

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: <https://orcid.org/0000-0002-6521-3977>.

ВЕЛИКІ ДАНІ: ПОНЯТТЯ, ОЗНАКИ ТА ВИКЛИКИ (КРИМІНАЛЬНО-ПРАВОВИЙ АСПЕКТ)

Анотація: В статті розглянуто зміст та ознаки феномену Великих Даних, поняття яких сьогодні перебуває у динамічному розвитку. Досліджено такі ознаки, як безпрецедентно значний обсяг інформації без жодних втрат даних, прихований потенціал повторної цінності, перехід до кореляцій замість причинності тощо. Оглянуто джерела Великих Даних. Акцентовано увагу на ефекті поєднання даних з різних джерел з метою підвищення якості самої інформації та ефективності її обробки. Стверджується, що через великий обсяг інформації прерогатива прийняття рішень поступово зміщується від людини до алгоритму, який набирає ознак суб'єкта правових відносин. Розглянуто співвідношення права на забуття та його протилежності "постійної пам'яті", а також конфлікт права на забуття з ідеологією Великих Даних. Доведено, що можливість спостереження за окремою особою сьогодні вбудована практично у всі алгоритми, програмні та апаратні засоби, технологічна природа яких дозволяє обмінюватися даними з виробником, оператором та(або) іншими зацікавленими особами, не вимагаючи втручання з боку їх власника або користувача.

Розглянуто варіанти інформованої згоди у випадку зміни мети використання персональних даних. Доведено, що в окремих випадках персональні дані виникають як побічний ефект функціонування комплексу складних компонентів кібернетичної системи. Стверджується, що з появою Великих Даних поступово втрачають свою ефективність такі стратегії забезпечення конфіденційності як інформована згода, диференціальна конфіденційність, відмова від участі та анонімізація.

Запропоновано розглядати суб'єкта персональних даних в якості точки перетинання інформаційних потоків та як сукупність соціальних зв'язків, у зв'язку з чим для одержання вичерпної інформації про нього стає достатнім проаналізувати та систематизувати всі доступні периферійні неперсональні дані. Доведено, що вилучення певної інформації з набору даних може дорівнювати її наявності. Досліджено окремі різновиди кримінальних правопорушень в сфері Великих Даних та засоби протидії їм. Сформульовано різновиди викликів, які можуть виникнути у зв'язку з використанням Великих Даних, зокрема, порушення приватності, конфіденційності та деанонімізація, заміна причинності на кореляцію в юридичній сфері.

Ключові слова: Великі Дані, інформація, інформаційний потік, повторна цінність інформації, алгоритм, штучний інтелект, правова персона, персональні дані, інформована згода, право на забуття, постійна пам'ять, диференціальна конфіденційність, відмова від участі, анонімізація, деанонімізація, кримінальне правопорушення, причинність, кореляція.

Summary: The article examines the content and features of the Big Data phenomenon, the concept of which is currently undergoing dynamic development. The author analyzes such features as an unprecedentedly large amount of information without any data loss, hidden potential for repeated value, transition to correlations instead of causality, etc. The sources of Big Data are reviewed. The author emphasizes the effect of combining data from different sources in order to improve the quality of the information itself and the efficiency of its processing. It is argued that due to the large amount of information, the prerogative of decision-making is gradually shifting from a person to an algorithm

which acquires the characteristics of a subject of legal relations. The author analyzes the correlation between the right to be forgotten and its opposite, “permanent memory”, as well as the conflict between the right to be forgotten and the ideology of Big Data. It is proved that the ability to monitor an individual is now built into virtually all algorithms, software and hardware, the technological nature of which allows data to be exchanged with the manufacturer, operator and/or other interested parties without requiring intervention by their owner or user.

The author considers the options for informed consent in case of changing the purpose of personal data use. It is proved that in some cases personal data arise as a side effect of the functioning of a complex of integrated components of a cybernetic system. It is argued that with the advent of Big Data, such privacy strategies as informed consent, differential privacy, opt-out and anonymization are gradually losing their effectiveness.

The author proposes to consider a personal data subject as a point of intersection of information flows and as a set of social ties, and therefore, to obtain comprehensive information about him or her, it is sufficient to analyze and systematize all available peripheral non-personal data. It is proved that the removal of certain information from a dataset may be equal to its presence. Certain types of criminal offenses in the field of Big Data and means of counteracting them are investigated. The author formulates the types of challenges that may arise in connection with the use of Big Data, in particular, violation of privacy, confidentiality and deanonymization, replacement of causality with correlation in the legal sphere.

Keywords: *Big Data, information, information flow, information reuse value, algorithm, artificial intelligence, legal person, personal data, informed consent, right to be forgotten, permanent memory, differential privacy, opt-out, anonymization, de-anonymization, criminal offense, causation, correlation.*

Постановка проблеми. Зростаючі за кількістю та якістю апаратні та програмні засоби генерують величезний масив даних, які разом з іншою інформацією, що свідомо і цілеспрямовано збирається, використовується та зберігається, утворюють феномен Великих Даних (Big Data), формальне визначення яких сьогодні перебуває у динамічному розвитку. Між тим, значна кількість джерел [1; 2; 5] пропонує практично однакове тлумачення Великих даних як наборів структурованої або неструктурованої інформації, настільки великих розмірів, що стандартні методи, способи та підходи, які у більшості випадків побудовані на рішеннях аналітики та системах управління базами даних, не можуть бути до них застосовані. В якості типової характеристики Великих Даних вказують [35, с. 82-92] на їх фізичний обсяг, швидкість зростання та можливість одночасної обробки різних типів даних. Тож, Великі Дані є сучасним, складним і системним явищем, вплив якого на суспільство та право має бути дослідженим та ефективно скерованим, разом з одночасним запобіганням появі плетива негативних обставин, які неодмінно супроводжують кожне досягнення науково-технічного прогресу.

Завдяки розвитку технологій сьогодні стає доступним для одержання, фіксації, архівування, вимірювання та іншого використання той значний обсяг інформації, який раніше – 1) не мав можливості бути сприйнятим людиною через обмеження її перцептивної системи [35, с. 82-92]; 2) неможливо було належним чином одержати, зафіксувати, зосередити та зберегти з наперед невизначеною метою; 3) перебував у доступі, але залишався поза увагою людини через відсутність зацікавленості або чітко сформульованої мети обробки тощо.

Джерелами Великих Даних виступають прилади і сенсори промислових та інших об'єктів, технологічні лінії виробництва, інфраструктурні системи (енергетичні, комунікаційні, торговельні, транспортні, охоронні тощо), соціальні мережі, мобільний зв'язок, будь які побутові датчики контролю та управління, устаткування відео-спостереження та медичних обстежень, бази даних підприємств, установ та організацій

тощо. До Великих Даних також можливо віднести будь-які різновиди інформації, в тому числі її уривки, координати GPS, запити до пошукових систем та відповіді на них, тривалість часу перегляду певного контенту, реакції на нього (уподобання, невподоба, схвалення, заперечення, коментарі) у відкритому або анонімному режимі тощо.

Завдяки різноманітності та мережевому характеру джерел у суспільствах високого ступеню технологічного розвитку інформація відстежується, фіксується, генерується, обробляється та зберігається у таких безпрецедентних масштабах, з якими людство ще не мало справи. Особливо показовим є те, що інформація у значних обсягах накопичується не тому, що її хтось замовляє або відстежує, але через доступність сучасних технологій, адже у значній кількості випадків є більш простим і менш витратним фіксувати та зберігати всю вхідну інформацію (наприклад, з камери відеоспостереження або з сенсору у розумному будинку), ніж фільтрувати та розподіляти її за певними критеріями.

З іншого боку, обсяги та швидкість генерації і одержання певних даних обумовлюють вимушеність зберігання їх в первинному сирому (невпорядкованому, швидкому, неточному, неструктурованому) вигляді як кращу альтернативу втрачання. Таким чином, відповідно до ідеології Великих Даних вся інформація, яка надходить або генерується і потрапила у систему обробки у той чи інший спосіб, має залишатися там і бути збереженою про всяк випадок на максимально тривале невизначене майбутнє.

Такий підхід до організації та здійснення інформаційної діяльності принципово відрізняється від всіх попередніх, аналогічно тому, якби несподівано кожна крапля питної води почала використовуватися людством максимально ефективно без жодних втрат завдяки появі передовій технології під умовною назвою Big Water (Великі Води).

Поряд з безпрецедентно значним обсягом, ще однією надзвичайною властивістю Великих Даних є можливість повторного використання інформації, яка була раніше здобута та збережена з іншою метою, але продовжує містити в собі певну неочевидну цінність. Виявлення нового аспекту такої цінності раніше одержаної та збереженої інформації розширює горизонти для ефективної обробки даних, але породжує нові правові виклики.

Інформаційна діяльність під гаслом “не втрачати нічого, використовувати всі види даних і стільки разів, скільки це буде визнано ефективним сьогодні або у майбутньому” [34, с. 72] відкриває нову епоху початкової та повторної цінності будь-якої інформації, виняткової прозорості та деанонізації, заміни підґрунтя для прийняття юридично значущих рішень з встановлення причинності до з’ясування кореляцій.

Тому в якості спроби подальшого розвитку понятійного апарату, вбачаються підстави запропонувати розглядати Великі Дані як комплексний феномен, що включає: 1) надзвичайно великий обсяг інформації, який перевищує всі раніше доступні для одержання, зберігання та обробки; 2) можливість використовувати повний обсяг інформації, замість окремої частини (вибірки); 3) здатність виявляти та використовувати повторну цінність будь-якої інформації, яка одержується та зберігається, в тому числі про всяк випадок; 4) нові підґрунтя для одержання умовиводів, зокрема, перехід від встановлення причинності до виявлення кореляцій.

Як і будь-яке досягнення науково-технічного прогресу, явище Великих Даних само по собі не є ані позитивним, ані негативним, натомість виступає логічним результатом людської діяльності з пізнання оточуючого світу. Між тим, воно: 1) реально існує і набирає оберти, тому необачно робити інший вигляд; 2) має здатність глибинно, масштабно та у непередбачуваний спосіб вплинути на більшість аспектів життя та діяльності людини, включаючи галузі економіки, політики, культури і права; 3) право

може зазнати змін фундаментального характеру, що матиме наслідком перегляд багатьох концепцій та підходів до розв'язання існуючих та майбутніх проблем.

Відбувається поступовий перехід від часів “малих даних” до новітньої пори Великих Даних. Старі правила, які діяли до цього і довели свою ефективність, мають бути критично переглянуті та змінені на користь суспільству з урахуванням актуальних реалій.

Метою статті є визначення змісту та ознак феномену Великих Даних з точки зору права, в тому числі у межах кримінально-правового аспекту, виявлення співвідношення ідеології Великих Даних з традиційними інформаційними політиками, принципами та стратегіями, зокрема, щодо забезпечення прав в інформаційній сфері шляхом застосування практик інформованої згоди, диференціальної конфіденційності, відмови від участі та анонімізації, а так само з усталеними підходами до визначення умов збирання, зберігання та використання різних видів інформації, виявлення початкової, повторної та потенційної цінності, формулювання мети обробки даних тощо.

Результати аналізу наукових публікацій. Дослідження феномену 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, К.І. Беляков, Д. Білик, С.М. Браїчевський, 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 зусиллями необмеженої кількості пошукувачів.

Виклад основного матеріалу.

Безпрецедентно значний обсяг даних. Інформація була, є і буде обертатися повсюди, в той час як однією з основних проблем залишається нездатність її належним чином сприйняти, зафіксувати, зберегти та ефективно використати. Втім, сьогодні обсяги інформації і можливості їх обробки в різній спосіб збільшуються у складній прогресії. Так, з моменту старту у 2000 р. проекту “Слоанівський цифровий огляд неба” (Sloan Digital Sky Survey, SDSS)¹ його телескоп за перші декілька тижнів зафіксував більше даних, ніж людство накопичило на протязі всієї історії астрономії [6]. У 2003 р. вперше був розшифрований геном людини, після чого ще десять років інтенсивних досліджень було витрачено на побудову послідовності з трьох мільярдів основних пар, але через наступні десять років та сама кількість ДНК аналізується кожні 15 хвилин [26]. На фондових ринках США кожного дня відбувається приблизно сім мільярдів обмінних операцій, з яких дві третини здійснюються за допомогою алгоритмічного або високочастотного трейдингу (algorithmic, high-frequency trading – комп'ютерних алгоритмів на базі математичних моделей) [13]. З огляду на дійсно великий обсяг інформації та необхідність прискореного реагування (секунди, мілісекунди або їх ще більш дрібні частки) рішення в такій конфігурації сил приймають не люди, але алгоритми. На Facebook завантажується більше десяти мільйонів нових фотографій кожну годину, коментарі та реакції на будь-яке повідомлення у кількості три мільярди

¹ Слоанівський цифровий огляд неба (Sloan Digital Sky Survey, SDSS) – масштабне дослідження зображень, спектрів та червоного зміщення галактик за допомогою 2,5-метрового оптичного телескопу в обсерваторії Апач-Пойнт (Нью-Мексико, США). Проект названо на честь американської філантропічної громадської організації Alfred P. Sloan Foundation.

на день залишають після себе потужний цифровий слід, який дозволяє вивчити вподобання користувачів так само ефективно, як і ідентифікувати анонімну особу [12].

Донедавна через брак технічних та інших можливостей загальний обсяг здобутої інформації примусово скорочувався до зручного мінімуму і вже потім піддавався відповідній обробці. Сьогодні є можливість збирати та обробляти (аналізувати, використовувати з первинною метою, зберігати, повторно аналізувати та використовувати тощо) всю інформацію, яку тільки є можливим залучити в процес. Фактично стають непотрібними вибірки, за допомогою яких раніше вирішували проблему відносної об'єктивності висновків, адже завдяки феномену Великих Даних з'являється можливість досліджувати дійсно повний обсяг інформації замість його окремої репрезентативної частини. Відмова від вибірок усуває ще один їх недолік, крім неповноти, що полягає у необхідності формування нового набору даних тоді, коли змінюється мета або виникає нова, і це вимагає витрачання додаткових ресурсів і зусиль з збирання і обробки інформації.

Можливості доступу до повного обсягу інформації та його опрацювання обумовлюють досягнення помітних результатів. Так, в сфері протидії злочинності використання Великих Даних замість ретельно скоординованої вибірки продемонструвало свою ефективність на прикладі виявлення нетипової поведінки як ознаки шахрайських дій з банківськими картками. Повне охоплення надало можливість виявити значення, які істотно відрізняються від інших, і саме це привернуло до них увагу і вказало на їх підозрілий характер. Система обробки даних міжнародних переказів компанії Хоом успішно виявила незначне перевищення середньої кількості транзакцій з використанням кредитних карт Discover Card в Нью-Джерсі та встановила певну закономірність, якої не повинно було бути. Досліджуваною активністю були дії зловмисників, які неможливо було би виявити поза вивчення всього обсягу інформації про всі транзакції в реальному часі та у необмеженій кількості [29].

У свою чергу, спостереження за нелінійними кореляціями на підставі Великих Даних привело до важливого висновку в сфері внутрішньої політики: щоб зробити населення країни щасливим, краще зосередитися на збільшенні доходів незаможних прошарків населення, ніж збільшувати доходи для всіх [15]. Якщо численні приклади підтверджують позитивний вплив поєднання апельсинового соку з аспірином на лікування певної хвороби, то така кореляція має право на зважене використання, навіть за відсутність розуміння причини позитивного ефекту.

В епоху Великих Даних стає важливим їх масштаб, адже чим більше інформації вдалося накопичити, тим більшу цінність можливо з неї одержати. Зміна масштабу веде до зміни властивостей та цінності певної інформації за принципом переходу кількості у якість. Таким чином, повний масив даних замість вибірки дає можливість працювати на нечувано глибокому рівні розуміння будь якої проблеми.

Тому нове мистецтво роботи з інформацією, зокрема з Великими Даними або алгоритмами штучного інтелекту, полягає у влучному формулюванні запиту (влучна постановка питання), розробці та застосуванні ефективних способів його обробки. У протилежному випадку значний обсяг інформації залишиться просто значним обсягом “сирих” (невпорядкованих, неточних, неструктурованих) даних, які не надають жодних відповідей, хоча і можуть їх містити.

Так, свого часу Amazon мала значний обсяг інформації про кожного клієнта (зокрема, про книжки, які були переглянуті або придбані, про час, який був витрачений на перегляд, про одночасні покупки тощо). Вказаний обсяг даних являв собою незначну цінність до того моменту, як був розроблений та запатентований метод сумісної фільтрації “предмет-предмет” для рекомендацій відвідувачам нових книг, фільмів та

інших товарів. Рекомендації від алгоритму значно перевищили за своєю ефективністю роботу відділу літературних редакторів та критиків, у зв'язку з чим останній згодом був розформований, адже для результату виявилось неважливим, чи розуміє застосунок справжню причину здійснення покупцем свого наступного вибору, або просто безпомилково його вгадує [34, с. 81].

У наведених прикладах цілком помітним вбачається зсув акценту з пошуку відповіді на питання “чому” (чому покупець здійснює саме такий вибір, чому відбувається певна транзакція, чому хвороба піддається кращому лікуванню тощо) на з'ясування змісту відповіді на питання “що” (який вибір покупець скоріш за все здійснить наступного разу). Така зміна пріоритетів є одним з принципів роботи з Великими даними: розуміння причини поступається прийняттю рішень на підставі кореляції, яка у свою чергу не може бути виявлена без використання Великих Даних.

Кореляція (correlation) або залежність (dependence) є позначкою статистичного взаємозв'язку між двома випадковими змінними або двовимірними даними, кількісним вираженням статистичного зв'язку між двома значеннями [37, с. 56]. Сильна кореляція вказує, що за зміну одних значень інші значення теж зміняться, слабка кореляція означає, що навіть за зміну одних значень інші практично не зміняться.

Кореляції допомагають 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], наприклад, між розміром взуття та довгостроковою пам'яттю, або улюбленим кольором та станом здоров'я. Навіть сильні кореляції жодним чином не пояснюють певне явище, адже вони пропонують не визначеність, але вірогідність. Оскільки така вірогідність може бути доволі високою, цінність кореляції пропорційно збільшується, що утворює підґрунтя для ефективного прогнозування. Так, якщо подія А супроводжується подією В, то спостереження за подією В дозволяє спрогнозувати появу події А. Але, що важливо і не є зайвим повторити, кореляційна вірогідність не вказує на причину і не підмінює її, втім може бути етапом встановлення останньої.

Таким чином, у більшості 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” [22]. В галузі права слід бути більш обережним з використанням кореляцій на підставі Великих Даних, ніж в сфері маркетингу, політичних або економічних прогнозів, але з огляду на обсяг і складність цього питання воно буде розглянуте у наступних публікаціях у розвиток цього дослідження.

Великі Дані також дозволяють побачити та використати нові зв'язки між тими окремими фрагментами інформації, які на перший погляд не демонструють жодної взаємопов'язаності між собою. За допомогою Великих Даних стає можливим побудувати складні ланцюжки взаємозв'язків всього з усім. Крім того, об'єднання різних даних з метою їх дослідження на предмет виявлення певних взаємозв'язків є ще одним з способів управління Великими даними, підвищення якості самої інформації та ефективності її обробки. Так, згода різних автовиробників на встановлення відповідних датчиків та надання інформації з них незалежній компанії Inrix, яка аналізує дорожній рух в режимі реального часу шляхом використання даних геолокації від транспортних засобів в більше ніж одній тисячі міст на семи континентах та у п'ятдесяти країнах, значно збільшила обсяги даних та якість їх аналізу на противагу тому, якби кожний виробник збирав та обробляв таку інформацію самостійно [16].

Зворотній бік алгоритмізації полягає у тому, що через дійсно великий обсяг інформації реальні важелі у прийнятті рішень поступово зміщуються від людини до алгоритму, зокрема, штучного інтелекту, який поступово та впевнено набирає ознак суб'єкта правових відносин [38, с. 123-136; 39, с. 19-41; 40, с. 78-96].

Відновлювальна цінність Великих Даних через повторне використання інформації. Завдяки суттєвому зменшенню вартості зберігання одиниці інформації (адже за останні п'ятдесят років вона зменшується вдвічі кожні два роки, у той час як щільність архівації збільшилася у п'ятдесят мільйонів разів [34, с. 63]) тотальне збереження даних стає більш привабливим, ніж їх знищення. Справжнє масштабне значення Великих Даних полягає у поєднанні чималих обсягів інформації з можливістю їх повного зберігання та повторного використання. Іноді останнє може відбуватися навіть з більшою ефективністю, ніж під час виконання первісних завдань. Сучасні технології успішно подолали межу значної кількості обмежень в сфері збирання та зберігання даних. Істотно більшу кількість даних стає можливим збирати пасивно та без надмірних зусиль. Інформацію, яка надходить за умовчанням (наприклад, від сенсору, датчику або камери спостереження), простіше одержувати у повному обсязі, ніж фільтрувати за допомогою складних програмних або апаратних рішень.

З свого боку, належне зберігання великих масивів інформації та доступність цих сховищ утворили підґрунтя для повторного використання будь яких даних, що будь коли були генеровані або зібрані. Це означає, що в епоху Великих Даних поступово зникають різновиди інформації з нульовою цінністю. Завдяки можливості повторного використання будь-які дані є цінними самі по собі з огляду на далекоглядну перспективу, оброблені та необроблені (невпорядковані, неточні, неструктуровані) дані стають важливими факторами формування своєї загальної продуктивності.

Тож, абсолютна цінність інформації не може бути визначена наперед і здатна перевищувати значущість мети її первинного збирання та обробки. Так, первинне використання складної моделі ІВМ, що поєднувала багатопланові вхідні дані (рівень заряду акумулятора, місцезнаходження автомобіля, час доби, доступні та найближчі станції підзарядки, прогноз погоди тощо) з інформацією про поточне споживання електричної енергії в мережі та статистичними даними про закономірності такого споживання [14], утворило підґрунтя для ефективного визначення оптимального місця і часу підзарядки електромобіля, в той час як вторинне використання – місце для побудови нових станцій для обслуговування [25]. В оглядній перспективі цінність даних може знайти своє відображення в фінансовому обліку як новий різновид активів та ресурсів.

Таким чином, завдяки Великим даним цінність будь-якої інформації необхідно розглядати з точки зору всіх відомих або невідомих, потенційних та альтернативних можливостей її використання. Виявлення нових аспектів цінності збереженої інформації є додатковим різновидом мистецтва обробки Великих Даних, що вимагатиме існування відповідних фахівців та професій. Однією з них є спеціаліст з обробки даних (Data Scientist), що поєднує в собі навички програмування, дизайну, аналітики, статистики та інфографіки, і відповідно до звіту WEF Future of Job [31] разом з фахівцями з Machine Learning буде найбільш витребуваним на ринку праці до 2025 року. Сьогодні у напрямку виявлення нових аспектів цінності інформації проводять дослідження практично всі суб'єкти обробки та використання персональних або інших даних, а так само об'єднання зацікавлених осіб (зокрема, Sunlight Foundation в США з 2006 по 2020 р.р. [32], Open Knowledge Foundation в Великобританії [23] тощо).

Великі Дані та право на забуття. Право на забуття (право бути забутим, right to be forgotten) є протилежністю “постійній пам’яті” (гарантуванні того, що ніхто і ніщо не укриється від свого минулого, адже записи завжди можливо дослідити та використати [34, с. 136]), та правом, що дозволяє особі з різних міркувань вимагати видалення своїх персональних або інших даних із сфери загального доступу, пошукових систем тощо. Поряд з визнанням та реалізацією такого права важливим вбачається врахування та забезпечення необхідного балансу з загальними соціальними інтересами (громадська або національна безпека, рецидив як різновид повторності, право громадськості бути проінформованою щодо особи публічного діяча тощо). У зв’язку з цим триває дискусія щодо доцільності законодавчого закріплення права на забуття через побоювання, що воно суперечить таким визнанням фундаментальними правам людини, як свобода слова, свобода доступу до інформації, а також через можливий негативний вплив з точки зору цензури та фальсифікації історії [17; 18; 27]. Її гострота була підсилена рішенням ЄСПЛ у справі “Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja González” 2014 р. [10] і має тенденцію до збільшення через появу додаткового аргументу зі сфери Великих Даних: будь яка інформація має зберігатися максимально довгі строки для того, щоб завдяки повторному використанню підвищити її цінність.

Поряд з цим існують сумніви з приводу збереження своєї ефективності таким способом забезпечення приватності, як “диференціальна конфіденційність” [8], що передбачає свідоме розмиття даних до такого рівня, коли запит великого обсягу інформації буде видавати лише приблизні результати і це на певний час має ускладнити процес зв’язування відповідних точок даних (окремих значень) з конкретним особами. Втім, уявлення про межі та критерії такого розмивання постійно змінюються. Оскільки у загальному обсязі Великих Даних значна частка інформації від початку її генерації або збирання є неперсоналізованою і у певному розумінні розмитою, то додаткові зусилля можуть взагалі позбавити дані певного сенсу, змісту та цінності.

Великі Дані та персональні дані. Факторами, які суттєво спростили та значно здешевили процес стеження за будь якими аспектами суспільного буття або за будь-якою особою, виступають Інтернет, Великі Дані та інші сучасні технології, а так само здешевлення вартості збирання та зберігання будь якої інформації, зростання потужності аналітичних інструментів тощо. Алгоритми Amazon, e-Bay, Facebook, мережі рітейлерів (“Розетка”, “Епіцентр” тощо) розуміються на вподобаннях та історії здійснених замовлень, Google відстежує пошукові запити та відвідані веб-сторінки, на платформі Twitter розміщують миттєві думки та реакції, Facebook та LinkedIn детально вибудовують мережу соціальних зв’язків та уважно спостерігають за нею тощо.

Можливість спостереження за окремою особою сьогодні вбудована практично у кожний прилад або інструмент, в тому числі у трекери активності, смартфони, розумні годинники, програмне забезпечення, застосунки, прилади, датчики та об’єкти “розумного дому” або мережі Інтернету речей (Всеосяжного Інтернету), всі інші засоби переносної електроніки (wearable electronics)² тощо. Їх технологічна природа дозволяє обмінюватися даними з виробником, оператором та(або) іншими зацікавленими особами (державою, роботодавцем, медичною установою або страховою компанією), не

² Wearable technology (переносна технологія) – технологія, призначена для використання під час носіння. До поширених типів натільних технологій належать розумні годинники та розумні окуляри. Натільні електронні пристрої часто розташовуються близько або на поверхні шкіри, де вони виявляють, аналізують і передають інформацію, таку як життєво важливі показники та/або дані про навколишнє середовище, і в деяких випадках дозволяють користувачеві отримати негайний зворотний зв’язок [9].

вимагаючи втручання з боку власника або користувача. Так, апаратні та програмні засоби відстежування активації подушки безпеки на транспортних засобах можуть зберігати докази певної поведінки водія у випадку необхідності її перевірки [33].

Комунікації в Інтернеті та, зокрема, соціальних мережах, вже поставили приватність під істотну загрозу, але кореляції на підставі Великих Даних таку загрозу збільшують на порядок і здатні за відсутності цілеспрямованого дешифрування ідентифікувати будь-якого суб'єкта інформаційних відносин.

Традиційний обсяг інформації про особу (прізвище, ім'я та по батькові, дата і місце народження, реквізити документів, банківські рахунки, етапи навчання, кар'єри та працевлаштування, освіта, професія, стать, національність, мова, ставлення до релігії тощо) завдяки соціальним мережам (Facebook, Twitter, LinkedIn тощо) доповнюється даними про особисті зв'язки, реакції (не-)уподобання, висловлені думки, розклад, пошукові запити, тривалість перегляду певного контенту тощо. Якщо персональні дані (Personal Data) є відомостями або сукупністю відомостей про фізичну особу, яка ідентифікована або може бути конкретно ідентифікована (ст. 2 Закону України "Про захист персональних даних" від 01.06.10 р. № 2297-VI), то в епоху Великих Даних вбачаються підстави визнати їх обсяг таким, що не має окреслених меж.

В значній кількості випадків персональні дані можуть виникати як побічний ефект функціонування комплексу складних компонентів кібернетичної системи, коли від початку генерації або збирання певна інформація є агрегованою в загальний масив мультимедійних даних суто технологічним способом і не має прямого відношення до персональних даних, але згодом останні виникають внаслідок її обробки та поєднання з іншою інформацією [35; с. 82-92]. Так, розумні електричні лічильники можуть збирати до трьох тисяч точок даних (окремих значень) на місяць в режимі реального часу, мають унікальний "підпис навантаження", який дозволяє відрізнити побутові прилади між собою (холодильник від телевізора, телевізор від комп'ютера або пральної машини тощо) і завдяки сукупній інформації окреслити індивідуальний малюнок поведінки певної особи, що може охоплювати персональні звички, медичні приписи, різновиди суспільно корисної або небезпечної поведінки тощо [19].

Оскільки, як було зазначено вище, цінність даних не обмежується їх первинним використанням, то повторні дії з ними з новою метою можуть являти собою порушення приписів чинного законодавства, адже на підставі положень Закону України "Про захист персональних даних" від 01.06.10 р. № 2297-VI володілець персональних даних (фізична або юридична особа) має визначити (тобто, добровільно обмежити) мету їх обробки, встановити склад цих даних та процедури вчинення певних дій з ними тощо. Ідеальною слід вважати ситуацію, коли згода на збирання та(або) використання даних містить достатньо конкретної інформації, в тому числі, про мету, тривалість зберігання та можливість повторної обробки, і не залишає місця для двозначної інтерпретації. У випадку виникнення необхідності повторного використання персональних даних, в тому числі з новою метою, не можуть братися до уваги для виправдання ігнорування її одержання такі фактори, як складність або значні витрати. Але у певних випадках зв'язок з суб'єктом персональних даних може бути втрачений з різних причин, що здатне утворити суттєві перешкоди у використанні Великих Даних відповідно до їх ідеології.

З іншого боку, якщо пропонувати і одержувати занадто широку згоду, в тому числі щодо можливості багаторазової зміни мети обробки, то це знівелює саме поняття інформованої згоди, оскільки суб'єкт персональних даних втрачає той контроль, який вона йому надає. На противагу надмірно широким або загальним формулюванням,

умови згоди повинні бути зрозумілими та конкретними, ґрунтуватися на вичерпній інформації про можливість і деталі повторного використання та(або) зміни мети, зазначати як саме і ким буде використовуватися певна інформація. У випадку, коли суб'єкт персональних даних інформовано погоджується на можливу зміну мети та повторне використання даних іншими особами, він має бути наділений правом змінити свою згоду або відкликати її у будь-який момент за певних умов.

В інших більш складних випадках через специфіку Великих Даних на ранніх етапах обробки певної інформації може зовсім не виникати питання про необхідність одержання згоди у зв'язку з відсутністю впевненості в тому, що окремі мультимедійні, мета- та інші дані мають відношення до конкретного суб'єкта персональних даних. До моменту повної деанонізації останнього дані можуть виглядати як такі, що не є персональними. Між тим, за допомогою кореляцій на підставі неперсональних даних можуть бути виявлені такі зв'язки, які з високою точністю ідентифікують певну особу, але факт порушення захисту може бути усвідомлений та виявлений добросовісним володільцем, одержувачем, розпорядником або третьою особою лише після одержання результату.

Втім, навіть вилучення певної інформації з набору даних може дорівнювати її наявності. Так, розмитий вигляд окремих будинків на Google Street View [36], який був запроваджений під тиском громадськості через побоювання привертання уваги потенційних грабіжників до заможних господарств або з інших слушних питань приватності [11], в результаті вказує правопорушникам на найбільш цікаві об'єкти, власники або користувачі яких висунули обґрунтовані вимоги щодо їх знеособлення.

Анонімізація (тобто, вилучення з наборів даних персональних ідентифікаторів, зокрема, прізвища, ім'я та по-батькові, адреси, дати і місця народження, номерів банківських рахунків, соціального забезпечення, реквізитів документів тощо) не є достатньо ефективною у випадку використання Великих Даних, які дозволяють здійснити повторне встановлення особи.

Так, для алгоритмів Netflix виявилось не занадто складним ідентифікувати особу жінки (матері двох дітей та прихованої лесбійки з консервативного Середнього Заходу США, яка здійснила заходи щодо знеособлення своїх персональних даних) під псевдонімом Jane Doe³, навіть за умови, що її особисті ідентифікатори були ретельно видалені з усіх баз даних [30]. На підставі порівняння даних від Netflix з іншими загально доступними відомостями дослідники з Техаського університету виявили, що оцінки анонімізованих користувачів відповідають оцінкам осіб з конкретними іменами на сайті Internet Movie Database (IMDb), тож їх вдалося ідентифікувати у 84 % випадків, а за наявності інформації про дату оцінки певного фільму – з точністю до 99 % [20; 21].

Так само будь-яку анонімну особу можливо повністю ідентифікувати за змістом її пошукових запитів або шляхом порівняння з даними від інших джерел, які включені в обсяг Великих Даних.

Наведені випадки демонструють недооцінку загроз у мінливому світі інформаційних технологій та необачність покладання на перманентну надійність існуючих засобів захисту. Але коли йдеться про Великі Дані, то виклик може бути аналогічним нездатності надійних сьогодні криптографічних паролів протистояти дешифруванню квантовим комп'ютером у найближчому майбутньому.

³ John Doe (англомовні ЗМІ, англо-саксонська правова сім'я Common Law Legal System) – термін, який використовується у випадку невизначеності (невстановленості, анонімності) особи, Jane Doe – еквівалент до особи жіночої статі, Baby Doe – відносно дитини.

У зв'язку з цим Пол Ом з Школи права університету Колорадо (Paul Ohm, Associate Professor, University of Colorado Law School) переконаний, що за наявності достатньої кількості даних ідеальна анонімізація неможлива всупереч будь-яким зусиллям [24]. Більше того, “соціальний граф” (тобто, зв'язки між людьми у соціальних мережах) так само вразливий до деанонімізації, якщо у процес ідентифікації задіяні Великі Дані [3].

Тож, з появою Великих Даних поступово втрачають свою ефективність такі стратегії забезпечення конфіденційності як інформована згода, відмова від участі та анонімізація. Якщо розглядати суб'єкта персональних даних в якості точки перетинання з будь-якими інформаційними потоками, оточенням, речами, приладами та контентом, а так само як сукупність соціальних зв'язків, то для одержання вичерпної інформації про нього стає достатнім проаналізувати та систематизувати всі доступні периферійні неперсональні дані, навіть на повної відсутності персоналізованої інформації.

З свого боку, кримінальні правопорушення в сфері Великих Даних можуть бути пов'язані з несанкціонованим доступом до конфіденційної інформації, її викраденням, зловживанням, порушенням правил збирання та обробки даних, а також з використанням аналітичних інструментів для незаконних цілей, наприклад, для дискримінації на основі особистих даних.

Одним з прикладів використання аналітичних інструментів з сумнівними цілями є випадок з Cambridge Analytica [4]: компанія збирала персональну інформацію мільйонів користувачів Facebook (зокрема, про їх інтереси, соціальні зв'язки, публічні записи, повідомлення, та інші дані, на підставі яких можливо відтворити вичерпний образ користувача) для створення детальних профілів, які надалі використовувалися для персоналізованих політичних рекламних кампаній та маніпулювання голосуванням на виборах. Ще одним прикладом суспільно небезпечного використання аналітичних інструментів є діяльність російської установи “Агентство Інтернет-досліджень”, більше відомої як “фабрика тролів” [28], що спеціалізується на створенні та поширенні дезінформації і пропаганди в соціальних мережах, використанні складних аналітичних інструментів для організації мережі фейкових акаунтів, виявленні вразливих для пропаганди груп людей, вимірюванні настроїв, інформаційному маніпулюванні тощо. Агентство звинувачують у втручанні в президентські вибори 2016 р. в США, а також у референдумі Brexit у Великобританії.

Використання аналітичних інструментів для незаконних цілей може також мати наступні форми:

- зловживання даними (наприклад, збирання даних про користувачів без їхньої згоди або використання даних, які були одержані з іншою метою, в тому числі для вчинення таких незаконних дій, як розсилання спаму, фішингу або для здійснення кібератак);
- фальсифікація даних (зміна або підробка даних з будь-якою метою);
- розкриття конфіденційної інформації (наприклад, дешифрування паролів, виявлення номерів соціального страхування та банківських реквізитів тощо).

Тож, з метою запобігання кримінальним правопорушенням з використанням Великих Даних та аналітичних інструментів необхідно постійно і напередженні вдосконалювати норми чинного законодавства, апаратні та програмні засоби захисту, розвивати та впроваджувати нові технології, зокрема, блокчейн, що можуть забезпечити більш високий рівень захисту від несанкціонованого доступу та зловживання, але не покладатися на жоден з існуючих варіантів в якості незмінного.

Великі Дані надають великі можливості, але долають рубежі захисту приватності та інших охоронюваних таємниць (зокрема, службової, комерційної, лікарської,

банківської, усиновлення тощо). Великі Дані та аналітичні інструменти можливо і необхідно використовувати для боротьби з неналежним використанням Великих Даних та аналітичних інструментів.

Аналіз Великих Даних здатний запропонувати свіжий погляд на існуючі проблеми та нові ідеї саме тому, що він не обтяжений межами традиційних підходів та усталеним способом мислення, а так само властивими йому упередженнями. Разом з тим, будь який новий метод, виріб чи технологію можливо використати з асоціальною метою. Тому треба звикати жити та взаємодіяти у світі, у якому стає можливим використати Великі Дані як з суспільно корисною, так і з суспільно небезпечною метою. Таким чином, щодо Великих Даних є підстави окреслити як мінімум наступні загрози: 1) порушення приватності, конфіденційності та деанонімізація; 2) несанкціонована повторна обробка інформації; 3) покладання на юридично значущу вірогідність (заміна причинності на кореляцію).

Висновки та пропозиції.

Формальне визначення Великих Даних сьогодні перебуває у динамічному розвитку, але у значній кількості випадків їх вбачається можливим розглядати як комплексне явище: 1) безпрецедентно значний обсяг інформації в структурованому та(або) неструктурованому (невпорядкованому, швидкому, неточному) вигляді, який перевищує всі раніше доступні для одержання, зберігання та обробки, до якого не можуть бути застосовані стандартні методи; 2) такий обсяг інформації, в якому не втрачаються жодні дані, вони зберігаються у повному обсязі, незалежно від того, чи є в них потреба у певний проміжок часу; 3) повний обсяг інформації замість окремої репрезентативної частини (вибірки); 4) прихований потенціал повторної цінності будь-якої інформації, що може бути виявлений після її обробки під новий запит; 5) нові методи обробки інформації, пов'язані зі збільшенням масштабу даних; 6) перехід від встановлення причинності до кореляцій як нового підґрунтя для одержання умовиводів.

Завдяки можливості повторного використання в епоху Великих Даних поступово зникають різновиди інформації з нульовою цінністю. Оброблені та необроблені (невпорядковані, неточні, неструктуровані) дані стають важливими факторами формування загальної продуктивності. Абсолютна цінність інформації не може бути визначена наперед і здатна перевищувати значущість мети її первинної обробки. В найближчій перспективі цінність даних може знайти своє відображення в фінансовому обліку як новий різновид активів та ресурсів.

Джерелами Великих Даних виступають будь-які явища, об'єкти або процеси, щодо яких або від яких можливо одержати певну інформацію, зокрема, прилади і сенсори промислових та інших об'єктів, технологічні лінії виробництва, інфраструктурні системи (торгівельні, енергетичні, транспортні тощо), соціальні мережі, мобільний зв'язок, будь-які побутові датчики контролю та управління, устаткування відеоспостереження та медичних обстежень, бази даних підприємств, установ та організацій, будь які різновиди інформації, в тому числі її уривки, координати GPS, запити до пошукових систем та відповіді на них, тривалість часу перегляду певного контенту, реакції на нього (уподобання, невподоба, схвалення, заперечення, коментарі) тощо. Обсяги та швидкість генерації і одержання інформації обумовлюють вимушеність зберігання її в первинному сирому (невпорядкованому, швидкому, неточному, неструктурованому) вигляді як кращу альтернативу втрачанням.

За сучаснийого розвитку технологій збирати та зберігати всі вхідні дані без виключень є більш простим і менш витратним рішенням, ніж фільтрувати та розподіляти їх за певними критеріями. Феномен Великих Даних є об'єктивним явищем, яке реально існує і набирає оберти, має здатність масштабно та у непередбачуваний

спосіб вплинути на більшість аспектів життя та діяльності людини, у зв'язку з чим право може зазнати змін фундаментального характеру, що матиме наслідком перегляд багатьох концепцій та підходів до розв'язання існуючих та майбутніх проблем.

Нове мистецтво роботи з інформацією, в тому числі з Великими даними або алгоритмами штучного інтелекту, полягає у влучному формулюванні запиту та визначенні способу його обробки, без чого значний обсяг інформації залишиться просто значним обсягом сирих (невпорядкованих, неточних, неструктурованих) даних, які не надають жодних відповідей. Виявлення нових аспектів цінності збережених даних є додатковим різновидом мистецтва обробки Великих Даних, що вимагатиме розвитку відповідних професій, зокрема, таких як фахівець з обробки даних (Data Scientist).

Кореляції допомагають аналізувати об'єкти через корисні статистичні закономірності, які можуть бути не пов'язані з причинністю. Зміна пріоритетів з пошуку відповіді на питання "чому" на з'ясування "що" є одним з принципів роботи з Великими даними: розуміння причини поступається прийняттю рішень на підставі кореляції, яка у свою чергу ґрунтується на використанні Великих Даних. Але в галузі права перехід від встановлення причинності до кореляцій на підставі Великих Даних може мати інший ефект, ніж в сфері маркетингу, фізики, природознавства або економічних прогнозів.

Великі Дані дозволяють побачити та використати нові зв'язки між тими окремими фрагментами інформації, які на перший погляд не демонструють жодної взаємопов'язаності між собою. Об'єднання різних даних з метою їх дослідження на предмет існування певних взаємозв'язків є ще одним зі способів управління Великими даними, підвищення якості самої інформації та ефективності її обробки. Але через дійсно великий обсяг інформації прерогатива прийняття рішень поступово зміщується від людини до алгоритму, який набирає ознак суб'єкта правових відносин.

Право на забуття (право вимагати видалення своїх персональних та інших даних із сфери загального доступу, пошукових систем), яке є протилежністю "постійній пам'яті", викликає дискусії з огляду на необхідність дотримання балансу з загальними соціальними інтересами (громадська або національна безпека, право громадськості бути проінформованою щодо особи публічного діяча тощо), забезпечення свободи слова, свободи доступу до інформації, запобігання негативному впливу та фальсифікації історії. У свою чергу, право на забуття входить в конфлікт з ідеологією Великих Даних, згідно з якою будь-яка інформація має зберігатися максимально довгі строки для того, щоб завдяки повторному використанню видобути її нову цінність.

Можливість спостереження за окремою особою сьогодні вбудована практично у всі аспекти і сфери життя. Технологічна природа програмного та апаратного забезпечення дозволяє обмінюватися даними з виробником, оператором та(або) іншими зацікавленими особами, не вимагаючи втручання або усвідомлення з боку їх власника або користувача.

Розуміння обсягу персональних даних доповнюється інформацією про особисті зв'язки, реакції (не-)уподобання, висловлені думки, розклад, пошукові запити, тривалість перегляду певного контенту тощо. В окремих випадках персональні дані виникають як побічний ефект функціонування комплексу складних компонентів кібернетичної системи, коли на початку збирання або обробки певна інформація може бути агрегованою в загальний масив мультимедійних даних суто технологічним способом і не мати прямого відношення до персональних даних, але згодом останні виникають внаслідок її аналізу та поєднання з іншою неперсоналізованою інформацією.

Використання збережених даних з повторною метою вимагає одержання нової інформованої згоди суб'єкта персональних даних, яка має містити достатній обсяг

відомостей і не залишати місця для двозначної інтерпретації. У певних випадках зв'язок з суб'єктом персональних даних може бути втрачений з різних причин, що здатне утворити суттєві перешкоди у використанні Великих Даних. У випадку одержання надмірно широкої за своїм обсягом згоди на використання персональних даних (фактично з будь-якої метою) нівелюється саме поняття інформованої згоди. У випадку згоди суб'єкта персональних даних на можливу зміну мети та повторне використання даних іншими особами, він має бути наділений правом так само змінити свою згоду повністю або частково, або відкликати її у будь-який момент за певних умов.

На ранніх етапах обробки окремої інформації питання про необхідність одержання згоди може не виникати у зв'язку з відсутністю впевненості в тому, що певні мультимедійні, мета- та інші дані мають відношення до конкретної особи. До моменту повної деанонімізації останньої дані можуть виглядати як такі, що не відносяться до категорії персональних. За допомогою кореляцій на підставі не персональних даних можуть бути виявлені такі зв'язки, які з високою точністю ідентифікують певну особу, але факт порушення захисту може бути усвідомлений лише після одержання результату.

Вилучення певної інформації з набору даних може дорівнювати її наявності. Диференціальна конфіденційність, як спосіб забезпечення приватності, або втратить свою ефективність, або за максимального ступеню розмиття даних позбавить Великі Дані їх сенсу, змісту та цінності.

Анонімізація не є достатньо ефективною у випадку використання Великих Даних, які дозволяють здійснити повторне встановлення особи. Анонімну особу можливо повністю ідентифікувати за змістом її пошукових запитів або шляхом порівняння з даними від інших джерел, узагальнена інформація відносно яких входить в обсяг Великих Даних.

З появою Великих Даних поступово втрачають свою ефективність такі стратегії забезпечення конфіденційності як інформована згода, відмова від участі та анонімізація. Якщо розглядати суб'єкта персональних даних в якості точки перетинання інформаційних потоків від будь-якого оточення, речей, приладів, явищ та контенту, а так само як сукупність соціальних зв'язків, то для одержання вичерпної інформації про нього стає достатнім проаналізувати та систематизувати всі доступні периферійні неперсональні дані, в тому числі за повну відсутність персоналізованої інформації.

Кримінальні правопорушення в сфері Великих Даних можуть бути пов'язані з несанкціонованим доступом до конфіденційної інформації, її викраденням, зловживанням, порушенням правил збору та обробки даних, а також з використанням аналітичних інструментів для незаконних цілей, наприклад, для дискримінації людей на основі особистих даних. Використання аналітичних інструментів для незаконних цілей може мати форми: зловживання даними (наприклад, збирання даних про користувачів без їхньої згоди або використання даних, які були одержані з іншою метою, в тому числі для вчинення таких незаконних дій, як розсилання спаму, фішингу або для здійснення кібератак); фальсифікація даних (зміна або підробка даних з будь-якою метою); розкриття конфіденційної інформації (наприклад, дешифрування паролів, виявлення номерів соціального страхування та банківських реквізитів тощо).

З метою запобігання кримінальним правопорушенням з використанням Великих Даних та аналітичних інструментів, необхідно постійно і на упередження вдосконалювати норми чинного законодавства, апаратні та програмні засоби захисту, розвивати та впроваджувати нові технології, зокрема, блокчейн, які здатні забезпечити більш високий рівень захисту від несанкціонованого доступу та зловживання, але не покладатися на жоден з існуючих варіантів в якості незмінного.

Великі Дані надають великі можливості, але долають рубежі захисту приватності та інших охоронюваних таємниць (службової, комерційної, лікарської, банківської, усиновлення тощо). Великі Дані та аналітичні інструменти можливо і необхідно використовувати для боротьби з неналежним використанням Великих Даних та аналітичних інструментів. Слід зважувати на наступні загрози та виклики в сфері використання Великих Даних: 1) порушення приватності, конфіденційності та деанонімізація; 2) несанкціонована повторна обробка інформації; 3) покладання на юридично значущу вірогідність (заміна причинності на кореляцію).

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. Проблематика змісту і значення для сфери права переходу від встановлення причинності до кореляцій на підставі Великих Даних буде розглянута у наступній публікації.

Використана література

1. Big Data / Oxford Learners' Dictionary. URL: <https://www.oxfordlearnersdictionaries.com/definition/english/big-data?q=Big+Data> (дата звернення: 24.02.2023).
2. Big Data / The dictionary by Merriam-Webster. URL: [https://www.merriam-webster.com/dictionary/big data](https://www.merriam-webster.com/dictionary/big%20data) (дата звернення: 24.02.2023).
3. Backstrom, Lars. Wherefore Art Thou R3579X? Anonymized Social Networks, Hidden Patterns, and Structural Steganography / Lars Backstrom, Cynthia Dwork, and Jon Kleinberg. *Communication of the Association of Computing Machinery*. December. 2011. P. 133.
4. Cadwalladr Carole, Graham-Harrison Emma. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian. The Cambridge Analytica Files*. Sat 17 Mar 2018. URL: <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election> (дата звернення: 24.02.2023).
5. Cappa, Francesco; Oriani, Raffaele; Peruffo, Enzo; McCarthy, Ian. Big Data for Creating and Capturing Value in the Digitalized Environment: Unpacking the Effects of Volume, Variety, and Veracity on Firm Performance. *Journal of Product Innovation Management*. 38 (1): 49-67. DOI:org/10.1111/jpim.12545. URL: <https://onlinelibrary.wiley.com/doi/epdf/10.1111/jpim.12545> (дата звернення: 24.02.2023).
6. Cukier, Kenneth. Data, data everywhere. *The Economist*. February 27, 2010. P. 1-14. URL: <https://www.economist.com/special-report/2010/02/27/data-data-everywhere> (дата звернення: 24.02.2023).
7. Damghani, B.M. The Non-Misleading Value of Inferred Correlation: An Introduction to the Cointelation Model. *Wilmott*, 2013: 50-61. DOI:org/10.1002/wilm.10252. URL: <https://onlinelibrary.wiley.com/doi/10.1002/wilm.10252> (дата звернення: 24.02.2023).
8. Dwork, Cynthia. A Firm Foundation for Private Data Analysis. *Communications of the Association of Computing Machinery* 54, 1. January 2011. Pp. 86-95. URL: <https://dl.acm.org/toc/cacm/2011/54/1> (дата звернення: 24.02.2023)
9. Düking P, Achtzehn S, Holmberg HC, Sperlich B. Integrated Framework of Load Monitoring by a Combination of Smartphone Applications, Wearables and Point-of-Care Testing Provides Feedback that Allows Individual Responsive Adjustments to Activities of Daily Living. *Sensors (Basel)*. 2018. May 19; 18(5):1632. DOI:org/10.3390/s18051632. URL: <https://pubmed.ncbi.nlm.nih.gov/29783763> (дата звернення: 24.02.2023).
10. Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja González. *Court of Justice of the European Union*. Press Release No 70/14. Luxembourg, 13 May 2014, Judgment in Case C-131/12. URL: <https://curia.europa.eu/jcms/upload/docs/application/pdf/2014-05/cp140070en.pdf> (дата звернення: 24.02.2023).

11. Graff, Amy. Why some houses on Google Street View are blurred out / SFGATE, July 15, 2019. URL: <https://www.sfgate.com/local/article/Google-Street-View-blurred-houses-14096605.php> (дата звернення: 24.02.2023).
12. Facebook IPO prospectus // Facebook. Form S-1 Registration Statement, US Securities And Exchange Commission. February 1, 2012. URL: <http://sec.gov/Archives/edgar/data/1326801/000119312512034517/d287954ds1.htm> (дата звернення: 24.02.2023).
13. Hilbert, M., Darmon, D. How Complexity and Uncertainty Grew with Algorithmic Trading / Entropy 2020, 22, 499. DOI:org/10.3390/e22050499. URL: <https://www.mdpi.com/1099-4300/22/5/499> (дата звернення: 24.02.2023).
14. IBM, Honda, and PG&E Enable Smarter Charging for Electric Vehicles // Press release. April 12, 2012. URL: <http://www-03.ibm.com/press/us/en/pressrelease/37398.wss> (дата звернення: 24.02.2023).
15. Inglehart, R., & Klingemann, H. D. (2000). Genes, culture, democracy, and happiness / MIT Press, Culture and subjective well-being, 165-183. URL: <https://psycnet.apa.org/record/2000-16279-006> (дата звернення: 25.02.2023).
16. Inrix: Unlock the Power of Big Data to Make an Impact, our data, insights and tools can help you shape the future of mobility. URL: <https://inrix.com> (дата звернення: 25.02.2023).
17. Mayes, Tessa. We have no right to be forgotten online. *The Guardian*. 18 Mar 2011. URL: <https://www.theguardian.com/commentisfree/libertycentral/2011/mar/18/forgotten-online-european-union-law-internet> (дата звернення: 25.02.2023).
18. Mitchell, David. The right to be forgotten will turn the internet into a work of fiction. *The Guardian*. 6 Jul 2014. URL: <https://www.theguardian.com/commentisfree/2014/jul/06/right-to-be-forgotten-internet-work-of-fiction-david-mitchell-eu-google> (дата звернення: 25.02.2023).
19. McNeil, Sonia K. Privacy And The Modern Grid. *Harvard Journal of Law & Technology*. 2011. Vol. 25, no. 1. URL: <http://jolt.law.harvard.edu/articles/pdf/v25/25HarvJLTech199.pdf> (дата звернення: 25.02.2023).
20. Narayanan, Arvind. Robust De-Anonymization of Large Sparse Datasets / Arvind Narayanan and Vitaly Shmatikov // Proceedings of the IEEE Symposium on Security and Privacy. 2008. P. 111. URL: http://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf (дата звернення: 25.02.2023).
21. Narayanan Arvind, Vitaly Shmatikov. How to Break the Anonymity of the Netflix Prize Dataset. Cornell University: Cryptography and Security, October 16, 2006. URL: <https://arxiv.org/abs/cs/0610105> (дата звернення: 25.02.2023).
22. Nikolić, D., Mureşan, R.C., Feng, W. and Singer, W. (2012), Scaled correlation analysis: a better way to compute a cross-correlogram. *European Journal of Neuroscience*, 35: 742-762. DOI:org/10.1111/j.1460-9568.2011.07987.x. URL: <https://onlinelibrary.wiley.com/doi/10.1111/j.1460-9568.2011.07987.x> (дата звернення: 25.02.2023).
23. Open Knowledge Foundation. URL: <https://okfn.org> (дата звернення: 25.02.2023).
24. Ohm, Paul. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization (August 13, 2009). *UCLA Law Review*, Vol. 57. P. 1701, 2010, U of Colorado Law Legal Studies Research Paper No. 9-12. URL: <https://ssrn.com/abstract=1450006> (дата звернення: 25.02.2023).
25. Luthy, Clay. Guest Perspective: IBM Working with PG&E to Maximize the EV Potential // PGE Currents magazine. April 13, 2012. URL: <http://www.pgecurrents.com/2012/04/13/ibm-working-with-pge-to-maximize-the-ev-potential> (дата звернення: 25.02.2023).
26. Pollack, Andrew. DNA Sequencing Caught in the Data Deluge. *New York Times*. November 30, 2011. URL: <http://www.nytimes.com/2011/12/01/business/dna-sequencing-caught-in-deluge-of-data.html?pagewanted=all> (дата звернення: 25.02.2023).
27. Powles, Julia. Floridi, Luciano. A manifesto for the future of the 'right to be forgotten' debate. *The Guardian*, 22 Jul 2014. URL: <https://www.theguardian.com/technology/2014/jul/22/a-manifesto-for-the-future-of-the-right-to-be-forgotten-debate> (дата звернення: 25.02.2023).
28. Renee DiResta, Kris Shaffer, Becky Ruppel, David Sullivan, Robert Matney, Ryan Fox, Jonathan Albright. Ben Jonson. The Tactics & Tropes of the Internet Research Agency / U.S. Senate

Documents, University of Nebraska Lincoln, New Knowledge, Oct. 2019. URL: <https://digitalcommons.unl.edu/senatedocs/2> (дата звернення: 25.02.2023).

29. Rosenthal, Jonathan. *The Economist: International banking: Renaissance*. Penguin UK, 2012. 64 p. P. 7-8. URL: <https://www.penguin.co.nz/books/the-economist-9780241964231> (дата звернення: 25.02.2023).

30. Singel, Ryan. Netflix Spilled Your Brokeback Mountain Secret, Lawsuit Claims. *Wired*. December 17, 2009. URL: <http://www.wired.com/threatlevel/2009/12/netflix-privacy-lawsuit> (дата звернення: 25.02.2023).

31. The Future of Jobs Report 2020. *World Economic Forum*. October 2020. URL: https://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf (дата звернення: 25.02.2023).

32. The Sunlight Foundation. URL: <https://sunlightfoundation.com> (дата звернення: 25.02.2023).

33. Vehicle data recorders. Watching your driving. *The Economist*. June 23, 2012. URL: <http://www.economist.com/node/21557309> (дата звернення: 25.02.2023).

34. Viktor Mayer-Schönberger, Kenneth Cukier. *Big Data: The Essential Guide to Work, Life and Learning in the Age of Insight*. Hachette UK, 2013. 256 p.p.

35. Брайчевський С.М. Персональні дані та мультимедіа. *Інформація і прав*. № 4(35)/2020. С. 82- 92.

36. Жила, Надія. Онлайн-карти можуть сховати будинок за запитом. Як це зробити? *Mezha Media*, 25 травня 2022 р. URL: <https://mezha.media/2022/05/24/onlayn-karty-skhovaty-budynok> (дата звернення: 25.02.2023).

37. Карташов М.В. Імовірність, процеси, статистика: посібник. Київ: Видавничо-поліграфічний центр “Київський університет”, 2008. 494 с.

38. Радутний О. Суб’єктність штучного інтелекту у кримінальному праві. *Право України*. 1/2018. С. 123-136.

39. 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.

40. 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.

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