcourt, Columbia University), ставить під сумнів уявлення про те, що предиктивна поліція та інші форми прийняття рішень на основі Великих Даних можуть бути неупередженими та об’єктивними [12]. Його основний аргумент полягає в тому, що використання Великих Даних у правоохоронній діяльності увічнює та посилює існуючу соціальну нерівність, замість того, щоб вирішувати її, піддає маргіналізовані групи підвищеному нагляду та дискримінації [13]. Так само, журналістка і письменниця Джулія Енгвін (Julia Angwin) [1] застерігає проти широкого використання технологій нагляду та прийняття рішень на підґрунті Великих Даних як загрози для приватного життя та громадянських свобод, поглиблення расової та соціальної нерівності, посилення існуючих упереджень та дискримінації у зв’язку з їх недостатньою прозорістю та підзвітністю. Професор права Американського університету Ендрю Фергюсон (Andrew Ferguson, American University) [8] стверджує, що предиктивна поліція, яка спирається на кореляції та алгоритми для прогнозування ймовірних місць скоєння злочинів, може бути як ефективною, так і проблематичною: з одного боку, вона є здатною допомогти правоохоронним органам більш ефективно розподіляти ресурси і потенційно запобігати

правопорушенням, але з іншого боку, може підсилити упередження та дискримінацію, що має наслідком надмірний поліцейський нагляд в одних громадах і нестачу ресурсів в інших, вимагає підсилення прозорості, підзвітності та спирається на етичні міркування в кримінальному правосудді. Професор права Бостонського університету Даніель Сітрон (Danielle Citron, Boston University) [6] зосереджується на дослідженні алгоритмічної упередженості та дискримінаційних практик на підставі використання алгоритмів і Великих Даних, що призводить до посилення он-лайн переслідувань і злочинів на ґрунті ненависті.

Навіть за поверхового огляду є помітним той факт, що кореляції на підставі Великих Даних стають дедалі більше переконливими і привабливими для застосування, точність прогнозів постійно зростає, їх беруть до уваги і все більше покладаються під час прийняття юридично значущих рішень. Втім, у кожному випадку, коли це можливо, необхідно чітко розмежувати сфери, в яких доцільно задовольнятися кореляціями, від тих, де обов'язковим є встановлення причинності. Тож, проблематика використання в сфері права кореляцій на підставі Великих Даних зберігає свою актуальність і потребує подальших досліджень.

**Метою статті** є визначення переваг та недоліків використання в галузі права, в тому числі, кримінального, статистично обґрунтованих кореляцій на підставі Великих Даних поряд або замість встановлення причинності.

**Виклад основного матеріалу.** Завдяки переходу кількості у якість Великі Дані є викликом усталеному способу взаємодії та комунікації, а так само формування аргументації та побудови ланцюжків рішень, який традиційно спирався на категорію причинності. Між тим, підґрунтя для прийняття рішень не завжди базується на виявленні причини (наприклад, для вдалого маркетингу не є принциповим, чим саме обумовлюються уподобання користувачів Amazon, Netflix або Megogo), достатнім виявляється встановлення кореляції та побудови на її підставі рекомендацій, комерційних пропозицій та прогнозів. Але якщо виявити причину, то певні події стає не тільки можливим передбачити, але й керувати ними (наприклад, шляхом формування та закріплення окремих нейронних зв'язків надавати підказки користувачам, що саме їм подобається чи повинно подобатися в культурній, економічній або політичній сферах).

Кореляції на підставі Великих Даних дозволяють одержати відповідь на питання “що?” замість “чому?” і цього у переважній кількості випадків виявляється цілком достатньо. Такий підхід є повною мірою науковим та відповідає одному з фундаментальних принципів сучасної фізики [29, с. 65]: відомі нам “закони природи описують зміни”, що означає необхідність поділу картини опису світу на дві складові, а саме, стани і закони, де стан відповідає на питання “що це таке?”, закон – на питання “як воно змінюється?”. Зміст цього принципу полягає у тому, щоб замість відповіді на питання “чому речі або явища саме такі?” одержати пояснення з приводу того, “що відбудеться далі?”, спираючись на підґрунтя другого і третього принципів (“закони природи універсальні, тобто діють скрізь і завжди” та “закони природи локальні, тобто поведінка об'єкта в найближчому майбутньому залежить від умов, які існують безпосередньо поблизу нього” (принцип локальності), які уможливають проведення доказових експериментів. Так, в біології здобуто значний обсяг інформації про те, що клітини функціонують як потужні хімічні заводи, електростанції та сховища даних, що мають доволі складні інтерфейси для взаємодії з зовнішнім світом [16], але до цього часу не є відомим, чому вони саме такі і, тим більше, як побудувати аналогічні джерела інформації та енергії на рівні макросвіту людини. Так само в галузі фізики вдалося з'ясувати, як поведуться фундаментальні інгредієнти побудови всесвіту на мікрорівні (електрон, фотон, верхній та нижній кварк,

глюон, гравітон), але відсутнє розуміння та обґрунтування глибинної причинності, що не виключає можливості передбачити їх наступне ймовірне місце розташування, поведінку або прогнозовану взаємодію.

Кореляція не пояснює причини явища, але описує його характеристики і прогнозує розвиток подій, завдяки чому досягнення результату вбачається більш важливим, ніж з'ясування його глибинної причини (наприклад, використання випадково виявленого, незрозумілого, але статистично підтвердженого ефекту поєднання аспірину з апельсиновим соком на лікування певної хвороби).

Так, на підставі обробки 65 мільярдів операцій 1,5 мільярдів власників банківських карток в 210 країнах компанією MasterCard було встановлено певну кореляцію: якщо власник транспортного засобу заправляє його паливом приблизно о 16:00, то впродовж години він скоріш за все витратить \$35 – 50 США у продуктовому магазині або ресторані [19]. Алгоритм “оцінка відданості лікуванню” від компанії FICO на підґрунті складових, які не мають прямого відношення до медицини (зокрема, тривалість шлюбу, проживання в одному місці, трудові відносини з одним роботодавцем, наявність власного транспортного засобу тощо), дозволяє спрогнозувати, чи буде пацієнт акуратно приймати запропоновані йому ліки [22]. До менш переконливих прикладів кореляції можливо віднести одночасне співпадіння занепаду піратства в ХІХ столітті та глобальних кліматичних змін у бік потепління (насправді, на зменшення поширеності піратства вплинула не екологія, але зусилля європейських країн силами своїх військово-морських флотів). Іншими словами, Великі Дані дозволяють формулювати передбачення на підставі кореляцій. Практичне використання зазначених прогнозів з часом буде збільшуватися і конкурувати з іншими методами досліджень або побудови системи аргументації, в тому числі з встановленням причинності. Аналіз Великих Даних здатний запропонувати свіжий погляд та нові ідеї саме тому, що він не обтяжений межами традиційних підходів, усталеним способом мислення та притаманними йому упередженнями. Галузеві експерти можуть стати менш авторитетними, ніж фахівці з аналітики даних [20, с. 81]. Великі Дані краще, ніж передчуття фахівців та їх висновки та підставі досвіду, можуть передбачити той чи інший варіант розвитку подій та явищ.

Таким чином, кореляції на підставі Великих Даних можуть: 1) вплинути на рішення, які приймає людина; 2) доповнити аргументацію для обґрунтування рішень, які приймає людина; 3) суперечити рішенням, які приймає людина на підставі своїх знань та досвіду.

Феномен Великих Даних формує нові способи мислення та пізнання оточуючого світу, зокрема: 1) здатність аналізувати всі дані, замість того щоб вдовольнитися лише частиною або неповними вибірками; 2) готовність до того, щоб мати справу з неупорядкованими даними на шкоду індивідуальній точності задля формування загального висновку; 3) відмову в окремих випадках від встановлення причинності та точності на користь довіри математично обґрунтованим кореляціям. Завдяки Великим Данам виникає можливість зменшити ризик впливу особистих упереджень, крім того, значний обсяг даних та обчислювальні потужності самі по собі можуть виступати факторами виявлення закономірностей або кореляцій.

Використання кореляцій може бути корисним у сфері права. Вони можуть допомогти виявити закономірності та взаємозв'язки між різними змінними, які мають відношення до юридичних справ або правових ситуацій (зокрема, фактори, які сприяють злочинності, або ймовірність певних результатів у конкретному судовому процесі), тому актуальними тенденціями виступають передбачене правосуддя (Predict Courts' Decisions), технології прогнозування (Prediction Technology), правосуддя на основі даних (Data-driven justice) тощо. Втім, окремі країни намагаються їм протистояти. Так, у

2019 р. Франція зробила спірний крок і стала першою країною у світі, яка відповідним Законом Франції від 23 березня 2019 р. (LOIn 2019-222 du 23 mars 2019 deprogrammation 2018 – 2022 etderé for merourla justice) прямо заборонила суддівську аналітику та дослідження індивідуальної поведінки суддів [17].

Кореляції можуть бути корисними в тих галузях права, де статистичний аналіз є актуальним та ефективним. Вона знайшла своє використання у кримінальному праві та кримінології для аналізу закономірностей у рівні злочинності або вивчення ймовірності скоєння кримінального правопорушення конкретною особою на підставі її статистичних показників, у цивільному та трудовому праві – для вивчення моделей дискримінаційної поведінки, практик найму чи просування по службі, в екологічному праві – для розробки правових механізмів запобігання впливу забруднення або інших екологічних факторів на здоров'я та безпеку населення, у праві інтелектуальної власності – для визначення ймовірності порушення або виявлення потенційних порушень патентного або авторського права тощо.

Напрямами застосування кореляцій можуть бути: 1) одержання об'єктивних даних для підтримки юридичних аргументів або рішень, що може бути особливо корисно для нейтралізації впливу емоцій та суб'єктивних упереджень; 2) підвищення ефективності юридичних процесів шляхом швидкого виявлення ключової інформації та оптимізації процесу прийняття рішень; 3) розробка прогностичних моделей, прогнозування ймовірності результату в судовому процесі на підставі аналізу попередніх рішень певного судді або прецедентів зі схожими характеристиками; 4) виявлення тенденцій, зв'язків між певними змінними та результатом, закономірностей у факторах, які впливають на суддівські рішення, інформація відносно чого може бути використана для розробки більш ефективних юридичних стратегій та підбору переконливих аргументів, а так само для виявлення упереджень в судовій системі; 5) індивідуалізація кримінальних вироків, виявлення факторів, які значною мірою корелюють з рецидивом або ймовірністю вчинення нового кримінального правопорушення; 6) покращення правових досліджень на підставі аналізу обсягів правових Великих Даних; 7) забезпечення репрезентативності та підвищення ефективності відбору присяжних шляхом аналізу даних про їх характеристики, упередження та попередні рішення; 8) покращення доступу до правосуддя через системні поліпшення правової системи на підставі аналізу обсягів правових Великих Даних тощо.

Разом з цим, як було доведено у попередніх публікаціях [37, с. 88-105], з метою підвищення ефективності Великих Даних найбільш продуктивною виявилася стратегія збирання та зберігання максимально повного обсягу всієї інформації без жодних прогалин та втрат. Завдяки такому збільшенню обсягу даних певним чином послаблюються вимоги щодо їх упорядкованості та деталізованості, допустимими стають їх неструктурованість (невпорядкованість, неточність), але залишається нульова толерантність до відвертої помилковості або неправдивості. Новий масштаб дозволяє охопити обробкою різновиди інформації, які раніше не потрапляли до сфери уваги або їх було складно чи неможливо одержати.

Оскільки процес виявлення кореляцій вимагає менше витрат і часу, то у значній кількості випадків їх використовують поряд або замість встановлення причинності. У порівнянні з кореляціями, обґрунтування яких спирається на очевидні математичні докази, процес відшукування та доведення наявності причинно-наслідкового зв'язку між певними подіями та явищами є доволі складним і не завжди переконливим. Але оскільки мозок людини апріорі пристосований до відсутності або недостатності всієї необхідної інформації, ментальна пастка полягає у тому, що він самостійно вибудовує необхідні співвідношення, які ним сприймаються в якості необхідно-причинних. За

Даніелем Канеманом (Daniel Kahneman) [15, с. 74-75], людині притаманні дві основні форми мислення: швидке, яке не вимагає значних зусиль, та повільне, що характеризується значними витратами на обміркування. Швидке мислення схильне шукати причинно-наслідкові зв'язки навіть там, де їх немає, упереджено сприймає інформацію відносно наявних знань та переконань. Такий спосіб мислення корисний та допомагає вижити у небезпечному оточенні за необхідності миттєвого прийняття рішень в умовах обмеженої інформації, але йому бракує виявлення справжніх причин та їх наслідків. Тож, мозок людини у повсякденному житті лінується та не є схильним міркувати повільно та методично, внаслідок чого помилково спирається на удавані причинно-наслідкові зв'язки, що приводить до необґрунтованих або помилкових висновків та рішень. У зв'язку з цим В. Майер-Шенбергер (Viktor Mayer-Schönberger) та К. Кукьєр (Kenneth Cukier) вважають [20, с. 100-101], що всупереч загально прийнятому уявленню внутрішнє відчуття причинності жодним чином не поглиблює розуміння всесвіту і у більшості випадків воно є лише скороченим шляхом, який утворює ілюзію розуміння, але залишає у фактичному невіданні. У світі “малих даних” проходить багато часу, ніж вдається встановити помилковість певних причинно-наслідкових зв'язків. Але кореляції на підставі Великих Даних здатні спростувати їх удаваність і довести відсутність причинного та статистичного зв'язку між двома подіями або явищами, що має суттєво вплинути на повільне мислення.

До появи Великих Даних причинність відіграла у праві ключову роль. Зокрема, у кримінальному праві України побудована цілісна система встановлення необхідного причинного зв'язку між суспільно небезпечним діянням та суспільно небезпечним наслідком, який відрізняють від випадкового, що не відображає закономірностей розвитку подій і містить лише абстрактну можливість. Так, перш за все, два розглядуваних явища (можливі причина та наслідок) штучно (уявно) ізолюють (цей процес і його результат мають назву гіпотетичної елімінації) від всіх інших подій і явищ, що утворюють мережу причинно-наслідкових ланок і перебувають в універсальній взаємодії. Необхідний причинний зв'язок має місце лише тоді, коли діяння виступає необхідною умовою, без якої неможливе настання наслідку (*conditio sine qua non*): а) причина повинна передувати у часі наслідку (але “після” не завжди означає “внаслідок” (*post hoc non est propter hoc*); б) причина має викликати наслідок (зв'язок породження); в) розглядувана подія є головною, визначальною умовою (фактором) наступної; г) певна подія з неминучістю викликає наступну.

Між тим, слід визнати, що за допомогою цих правил вдається переконливо обґрунтовувати остаточне рішення лише у доволі простих ситуаціях. В інших випадках вдаються до переконання, умовляння або примушення учасників обговорення до колективної згоди імперативними засобами. Така складність ілюструється наступними прикладами.

*Ситуація 1.* Водій З. автомобілю Lexus виїхав на перехрестя на червоне світло світлофору, водій Д. автомобілю Volkswagen рушив на жовтий сигнал світлофора, між транспортними засобами відбулося зіткнення, автомобіль Lexus викинуло на тротуар, п'ятеро перехожих померли на місці, одна людина померла у лікарні, шестеро отримали травми. Водіїв З. та Д. було визнано винними і призначено покарання у вигляді 10 років позбавлення волі кожному.

*Ситуація 2.* Б. під час бійки наніс Т. ножове проникаюче поранення в передню стінку шлунку. У лікарні лікар Л. констатував стан сильного алкогольного сп'яніння пораненого, але був змушений дати йому наркоз, щоб провести необхідну операцію. Під час операції Т. помер через асфіксію від виверження зі шлунку.

В теорії кримінально права за межами України [11; 21; 28] розрізняють два види причинного зв'язку: фактичний та юридичний. Фактичний причинний зв'язок стосується тесту “якби не”, який запитує, чи настала би шкода, якби не поведінка певної особи, поведінка якої має бути необхідною умовою для настання шкоди. Якщо шкода виникла би незалежно від поведінки особи, то остання не є її фактичною причиною. В окремих юрисдикціях замість тесту “якби не” використовують тест “істотного (суттєвого) фактору”, який встановлює, чи була поведінка особи суттєвим фактором у заподіянні шкоди, навіть якщо були й інші фактори, що сприяли цьому. Юридичний причинний зв'язок, з іншого боку, стосується того, чи була поведінка особи безпосередньою причиною шкоди. Термін “безпосередня причина” означає, що шкода була передбачуваним наслідком поведінки особи. Питання полягає в тому, чи була поведінка особи прямою та істотною причиною шкоди, і чи перебувала шкода в межах відповідальності цієї особи. Для притягнення до кримінальної відповідальності сторона обвинувачення повинна довести як фактичний, так і юридичний причинний зв'язок поза розумним сумнівом. Але універсального правила для встановлення причинно-наслідкового зв'язку так само не існує, підхід до його встановлення залежить від конкретних фактів та обставин кожної справи.

Тож, у значній кількості випадків доволі складно довести як кореляції, так і причинність, натомість це можливо здійснити лише з певним ступенем вірогідності. Між тим, на боці кореляцій виступає той фактор, що на відміну від встановлення причинно-наслідкових зв'язків, існують переконливі математичні та статистичні методи для аналізу кореляцій, а так само належний інструментарій для впевненої демонстрації обґрунтованості висновків. Кореляції на основі Великих Даних можуть бути використані в кримінальному матеріальному та процесуальному праві, криміналістиці та криминології у різний спосіб, зокрема: 1) як докази в суді для підтримки або спростування аргументів сторони обвинувачення або захисту; 2) як підґрунтя для формулювання версій, слідчих зачіпок щодо зв'язків між людьми, місцями і подіями; 3) для оцінки ризику вчинення повторного кримінального правопорушення або потенційної майбутньої злочинної поведінки, що може вплинути на рішення про заставу, винесення вироку або обрання запобіжного заходу; 4) для створення мап злочинності та гарячих точок, що може допомогти правоохоронним органам ефективніше спрямовувати свої ресурси; 5) для виявлення закономірностей у злочинній діяльності з метою розкриття та запобігання вчиненню кримінальних правопорушень.

Використання кореляцій знайшло застосування у “прогностичному поліцейському контролі” (predictive policing) – діяльності правоохоронних органів, яка спирається на математичні засоби, прогностичну аналітику та інші аналітичні методи для виявлення потенційної злочинної діяльності [25]. Відповідно до звіту RAND найбільш ефективними виявилися наступні методи прогностичної роботи поліції: 1) прогнозування злочинів, 2) прогнозування правопорушників, 3) прогнозування особистості злочинців, 4) прогнозування жертв злочину [23]. За допомогою аналізу значних обсягів даних здійснюється планування патрулювання та контролю над окремими особами або групами для додаткової перевірки саме тому, що на них вказав алгоритм як на схильних до вчинення правопорушення. На території США поліцейські департаменти таких штатів, таких як Каліфорнія, Вашингтон, Південна Кароліна, Алабама, Арізона, Теннессі, Нью-Йорк та Іллінойс, використовують практику інтелектуальної поліцейської діяльності [9]. Поліцією Нью-Йорка впроваджено програму відстеження злочинів *Patternizr*, мета якої полягає у виявленні спільних рис правопорушень у виконанні тих самих осіб, для чого вона генерує можливу схему окремих різновидів суспільно небезпечної поведінки, але надалі



офіцер має вручну відшукати можливі шаблони для виявлення певних зв'язків [10]. У Мемфісі програма Blue CRUSH (Crime Reduction Utilizing Statistical History, “Зниження злочинності на основі статистичних даних”) надає поліцейським відносно точні дані про зони потенційної загрози з точки зору місця (у межах кількох кварталів) та часу (у межах кількох годин конкретного дня тижня) [27]. ГУ НПУ в Харківській області використовує інтелектуальну систему кримінального аналізу даних RICAS (Realtime intelligence crime analytics system), яка здійснює аналіз кримінальної обстановки (crime pattern analysis), аналіз загального профілю (general profile analysis), аналіз конкретного розслідування (case analysis), порівняльний аналіз (comparative analysis), аналіз групової злочинності (offender group analysis), аналіз особливостей профілю (specific profile analysis), аналіз розслідувань (investigation analysis) тощо [24]. Інтелектуальний поліцейський контроль використовується в Нідерландах. Започаткований у 2017 р. датський проект POL-INTEL базується на системі Gotham від Palantir Technologies, яка також використовується державною поліцією Німеччини та Європоллом [22].

Зазначені системи та подібні до них у повній мірі відповідають охоронній та превентивній функціям кримінального права, які передбачають, у тому числі, пріоритетність реагування на упередження та запобігання доведенню кримінального правопорушення до завершення. Такі заходи розглядаються як більш ефективні у порівнянні з ретроспективною (негативною) юридичною відповідальністю. Але у випадку виявлення особи потенційного правопорушника на переконливому підґрунті всіх кореляційних ознак ще до моменту вчинення суспільно небезпечного діяння, слід згадати про повну відсутність єдиної підстави кримінальної відповідальності, яка передбачена ч. 1 ст. 2 КК України. А так само про кару, як мету покарання (ч. 2 ст. 50 КК України) у той час, як карати поки немає за що, а покарання на майбутнє переходить всі межі. Тоді наступною привабливою формою реагування з боку держави в якості спеціальної превенції може вбачатися усне або письмове попередження, профілактична бесіда тощо. Але і такий підхід має суттєві вади, адже у зв'язку з фіксацією факту превентивних заходів щодо певної особи правові та інші соціальні наслідки для неї можуть бути вкрай негативними (зокрема, ними можуть бути “постійна пам'ять” [37, с. 88-105], відмова у прийнятті на роботу, звільнення з посади, пониження рейтингу), що обмежуватиме її права і свободи ще до і, головне, замість вчинення певного акту поведінки. Утримання балансу між інтересами суспільства та правами і свободами окремої особи в розглядуваному випадку утворює складний правовий виклик.

В галузі пеналізації сьогодні відсутній будь-який протокол (правила, формула, алгоритм, метод) щодо визначення на законодавчому рівні виду і розміру покарання за конкретне кримінальне правопорушення. Зазвичай за основу приймається попередня традиція (зокрема, відповідна норма в КК України 1960 р. або у попередніх редакціях чинного КК України) або суб'єктивно обраний результат зіставлення з суміжними або умовно схожими іншими кримінальними правопорушеннями. Такий підхід є притаманним не Великим Даним, але “малим”. У свою чергу практика спирається на діапазон, який визначений у санкції відповідної статті КК України. Кореляції на підставі аналізу Великих Даних реально здатні допомогти визначити співвідношення між будь-якими деліктами за ступенем їх суспільної небезпечності, що суттєво збагатить законодавчий інструментарій юридичної техніки. Але використання кореляцій під час індивідуального призначення покарання конкретній особі вбачається не повною мірою виправданим.

У зв'язку з можливістю здійснення вибору і зміни людиною своєї поведінки ідеальне прогнозування її майбутніх вчинків є вкрай складним. На підставі кореляцій не

може бути притягнута до відповідальності будь-яка особа до того, як вона вчинила конкретний вчинок, інакше це спростує ідею презумпції невинуватості, що покладена у підґрунтя правової системи. Необхідно передбачити правові запобіжники та інші механізми протидії подоланню певних розумних меж у тому випадку, коли кореляції на підставі індивідуальної інформації про певну особу спрогнозують високу ймовірність вчинення нею кримінального або іншого правопорушення. Тож, існує небезпека “диктатури” Великих Даних, коли даним надається більше змісту та значення, ніж всьому іншому, коли захоплення кореляціями на підставі непрозорих алгоритмів здатне привести до критичних зловживань, приклади чого вражають навіть в епоху малих даних (так, п’ятизначні номери татуювань в’язнів нацистських концтаборів відповідали номерам комплексної системи обліку на перфокартах IBM Hollerith, яка була використана для масових вбивств та катувань у промислових масштабах [2; 3]).

У випадку непрозорості певного алгоритму, а також відсутності доступу до вхідних даних та(або) обраних критеріїв, ускладнюється або унеможлиблюється вірогідність спростування або оскарження певного висновку з юридично значущими наслідками. Тому у зв’язку з поширенням використання Великих Даних та кореляцій на їх підставі до відомих викликів порушення приватності (конфіденційності) та непрозорості алгоритму додається також загроза некоректного обчислення юридично значущої вірогідності.

Поряд з перевагами та об’єктивним підвищенням ефективності від використання кореляцій у галузі права слід обов’язково звернути увагу на можливі недоліки, упередження або помилки. Завдяки Великим Даним алгоритми дійсно одержують корисну можливість прогнозувати будь-які події, наприклад, вірогідність серцевого нападу, погіршення стану здоров’я, невиконання покладених обов’язків або вчинення правопорушення. На цьому статистично доведеному підґрунті можуть бути прийняті юридично, політично та економічно значущі рішення, наприклад, щодо збільшення вартості страхування, дострокового звільнення або відмови у посаді, залишення заяви про кредитування без задоволення, застосуванні більш суворого покарання поміж декількох альтернативних, відмові у достроковому звільненні від відбування покарання тощо.

Превентивна функція окремих галузей права, зокрема, кримінального, спирається на пріоритетність запобігання настанню шкідливих наслідків у порівнянні з їх наступним відшкодуванням та прагненням повернення до попереднього стану: загальна превенція має спрацьовувати на осіб, які не вчинили правопорушення і тим самим утримувати їх від нього, спеціальна – на осіб, які його вчинили та на своєму негативному досвіді відчули суворість та невідворотність реакції з боку держави. Оскільки превенція завжди спрямована у майбутнє, виникає дилема щодо обрання критерію для прийняття рішень, якими можуть бути як висновок алгоритмів на підставі Великих Даних, так і конкретні вчинки, які відбулися і можуть вказувати на певний напрямок розвитку розглядуваної ситуації.

З появою Великих Даних людство дійсно вступає у новий світ, в якому складне дослідження непереконливої причинності може поступатися спіранню на непрозорі кореляції. Внаслідок цього непрозорість рішення уповноваженої особи або органу (лікар, керівник, орган досудового розслідування, суд певної ланки тощо) може базуватися на висновку “чорної скриньки” непрозорого алгоритму, у зв’язку з чим втрачається можливість аналізу, оскарження та перегляду такого рішення, яка в окремих випадках хоча і є примарною, але залишає надію. “Чорні скриньки” алгоритмів не повинні за своїми характеристиками тяжіти до “чорних дір” космосу, але мають відповідати ознакам прозорості, підзвітності, цілісності, спостережуваної доказовості тощо.

Позитивний ефект від використання кореляцій не повинен слугувати приводом для

заміни в системі кримінального права, а так само інших галузей, традиційного поняття “ймовірна причина” на “ймовірнісна причина”. Основна різниця між ними полягає у тому, що перша передбачає наявність безпосереднього зв'язку між однією подією та іншою, тоді як ймовірнісна причина виражає ступінь вірогідності того, що певна подія може бути причиною іншої (наприклад, ймовірною причиною аварії може бути допуск до керування особи, яка перебуває в стані алкогольного сп'яніння, ймовірнісною – складні погодні умови, які збільшують ризик керування джерелом підвищеної небезпеки).

Між кореляцією та причинністю необхідно закріпити ієрархію співвідношення, за яку кореляції не є беззаперечним доказом причинно-наслідкового зв'язку і повинні використовуватися в поєднанні з іншими типами доказів та відповідною юридичною аргументацією. Кореляції не мають бути використані в якості аргументу, який має переважну силу над іншими. Предиктивна аналітика вже продемонструвала певні недоліки і здатна викликати більш суттєві негативні наслідки. Кореляції на основі Великих Даних слід ретельно аналізувати, щоб переконатися, що вони не є упередженими або не ґрунтуються на помилках або хибних припущеннях, для чого вони мають бути прозорими та придатними для перевірки. Етичні проблеми використання кореляцій можуть бути пов'язані питаннями справедливості. Тому до використання кореляцій у сфері права слід підходити з обережністю, підкріплюючи рішення ретельним аналізом всіх обставин, глибоким розумінням не тільки правових принципів, але й основ роботи алгоритмів. Задля досягнення цієї мети правники мають оволодіти базовими знаннями з програмування, тестування, інженерії оперативного управління (prompt engineering) тощо.

Тому вбачається доречним тримати під постійною увагою наступні основні недоліки використання кореляцій в сфері права:

1. *Кореляція не вказує на причинно-наслідковий зв'язок і не доводить його існування.* Кореляція між двома змінними не обов'язково означає, що одна з них спричиняє (викликає, породжує) іншу. Зв'язок може бути хибним, якщо існує третій фактор, який відповідає за кореляцію, що об'єктивно спостерігається.

2. *Ризик упередженості.* На кореляції можуть впливати упередженість даних або вибір змінних. Дослідники можуть вирішити виключити або включити певні змінні, щоб досягти бажаного результату або приховати перешкоджаючі фактори, які могли б пояснити кореляцію.

3. *Відсутність контексту.* Кореляції часто використовуються для того, щоб робити прогнози або висновки про групу людей чи події. Однак вони можуть не враховувати конкретні обставини або індивідуальні відмінності, які можуть вплинути на результати.

4. *Обмежена сфера застосування.* Кореляції описують лише взаємозв'язок між двома змінними і не дають повного розуміння проблеми, що розглядається, вони можуть не враховувати важливі фактори, які мають відношення до ситуації.

5. *Неправильна інтерпретація.* Збільшення обсягу даних сприяє відшукуванню статистичних зв'язків між різними факторами, але ці зв'язки можуть виявитися помилковими, або недостатньо обґрунтованими (наприклад, кореляція між збільшенням обсягу продажів морозива у літній період та збільшенням кількості смертей від утоплень у водоймах не вказує на те, що морозиво є причиною смерті).

6. *Недостатня якість даних.* Кореляції є ефективними настільки, наскільки є якісними ті дані, які використані для їх створення. Неточні, неповні, застарілі або не релевантні дані можуть призвести до оманливих кореляцій. Наприклад, кореляції можуть свідчити про те, що люди з певною сукупністю характеристик є більш схильними до протиправної діяльності, однак не підтверджуватися щодо конкретної особи, але рішення,

що ґрунтуються виключно на кореляціях, здатні формально обґрунтувати несправедливе ставлення до неї.

7. *Неточність шаблонів.* Кореляції можуть ґрунтуватися на шаблонах даних, які не завжди є надійними індикаторами причинно-наслідкового зв'язку або навіть самої кореляції.

8. *Порушення конфіденційності.* Кореляції на підставі Великих Даних можуть розкрити конфіденційну інформацію, навіть якщо ця інформація не була безпосередньо надана або не призначалася для розкриття (так, співставлення купівельних звичок з даними про місцезнаходження торговельних закладів може ідентифікувати домашню адресу, аналіз історії пошуку з медичними показниками здатний виявити інформацію про стан здоров'я тощо). Як наслідок, кореляція може потенційно порушити право особи на приватність, розкриваючи інформацію, якою вона не мала наміру ділитися або навіть не знала, що її збирають.

9. *Втручання в сферу захисту інтелектуальної власності та охоронюваних таємниць.* Кореляції на основі Великих Даних потенційно можуть виявити закономірності та зв'язки, які є захищеними законами про інтелектуальну власність, зокрема, торговими марками, патентами або комерційною таємницею. Якщо аналіз Великих Даних дозволяє виявити кореляцію між двома змінними і вона є цінною для певної сфери діяльності, то існує можливість запатентувати відповідний алгоритм або метод і використовувати його для виявлення інших подібних кореляцій, які у незалежний спосіб розкривають певну охоронювану таємницю.

10. *Неналежний рівень прозорості і підзвітності.* (1). Відшукування та використання кореляцій на основі Великих Даних може спиратися на складні алгоритми та методи обробки, які важко зрозуміти та пояснити. Такий підхід може бути пояснений через охорону державної, службової, комерційної, банківської таємниці або іншої конфіденційної інформації, але поруч з цим він є здатним привести до прийняття рішень за принципом "чорної скриньки", коли вхідні дані та внутрішні процеси, які використані для прийняття певного рішення, не є прозорими, в тому числі для осіб, на яких таке рішення здійснює вплив і має правові наслідки. Зазначена непрозорість може ускладнити розуміння підстав для прийняття рішення, його перевірку та оскарження. (2). Використання кореляцій може призвести до створення "таємних законів" або прийняття рішень, які не підлягають перегляду або нагляду з боку громадськості, судів чи інших публічних органів. (3). Використання кореляцій може створити проблеми з підзвітністю, оскільки ускладнює розподіл відповідальності за конкретні рішення. Якщо рішення приймається на основі кореляцій, які згенеровані алгоритмом штучного інтелекту або системою машинного навчання, то може бути відсутня фактична підстава юридичної відповідальності у поведінці конкретної уповноваженої особи (можливо, до того часу, поки штучний інтелект не одержав відповідний статус суб'єкта правових відносин [33, с. 123-136; 35, с. 19-41; 34, с. 78-96]), що дозволить останній уникати відповідальності за неправові рішення та вчинки.

11. *Дискримінаційні практики.* Використання кореляцій на основі Великих Даних може призвести до дискримінаційних результатів за ознакою раси, статі, ставлення до релігії чи інших захищених характеристик, якщо дані, які задіяні в процесі аналізу, є упередженими або якщо алгоритми не мають запобіжників для пом'якшення упередженості.

12. *Етичні міркування.* Використання кореляцій у сфері права може викликати етичні занепокоєння щодо значної кількості правових практик за ознакою справедливості (яка є оціночним поняттям, але згадується у рішеннях Верховного Суду, Конституційного Суду України та ЄСПЛ).

Найбільш недосконале використання кореляцій у праві може мати місце в індивідуальних ситуаціях, коли ставки високі, а можливість помилки або недостатня

коректність можуть призвести до істотних хибних наслідків. Зокрема, у кримінальних справах, коли йдеться про свободу та добре ім'я, в тому числі впродовж всього досудового розслідування та судового розгляду, покладатися виключно лише на кореляції може бути ризиковано і не виправдано. Важливо мати більш конкретні та переконливі докази, які можуть підтвердити або спростувати кореляцію. Покладання виключно на кореляції в таких справах може призвести до неправомірного засудження або виправдання. Аналогічно, у цивільних та господарських справах про значні збитки, великі суми або суттєві майнові права, так само слід використовувати кореляції лише як один із багатьох доказів, жодний з яких не має задалегідь переважної сили.

Доцільно встановити запобіжники захисту від застосування висновків на підставі кореляцій та кількісних показників для тих осіб, які можуть через них зазнати утисків або шкоди (індекси впливовості або цитування, рейтинги навчальних закладів між собою, оцінювання працівника в організації роботодавця, відмова у працевлаштуванні, обрання певного протоколу лікування серед інших, репутаційні та правові ризики для фізичної особи, відносно якої відбувся прогноз вчинення нею кримінального правопорушення, але об'єктивно воно не відбулося тощо). Але у більшості випадків потрібно не стільки адаптувати старі правила та підходи, скільки розробляти та встановлювати нові.

Кореляції на підставі Великих Даних здатні вказати перспективні напрямки пошуку причинності, в окремих випадках перевищити за своєю ефективністю більшість інтуїтивно встановлених причинно-наслідкових зв'язків. Вони можуть допомогти сформуванню нове бачення певної проблеми. Але бажано уникати потрапляння у пастку пояснення причинності через кореляцію. Зокрема, якщо кореляція доводить, що люди з малим розміром взуття почуваються більш щасливими, а з великим – більш впевненими, то було би помилково носити менший розмір для відчуття щастя, так само як більший – для відчуття впевненості.

Призначення Великих Даних та кореляцій на їх підставі можливо вбачати здебільше в інформуванні, ніж у поясненні. Для справжнього захоплення кореляціями є всі підстави, але воно не повинне підмінювати відповіді на питання “чому?” поясненнями “що?”.

В одних випадках кореляції можуть повністю задовольнити запит, в інших – слугувати менш витратним та більш ефективним механізмом обробки даних для встановлення причинно-наслідкових зв'язків та взаємозалежних стосунків. Завдяки кореляціям утворюється підґрунтя для встановлення важливих лінійних та нелінійних відносин, а так само для заглиблення у пошуках причинності.

Оскільки дослідження кореляцій не повинне підмінювати собою встановлення причинності, то ієрархія між причинністю та кореляцією за критерієм точності повинна мати наступний вигляд: 1) кореляція (менша точність, одна з передумов встановлення причинності); 2) причинність (більша точність, яка може спиратися на кореляцію).

Кореляція повною мірою є ефективною тоді, коли є необхідність дослідити саме явище безвідносно до його причини. Але за необхідності пошуку причини можливо і доцільно почати з кореляційного аналізу Великих Даних. Найбільш вдало кореляції на підґрунті Великих Даних можуть бути використані у тих випадках, коли не є принциповою індивідуалізація (наприклад, смаки великої кількості споживачів). Натомість, кореляції слід вкрай обережно використовувати там, де індивідуалізація є дуже важливою (наприклад, всі персональні рішення, зокрема, щодо притягнення до кримінальної відповідальності, звільнення від неї, призначення покарання або звільнення від нього тощо).

До появи Великих Даних та кореляцій на їх підставі одна з найбільших небезпек в інформаційній сфері полягала у порушенні режиму приватності або конфіденційності. Сьогодні до неї додалися загроза деанонізації [37, с. 88-105] та неналежне використання

обчисленої вірогідності з високим ступенем ймовірності під час прийняття юридично значущих рішень.

Кореляції здійснюються алгоритмами, але їх споживачами виступає людина, яка, як правило, потребує пояснення причин. Але з часом звичка до використання кореляцій здатна підмінити собою звичний фундамент шляхом зменшення значення причинності, що слід вважати доволі загрозливим. Якщо переконливі кореляції на підставі аналізу Великих Даних сформулюють висновок про те, що поведінка певної особи є повністю передбачуваною та зумовленою, то наступним кроком може стати досягнення одного з негативних наслідків: 1) утворення підстави для притягнення до відповідальності або застосування превентивних заходів до особи, яка ще не вчинила кримінальне правопорушення або не має реального наміру продовжувати свою протиправну діяльність; 2) усунення (невизнання, заперечення, ігнорування) ознаки волі та свободи вибору.

Переконливе передбачення за допомогою кореляцій на підставі Великих Даних може мати доволі високий ступінь вірогідності та утворювати спокусу надавати оцінку не фактичним обставинам, але прогнозам щодо майбутньої поведінки. Але заперечення свободи вибору та свободи волі неминуче приводить до іншої пастки і означатиме відсутність підстав притягнення до юридичної відповідальності. Разом з тим, спирання на звичай в якості джерела права [31, с. 130-135] дозволяє відшукати опору в культурній традиції: будь яке ворожіння (таро, руни, Книга змін тощо), в тому числі наукове ворожіння [32, с. 58-67; 36, с. 99-16], не пропонує остаточного вирішення проблеми наперед, але вказує на певний інерційний сценарій, тобто, найбільш ймовірний варіант розвитку подій, якщо все залишиться у тій конфігурації, як це склалося на момент запиту. Оскільки за людиною визнається, хоча і не абсолютна, але реальна можливість змінювати своє майбутнє власними вчинками, тому відповідальність за майбутній вчинок буде протирічити свободі вибору. Кореляції на підставі Великих Даних, як й інший інструментарій народних та наукових традицій, дозволяють краще зрозуміти поточні та майбутні ризики, скорегувати дії в той чи інший спосіб. Але вони не замінюють собою причинності і не утворюють причинність на майбутнє.

Застосування кореляцій не повинне перешкоджати наданню зацікавленій особі можливості спростувати їх прогноз, який об'єктивно може бути обґрунтованим та переконливим, але не остаточним. За використання кореляцій ми маємо розширити заходи і обсяг забезпечення прав кожної особи, відносно якої застосовується аналіз і приймається рішення. Вдосконалення такого захисту має ґрунтуватися на наступних принципах: 1) відкритість, яка має стосуватися всього набору даних, критеріїв та алгоритму їх обробки; 2) розкриття всіх факторів, які можуть спростувати результати кореляції; 3) можливість одержання пояснень, в тому числі щодо логіки процесу виконання; 4) можливість повторного відтворення та спростування.

До того моменту, коли і якщо буде обґрунтований інший підхід, слід уникати застосування індивідуальної відповідальності на підставі колективних даних. Під час використання кореляцій на підставі Великих Даних повинна бути забезпечена гарантія базування індивідуальної відповідальності на об'єктивних проявах поведінки на відміну від покладання на прогнози. Право має бути готовим до наступу епохи Великих Даних на презумпцію невинуватості у часи, коли світогляд, що базується на підґрунті з'ясування причини, може повільно поступатися світогляду, який спирається на Великі Дані та кореляції.

### **Висновки та пропозиції.**

Завдяки переходу кількості у якість Великі Дані є викликом усталеному способу взаємодії, комунікації, аргументації, побудови ланцюжків доказів та прийняття рішень,

який традиційно спирався на категорію причинності. Мистецтво роботи з Великими Даними включає у себе формулювання максимально ефективного запиту до алгоритму обробки на підставі інженерії оперативного управління (prompt engineering). Феномен Великих Даних остаточно підтверджує назву сучасного суспільства як інформаційного. Необхідно приділяти увагу не тільки апаратному та програмному забезпеченню процесу обслуговування Великих Даних, але зосередитися на його наслідках, в тому числі тих, які мають правове значення.

Великі Дані виступають підґрунтям для кореляцій, які здатні конкурувати з встановленням причинності під час прийняття юридично значущих рішень. Кореляція є позначкою статистичного взаємозв'язку між двома випадковими змінними або значеннями. Великі Дані сприяють переходу від причинності до кореляції, але в галузі права слід більш обережно підходити до використання останніх.

Кореляції на підставі Великих Даних дозволяють одержати відповідь на питання “що?” замість “чому?” і цього у значній кількості випадків виявляється цілком достатньо. Кореляції допомагають аналізувати об'єкти і явища не через з'ясування фундаментальних принципів їх внутрішньої побудови, але через корисні статистичні закономірності, які можуть бути не пов'язані з причинністю. Сильні кореляції жодним чином не пояснюють певне явище, оскільки пропонують вірогідність замість визначеності. У певних випадках кореляції можуть допомогти з'ясувати причинний зв'язок.

Кореляції на підставі Великих Даних можуть: 1) вплинути на рішення, які приймає людина; 2) доповнити аргументацію для обґрунтування рішень, які приймає людина; 3) суперечити рішенням, які приймає людина на підставі своїх знань та досвіду.

Феномен Великих Даних формує нові способи мислення та пізнання оточуючого світу, зокрема: 1) здатність аналізувати всі дані, замість вдоволення лише частиною або неповними вибірками; 2) готовність мати справу з неупорядкованими даними на шкоду індивідуальній точності задля формування загального висновку; 3) відмова в окремих випадках від встановлення причинності та певного рівня точності на користь довіри математично обґрунтованим кореляціям.

Кореляції можуть бути корисними в тих галузях права, де статистичний аналіз є ефективним, для виявлення закономірностей та взаємозв'язків між різними змінними (фактори, що сприяють злочинності, ймовірність певного результату у судовому процесі тощо). Актуальними тенденціями виступають передбачене правосуддя (Predict Courts' Decisions), технології прогнозування (Prediction Technology), правосуддя на основі даних (Data-driven justice) тощо. Напрямами застосування кореляцій в галузі права можуть бути: 1) одержання об'єктивних даних, які можна використовувати для підтримки юридичних аргументів або рішень, для нейтралізації впливу емоцій та суб'єктивних упереджень; 2) підвищення ефективності юридичних процесів шляхом своєчасного виявлення ключової інформації та оптимізації процесу прийняття рішень; 3) розробка прогностичних моделей, прогнозування ймовірності результату в судовому процесі на підставі аналізу прецедентів зі схожими характеристиками; 4) виявлення тенденцій, зв'язків між певними змінними та результатом, закономірностей у факторах, які впливають на суддівські рішення, для виявлення упереджень в судовій системі; 5) індивідуалізація кримінальних вироків, виявлення факторів, які найсильніше корелюють з рецидивом або ймовірністю вчинення нового кримінального правопорушення; 6) покращення правових досліджень; 7) забезпечення репрезентативності та підвищення ефективності відбору присяжних шляхом аналізу даних про їх характеристики та упередження; 8) покращення доступу до правосуддя через системні поліпшення правової системи тощо.

До появи Великих Даних причинність відіграла у праві ключову роль. У кримінальному праві України побудована цілісна система встановлення необхідного причинного зв'язку між суспільно небезпечним діянням та суспільно небезпечним наслідком. Але за її допомогою вдається переконливо обґрунтувати остаточне рішення лише у доволі простих ситуаціях. В інших випадках фактично вдаються до переконання, умовляння або примушення учасників дискусії до колективної згоди імперативними засобами (зокрема, рішенням суду, яке набрало законної сили).

У значній кількості випадків складно довести як кореляції, так і причинність. Але розрахунок кореляцій здійснюється з більшою швидкістю та з меншими витратами, крім того, на відміну від процесу встановлення причинно-наслідкових зв'язків, існують математичні та статистичні методи для аналізу кореляцій, переконливий інструментарій для впевненої демонстрації обґрунтованості висновків.

Кореляції на основі Великих Даних можуть бути використані в кримінальному матеріальному та процесуальному праві, криміналістиці та кримінології у різний спосіб, зокрема: 1) як докази в суді для підтримки або спростування аргументів сторони обвинувачення або захисту; 2) як підґрунтя для формулювання версій, слідчих зачіпок щодо зв'язків між людьми, місцями і подіями; 3) для оцінки ризику вчинення повторного кримінального правопорушення або потенційної майбутньої злочинної поведінки, що може вплинути на рішення про заставу, винесення вироку або обрання запобіжного заходу; 4) для створення мап злочинності та гарячих точок, що може допомогти правоохоронним органам ефективніше спрямовувати свої ресурси; 5) для виявлення закономірностей у злочинній діяльності з метою розкриття та запобігання вчиненню кримінальних правопорушень.

В сфері пеналізації відсутній будь-який протокол (правила, формула, алгоритм, метод) з приводу визначення на законодавчому рівні виду і розміру покарання за конкретне кримінальне правопорушення, тож кореляції можуть сприяти цьому, але використання кореляцій під час індивідуального призначення покарання вбачається невинуватим. На підставі кореляцій не може бути притягнута до відповідальності будь-яка особа до того, як вона вчинила конкретний вчинок, інакше це спростує ідею презумпції невинуватості, що покладена у підґрунтя правової системи.

Існує небезпека “диктатури” Великих Даних, коли даним надається більше змісту та значення, ніж всьому іншому, коли захоплення кореляціями на підставі непрозорих алгоритмів здатне привести до критичних зловживань. З появою Великих Даних людство вступає у новий світ, в якому складне дослідження непереконаливої причинності може поступатися спіранню на непрозорі кореляції, де непрозорість рішення уповноваженої особи або органу базується на висновку “чорної скриньки” непрозорого алгоритму.

Кореляції не є беззаперечним доказом причинно-наслідкового зв'язку і повинні використовуватися в поєднанні з іншими типами доказів та відповідною юридичною аргументацією. Кореляції не мають бути використані в якості аргументу, який має переважну силу над іншими. Етичні проблеми використання кореляцій можуть бути пов'язані питаннями справедливості. Тому до використання кореляцій у сфері права слід підходити з обережністю, підкріплюючи рішення ретельним аналізом всіх обставин, глибоким розумінням не тільки правових принципів, але й основ роботи алгоритмів. Задля досягнення цієї мети правники мають оволодіти базовими знаннями з програмування, тестування, інженерії оперативного управління (prompt engineering) тощо. Основними недоліками та вадами використання кореляцій в сфері права можуть бути наступні: 1) кореляція не вказує на причинно-наслідковий зв'язок і не доводить його існування; 2) ризик упередженості; 3) відсутність контексту; 4) обмежена сфера



застосування; 5) неправильна інтерпретація; 6) недостатня якість даних; 7) неточність шаблонів; 8) порушення конфіденційності; 9) втручання в сферу захисту інтелектуальної власності та охоронюваних таємниць; 10) неналежний рівень прозорості і підзвітності; 11) дискримінаційні практики; 12) етичні міркування.

Вбачається доцільним встановити запобіжники захисту від застосування висновків на підставі кореляцій та кількісних показників для тих осіб, які можуть через них зазнати утисків, шкоди або іншого порушення прав та свобод. У більшості випадків потрібно не стільки адаптувати старі правила та підходи, скільки розробляти та встановлювати нові.

Призначення Великих Даних та кореляцій на їх підставі здебільше полягає в інформуванні, ніж у поясненні. Оскільки дослідження кореляцій не повинне підмінювати собою встановлення причинності, то за критерієм точності ієрархія між причинністю та кореляцією повинна мати наступний вигляд: 1) кореляція (менша точність, одна з передумов встановлення причинності); 2) причинність (більша точність, яка може спиратися на кореляцію).

Кореляція повною мірою є ефективною тоді, коли є необхідність дослідити явище безвідносно до його причини. Найбільш вдало кореляції на підґрунті Великих Даних можуть бути використані у тих випадках, коли не є принциповою індивідуалізація. Натомість, їх необхідно вкрай обережно використовувати там, де індивідуалізація є дуже важливою (всі персональні рішення, зокрема, щодо притягнення до кримінальної відповідальності, звільнення від неї, призначення покарання або звільнення від нього тощо). Вдосконалення захисту прав і свобод людини у випадку використання кореляцій має ґрунтуватися на наступних принципах: 1) відкритість, яка має стосуватися всього набору даних, критеріїв та алгоритму їх обробки; 2) розкриття всіх факторів, які можуть спростувати результати кореляції; 3) можливість одержання пояснень, в тому числі щодо логіки процесу виконання; 4) можливість повторного відтворення та спростування.

Кореляції на підставі Великих Даних, як й інший інструментарій народних та наукових традицій, дозволяють краще зрозуміти поточні та майбутні ризики, скорегувати дії в той чи інший спосіб. Але вони не замінюють собою причинності і не утворюють причинність на майбутнє. Під час використання кореляцій на підставі Великих Даних повинна бути забезпечена гарантія базування індивідуальної відповідальності на об'єктивних проявах поведінки на відміну від покладання на прогнози. Право має бути готовим до наступу епохи Великих Даних на презумпцію невинуватості у часи, коли світогляд, що базується на підґрунті з'ясування причини, може повільно поступатися світогляду, який спирається на Великі Дані та кореляції.

**Перспективи подальших наукових досліджень.** Порушені питання та надана їм авторська оцінка є дискусійними та відкритими для конструктивної критики і широкого обговорення з огляду на їх актуальність та важливість для забезпечення подальшого розвитку інформаційного суспільства.

### Використана література

1. Angwin, Julia. *Dragnet Nation: A Quest for Privacy, Security, and Freedom in a World of Relentless Surveillance*. New York: Times Books, 2014. 304 p.
2. Allen, Michael. *Stranger than Science Fiction: Edwin Black, IBM, and the Holocaust. Technology and Culture*. The Johns Hopkins University Press. Vol. 43. No.1 (Jan., 2002), pp. 150-154. URL: <https://www.jstor.org/stable/25147861> (дата звернення: 30.03.2023).
3. Black, Edwin. *IBM and the Holocaust: The Strategic Alliance Between Nazi Germany and America's Most Powerful Corporation*. New York: Crown, 2001. 519 pp.

4. Berk, R., & Hyatt, J. (2015). Machine Learning Forecasts of Risk to Inform Sentencing Decisions. *Federal Sentencing Reporter*, 27(4), 2015. pp. 222-228. URL: <https://doi.org/10.1525/fsr.2015.27.4.222> (дата звернення: 30.03.2023).
5. Berk, R.A., Barnes, G., Jenkins, J., & Sherman, L. Forecasting murder within a population of probationers and parolees: a high stakes application of statistical learning. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 181(1)/2018, 127-155 pp. URL: <https://www.jstor.org/stable/30136747> (дата звернення: 30.03.2023).
6. Citron, Danielle Keats. Hate Crimes in Cyberspace. Harvard University Press, 2014. 352 p.
7. Damghani, B.M. The Non-Misleading Value of Inferred Correlation: An Introduction to the Cointelation Model. *Wilmott*, Volume 2013, Issue 67, Sept. 2013. pp. 50-61. URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2429120](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2429120) (дата звернення: 01.04.2023).
8. Ferguson, A.G. The rise of big data policing: Surveillance, race, and the future of law enforcement. New York: NYU Press, 2017. 272 p.
9. Friend, Zach. "Predictive Policing: Using Technology to Reduce Crime". *FBI Law Enforcement Bulletin*, Federal Bureau of Investigation, 8 February 2018. URL: <https://leb.fbi.gov/articles/featured-articles/predictive-policing-using-technology-to-reduce-crime> (дата звернення: 01.04.2023).
10. Griffard, Molly. A Bias-Free Predictive Policing Tool?: An Evaluation of the Nypd's Patternizr. *Fordham Urban Law Journal*, Dec. 2019, 47 (1): 43-83. URL: <https://login.proxylib.csueastbay.edu/login?qurl=https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=a9h&AN=140930218&site=ehost-live&scope=site> (дата звернення: 01.04.2023).
11. Hart H.L.A., Leslie Green, Joseph Raz, Penelope A. Bulloch. The Concept of Law. Oxford University Press, Clarendon Law Series, 3rd edition, December 29, 2012. 380 p.
12. Harcourt, B.E. The collapse of the harm principle. *Journal of Criminal Law and Criminology*, 101(3)/2011, pp. 761-803. URL: <https://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=7037&context=jclc> (дата звернення: 30.03.2023).
13. Harcourt, B.E. The counterrevolution: How our government went to war against its own citizens. Basic Books, Feb 27, 2018. 336 p.
14. Nikolić, D., Murešan, R.C., Feng, W. and Singer, W. Scaled correlation analysis: a better way to compute a cross-correlogram. *European Journal of Neuroscience*, 35/2012: 742-762 pp. URL: <https://onlinelibrary.wiley.com/doi/10.1111/j.1460-9568.2011.07987.x> (дата звернення: 30.03.2023).
15. Kahneman, Daniel. Thinking, Fast and Slow. New York: Farrar, Straus and Giroux, 2011. 512 p.
16. Lane, Nick. The Vital Question: Energy Evolution and the Origins of Complex Life. New York: W.W. Norton & Company, 2016. 360 p.
17. Langford, Malcolm; Madsen, Mikael Rask: *France Criminalises Research on Judges, Verf Blog*, June 22, 2019. URL: <https://verfassungsblog.de/france-criminalises-research-on-judges> (дата звернення: 30.03.2023).
18. Levine E.S., Jessica Tisch, Anthony Tasso, Michael Joy. The New York City Police Department's Domain Awareness System. *Interfaces* 47(1)/2017, pp. 70-84. URL: <https://pubsonline.informs.org/doi/10.1287/inte.2016.0860> (дата звернення: 30.03.2023).
19. Master Card Spending Pulse TM. URL: <https://www.mastercardservices.com/en/solutions/data-networks/spendingpulse> (дата звернення: 30.03.2023).
20. Mayer-Schönberger, Viktor Kenneth Cukier. Big Data: The Essential Guide to Work, Life and Learning in the Age of Insight. HachetteUK, 2013. 256 p.p.
21. Moore, Michael S. Causation and Responsibility: An Essay in Law, Morals, and Metaphysics. Oxford University Press, 2009. 530 p.
22. Nelsen, Arthur. Feature-Pushback against AI policing in Europe heats up over racism fears. *Thomson Reuters Foundation*, October 20, 2021. URL: <https://www.reuters.com/article/europe-tech-police-idINL8N2R92HQ> (дата звернення: 30.03.2023).
23. Perry, Walter L., Brian McInnis, Carter C. Price, Susan Smith, and John S. Hollywood, Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations. *Santa Monica, CA: RAND Corporation*, 2013. URL: [https://www.rand.org/pubs/research\\_reports/RR233.html](https://www.rand.org/pubs/research_reports/RR233.html) (дата звернення: 30.03.2023).

24. RICAS (Realtimeintel lidence crime analytics system) інтелектуальна система кримінального аналізу даних. URL: <https://ricas.org/uk> (дата звернення: 30.03.2023).
25. Rienks R. Predictive Policing: Taking a chance for a safer future. *Issuu*, Jul 14, 2015. URL: [https://issuu.com/rutgerrienks/docs/predictive\\_policing\\_rienks\\_uk](https://issuu.com/rutgerrienks/docs/predictive_policing_rienks_uk) (дата звернення: 30.03.2023).
26. Thurm, Scott. Next Frontier in Credit Scores: Predicting Personal Behavior. *Wall Street Journal*, October 27, 2011. URL: <http://online.wsj.com/article/SB10001424052970203687504576655182086300912.html> (дата звернення: 30.03.2023).
27. Vlahos J. The department of pre-crime. *Scientific American*, 2012 Jan; 306(1): 62-70. URL: <https://pubmed.ncbi.nlm.nih.gov/22279837> (дата звернення: 30.03.2023).
28. Wright, Richard W. Causation in the law. *Routledge Encyclopedia of Philosophy Online*, Taylor and Francis, 2001. URL: <https://www.rep.routledge.com/articles/thematic/causation-in-the-law/v-1> (дата звернення: 30.03.2023).
29. Вільчек, Френк. Основи. 10 ключів до реальності ; пер. з англ. Миколи Климчук. Київ: Лабораторія, 2021. 208 с.
30. Карташов М.В. Імовірність, процеси, статистика: посібник. Київ: Видавничо-поліграфічний центр “Київській університет”, 2008. 494 с.
31. Радутний О.Е. Звичай як джерело кримінального права. *Держава і право: збірник наукових праць. Юридичні і політичні науки. Вип. 45.* Київ: Ін-т держави і права ім. В.М. Корецького НАН України, 2009. 668 с. С.130-135.
32. Радутний О.Е. Стан інформаційно-законодавчої діяльності на прикладі Кримінального кодексу України. *Інформація і право. № 3(18)/2016.* С. 58-67 .
33. Радутний О. Суб’єктність штучного інтелекту у кримінальному праві. *Право України. 1/2018.* С. 123-136 .
34. Радутний О.Е. Мораль і право для штучного інтелекту та цифрової людини: закони робототехніки та “проблема вагонетки”. *Інформація і право. № 3(30)/2019.* С. 78-96.
35. Радутний О.Е., Якулявічене Л. Права людини крізь призму штучного інтелекту, роботизації та цифрової людини / Права людини в умовах цифрової трансформації суспільства : монографія / [Д.В. Лученко, О.В. Капліна, В.Я. Настюк та ін.] / за ред. проф. Д.В. Лученка. Харків : НЮУ імені Ярослава Мудрого, 2022. 272 с. С. 19-41.
36. Радутний О.Е. Публічні заклики та заперечення як форми інформаційної колабораційної діяльності за Кримінальним кодексом України. *Інформація і право. № 2(41)/2022.* С. 99-116.
37. Радутний О.Е. Великі дані: поняття, ознаки та виклики (кримінально-правовий аспект). *Інформація і право. № 1(44)/2023.* С. 88-105.

~~~~~ \* \* \* ~~~~~