

СОЦІАЛЬНІ НАСЛІДКИ BIG DATA: ДАНІФІКАЦІЯ СОЦІАЛЬНОГО ЖИТТЯ

Кислова Ольга Миколаївна – кандидат соціологічних наук, доцент кафедри методів соціологічних досліджень соціологічного факультету Харківського національного університету імені В. Н. Каразіна, майдан Свободи, 4, Харків, 61122, Україна, olga.kyslova@karazin.ua, ORCID ID <https://orcid.org/0000-0003-2241-1313>

Olga Kyslova – PhD in Sociological Science, Associate Professor, Department of Methods of Sociological Research V. N. Karazin Kharkiv National University, 4, Svobody Sq., Kharkiv, 61122, Ukraine, olga.kyslova@karazin.ua, ORCID ID <https://orcid.org/0000-0003-2241-1313>

Стаття присвячена аналізу соціальних наслідків стрімкого поширення великих даних, які в сукупності поступово призводять до формування нової фази еволюційного розвитку цивілізації – суспільства, заснованого на даних (datafied society). Показано, що головним чинником соціальних трансформацій є даніфікація (datafication), яка передбачає переведення в кількісну форму (в «дані») всіх явищ та аспектів оточуючого світу, навіть таких, що раніше мислилися винятково «якісними», кількісно не оцінюваними: людських дій, відносин, бажань, настроїв тощо. У цьому контексті змінюється інтерпретація «даних», вони перестають бути тим, що дано, що існує незалежно від чого б то не було. Сьогодні дані «виробляються», «видобуваються» з соціального життя. При цьому їхнє виробництво зумовлюється двома факторами: 1) зовнішньою інфраструктурою, за допомоги якої вони збираються, обробляються та зберігаються; 2) специфікою генерування їхньої вартості (цінності), що включає перш за все монетизацію, а крім того, засоби державного контролю, культурне виробництво, громадянські можливості тощо. В сучасних умовах все частіше говорять про розвиток нового напрямку інтелектуального аналізу даних, що отримав назву «майнінг життя» («life mining»), під яким розуміють вилучення шаблонів з комбінованих цифрових слідів, залишених людьми в інтернеті, і прогнозування на їхній основі реальних результатів: майбутніх дій і емоційних станів; відновлення відсутньої інформації, наприклад особистої інформації, яку людина не хотіла б оприлюднити. Наявність інструментів «майнінгу життя» та дедалі ширше їхнє застосування призводить до масового привласнення соціального життя тими, хто володіє великими даними, до формування феномену диктатури даних. У зв'язку з цим виникає нагальна потреба залучення соціологів до дискусій відносно соціальних наслідків великих даних, а також до прогнозування потенційних шляхів розвитку нашого суспільства, оскільки без цього неможливий опір тим, хто прагне встановити диктатуру даних та створити суспільство безпрецедентного нагляду та контролю.

Ключові слова: великі дані, даніфікація, майнінг життя, диктатура даних.

The article is devoted to the analysis of the social consequences of the rapid dissemination of big data that gradually leads to the formation of a new phase of the evolutionary development of civilization - a datafied society. It is shown that the main factor in social transformations is datafication, which involves the transition to a quantitative form (into "data") of all phenomena and aspects of the world around us, even those that were previously understood as purely "qualitative", not quantifiable: human actions, relationships, desires, moods, etc. In this context, the interpretation of "data" is changing, they cease to be what is given, what exists independently of anything. Today, data is "produced", "extracted" from social life. Moreover, their production is determined by two factors: 1) the external infrastructure with which they are collected, processed and stored; 2) the specifics of generating their value, including primarily monetization, and in addition, means of state control, cultural production, and civic opportunities. Today, more and more people are talking about the development of a new area of intelligent data analysis, called "life mining", which means extracting patterns from combined digital footprints left by people on the Internet and predicting real results based on them: future actions and emotional states; restoration of missing information, for example, personal information that a person would not want to make public. The existence of "life mining" tools and their wider application leads to the massive appropriation of social life by those who possess big data, to the formation of the phenomenon of data dictatorship. In this regard, there is an urgent need to involve sociologists in discussions concerning the social consequences of big data, as well as in predicting the possible ways for society to develop, since without this, resistance to those who want to establish a data dictatorship and create a society of unprecedented supervision and control is impossible.

Keywords: big data, datafication, life mining, data dictatorship

Сьогодні активно дискутуються питання, пов'язані з великими даними, а також тими революційними змінами в науці та суспільстві, які відбуваються внаслідок їхнього стрімкого поширення [5; 17]. Використання великих даних не тільки відкриває нові можливості (пізнання, навчання,

спілкування, вирішення побутових проблем), а й зумовлює появу нових проблем. Так, Віктор Майер-Шенбергер та Кеннет Кук'єр відзначають: «Великі дані мають велике практичне значення як технологія, яка слугує вирішенню актуальних повсякденних проблем, але водночас породжує ще більше нових» [5]. Ці проблеми можна умовно розділити на два напрямки: 1) методологічні, зумовлені появою нових підходів до отримання наукового знання¹, що перебудовує звичний порядок проведення наукових досліджень, зокрема соціологічних; 2) соціальні, які виникають у зв'язку з тим, що дані про соціальне життя монетизуються та перетворюються на ресурс, за який зараз йде жорстока боротьба. Останнім часом навіть запропоновано ввести новий показник розвитку економіки країн – валовий продукт даних, який ще називають «новий ВВП» (англ. new GDP – Gross Data Product) [10].

Мета даної статті – дослідити соціальні наслідки великих даних, які у сукупності поступово призводять до формування нової фази еволюційного розвитку цивілізації – суспільства, заснованого на даних (Datafied Society).

Великі дані (Big Data) – техно-соціальний феномен, розуміння якого зазнало певних трансформацій: від визнання надзвичайності їхніх об'ємів, які не дозволяють здійснювати обробку звичними методами та інструментами, до фокусування уваги на соціогуманітарних аспектах, на «соціальному боці». «Наразі великі дані розуміються як сукупність цифрових слідів, які залишають люди, коли використовують інформаційні технології: здійснюють інтернет-серфінг, завантажують мобільні додатки чи музику, спілкуються з друзями в соціальних мережах, користуються GPS, купують товари в інтернет-магазинах тощо. Великі дані генеруються інтернетом, проте вони містять інформацію не стільки про інтернет, скільки про суспільство та суспільні процеси, які відображаються в інтернеті. Іншими словами, великі дані – даніфікована інформація про все і про всіх» [3].

Розуміння великих даних нерозривно пов'язане з осягненням сутності *даніфікації*, яку інтерпретують як процес перетворення в кількісну форму (в «дані») всіх явищ та аспектів оточуючого світу, навіть таких, що раніше мислилися винятково «якісними», кількісно не оцінюваними. Термін «даніфікація» («datafication») був введений у науковий обіг В. Майер-Шенбергером та К. Кук'єром у 2013 році в контексті висвітлення новітніх тенденцій суспільних трансформацій, зумовлених поширенням великих даних. Цим терміном вони позначили «...процес подання явищ в кількісній формі, яка дозволяє їх аналізувати» [5].

Даніфікація проковує поглянути на дані під незвичним кутом зору. Відомо, що слово «дані» є множиною від лат. datum – те, що дано, існує незалежно від чого б то не було. Проте останнім часом тлумачення даних змінюється. Сьогодні дані «виробляються», «видобуваються» з соціального життя [11; 12; 18; 24; 25; 30; 31]. При цьому виробництво даних зумовлюється двома важливими факторами: 1) зовнішньою інфраструктурою, за допомогою якої вони збираються, обробляються та зберігаються; 2) специфікою генерування їхньої вартості (цінності), що включає перш за все монетизацію, а крім того, засоби державного контролю, культурне виробництво, громадянські можливості тощо.

Інфраструктура виробництва даних належить і контролюється переважно корпораціями та державами, які можуть вирішувати, як і які дані будуть накопичуватися та слугуватимуть основою для подальшого аналізу, результати якого, у свою чергу, згодом стануть основою та обґрунтуванням управлінських рішень. Алгоритмічне регулювання (інші назви – алгоритмічне управління, алгоритмічна влада; англ. algorithmic regulation, algorithmic governance) – новітня форма управління, в якому використання комп'ютерних алгоритмів застосовуються до нормативно-правових актів та правоохоронних органів [29].

Сьогодні IT-компанії та урядові установи активно здійснюють видобуток (майнінг) швидко зростаючих даних, зібраних через соціальні медіа та комунікаційні платформи, такі як Facebook, LinkedIn, Twitter, Skype, YouTube та безкоштовні сервіси електронної пошти, з метою відстеження інформації про поведінку людини: «Зараз можна збирати таку інформацію, яку раніше не могли відстежити, будь то стосунки, виявлені телефонними дзвінками або настройками, розкритими через твіти» [5]. Дані, зібрані інтернетом, тобто наші цифрові сліди, перетворилися на засіб доступу до приватної та навіть інтимної інформації: до наших відносин, переживань і настроїв.

Так, Facebook даніфікує та аналізує людські відносини, які представляються у вигляді соціального графа користувачів, який можна не лише візуалізувати, а й обчислити з використанням математичного апарату теорії графів. Зокрема, інструмент «Graph Search Facebook» (який ще називають «розумний пошук у Facebook») дає можливість буквально «проникнути в голови» користувачів цієї мережі, дізнатися про їхні інтереси та переваги, де вони живуть і чим заробляють на життя, а також багато іншого. І хоча офіційно розвиток Graph Search Facebook був зупинений у грудні 2014 року, проте всі його функції працюють і сьогодні, а головне – Facebook володіє найціннішими даними, застосування яких спроможне впливати на різноманітні соціальні процеси. Найвідомішим прикладом цього є скандал в 2017 році,

¹ Мається на увазі обчислювальна та цифрова епістемології (докладніше див. [17]).

спричинений витоком даних з Facebook та їхнім застосуванням британською компанією Cambridge Analytica з метою впливу на думку виборців шляхом точно розрахованої персоналізованої реклами. До цього скандалу більшість критичних зауважень на адресу компанії Facebook були спрямовані на те, які дані вона здатна отримати, проте сьогодні стало зрозумілим, що слід враховувати і те, що вона з ними робить і які наслідки це викликає.

Twitter спеціалізується на даніфікації настроїв, є ініціатором розвитку методів аналізу тональності текстів – методів сентимент-аналізу (англ. Sentiment analysis), які часто називають методами видобутку думок та уподобань (англ. Opinion mining). Всі ці методи є різновидом автоматизованого контент-аналізу, що використовує комп'ютерну лінгвістику з метою виявлення в текстах емоційно забарвленої лексики, яка дає можливість оцінити емоції та думки авторів по відношенню до об'єктів, про які йдеться в цих текстах. Застосування таких методів широко поширене – від бізнес-послуг до політичних кампаній. Twitter оперує даними понад 100 мільйонів людей, які щодня відправляють 250 мільйонів коротких твітів – коротких повідомлень відносно того, що вони робили, що вони роблять зараз, що збираються робити найближчим часом. Ці дані Twitter використовує не менш цікаво, ніж Facebook. Зокрема, прогнозує поведінку людей у найближчій перспективі, при цьому інтерес представляє не поточна чи минула діяльність людей, а їхні плани на майбутнє, які «видобуваються» у вигляді низки найбільш ймовірних видів діяльності [27]. Такі прогнозні дослідження (predictive analytics) можуть мати неоднозначні застосування. З одного боку, вони можуть сприяти запобіганню правопорушень та кримінальних злочинів, з іншого боку, вони здатні занурити нас у реальність, сконструйовану фантазією Стівена Спілберга у кінофільмі «Особлива думка» (2002 р.), де показано, що можна створити такий соціальний порядок, коли покарання настає до здійснення злочину (лише на основі побудованого прогнозу).

Отже, прогнозні дані можуть бути використані для покращення суспільства, але ними також можна зловживати. В цьому контексті дуже актуальним є попередження В. Майер-Шенберґера: «Найбільшим і найскладнішим завданням суспільства є захист вільної волі, самої основи суспільної свободи від тих, хто неправильно застосовує Big Data» [22].

Прогнозна аналітика, заснована на великих даних, має величезні можливості, які до повсюдної даніфікації здавалися просто фантастичними. Проте сьогодні вони стали реальністю. «Ми спостерігаємо появу нового типу видобутку (mining) даних, який, можливо, найкраще характеризується терміном «майнінг життя» («life mining»), під яким розуміють отримання корисних знань із комбінованих цифрових слідів, залишених людьми, які значну частину свого життя проводять в інтернеті» [27, р. 2].

«Майнінг життя» («life mining») здійснюється на основі даніфікованої інформації, його принципи закладено в ідеологічні припущення, вкорінені в пануючих нині соціальних нормах взаємодії, що здійснюється на основі сучасних інформаційних технологій. Користувачі надають особисту інформацію компаніям та отримують натомість послуги. Особисті дані в обмін на послуги зв'язку – певна форма бартеру. Проте пересічні користувачі не усвідомлюють повною мірою наслідки – підвищення ефективності соціального контролю та поширення маніпулятивного маркетингу. При цьому прийняття рішень, що ґрунтуються на великих даних (algorithmic regulation, algorithmic governance), часто незрозумілі звинувачення. Вимальовується ймовірність поширення нового виду дискримінації – дискримінації на основі передбачення проступку чи злочину, яких в дійсності не було [28]. Таким чином, можна відзначити дискримінаційний потенціал великих даних, які дозволяють нам створювати превентивні узагальнення людської поведінки в раніше неможливих масштабах.

Сьогодні наукове співтовариство все більше усвідомлює необхідність відповідального використання алгоритмів, керованих великими даними, важливість урахування законів та етичних норм, а головне – встановлення регуляторних рамок, які запобігатимуть ймовірним зловживанням «алгоритмічною владою». Прикладами таких регуляторних рамок є Загальний регламент Європейського Союзу про захист даних (GDPR), Нью-Йоркський закон про автоматизовані системи рішень (ADS) та принцип чистої нейтральності, який спрямований на захист прав осіб, що потерпають від втручання у власні «думки та поведінку» внаслідок «майнінгу життя», що здійснюється на основі безперервного збору та аналізу особистих даних [32]. «Ці регуляторні рамки є визначними прикладами світової тенденції: уряди починають визнавати необхідність регулювання алгоритмічної технології, керованої даними» [8].

Надання послуг у державному секторі, нагляд та управління поведінкою громадян є ключовим напрямком алгоритмічного управління. Політологи та правознавці досліджують автоматизовані процедури надання державних послуг та прийняття адміністративних рішень, які, за думкою їхніх розробників, повинні підвищувати ефективність прийняття рішень, спрощувати та поліпшувати роботу державних служб. Однак ці очікування часто не виправдовуються. Дослідження показали, що розгортання алгоритмічних систем у державному секторі спричинило багато непередбачуваних наслідків [15]. Крім того, застосування алгоритмічних інструментів в уряді часто покладається на нові форми спостереження за населенням, що здійснюється державними та корпоративними суб'єктами. Підставою для багатьох проєктів надання цифрових послуг та вибору політики на основі алгоритмів є системи оцінювання та

прогнозування поведінки, уподобань та думок громадян. Вони використовуються для розподілу соціальних пілг, для боротьби з ухиленням від сплати податків та шахрайствами, для інформування правоохоронних органів та запобігання тероризму, прикордонного контролю та управління міграцією. Проте дослідження демонструють, що практики оцінювання на основі алгоритмів перелаштовують владні відносини, причому частіше на шкоду бідним та тим, хто має низький рейтинг (*underscored*) [16].

Останнім часом майнінг життя (*life mining*) все частіше розглядається як нова парадигма розуміння соціальності та соціальної поведінки [25]. Це зумовлено тим, що сучасні люди все більше занурюються в цифрове середовище та все сильніше відчувають залежність від цифрових пристроїв. Вони практично невіддільні від своїх смартфонів, активно спілкуються з веб-сайтами соціальних мереж та додатками для обміну повідомленнями, здійснюють платежі за допомогою онлайн-платформ та кредитних карток. Все це означає, «... що поведінку людей, зв'язок, географічне положення та навіть фізіологічні стани можна легко записати, створюючи великі зразки цифрових слідів» [18]. Дослідження М. Косінського та інших [19] показує, як приватні риси та ознаки можна «витягнути» з цифрових слідів. Наприклад, у цьому дослідженні Facebook Likes використовувались для автоматизованого прогнозування цілої низки особистих ознак, зокрема: сексуальної орієнтації, етнічної приналежності, релігійних та політичних поглядів, інтелекту, почуття щастя, вживання наркотичних речовин, батьківської розлуки, віку тощо.

Пересічні користувачі інтернету вважають, що вони можуть приховати певні відомості про своє життя, такі як їхня сексуальна орієнтація чи вік, якщо вони це не розголошували. Але ця інформація може бути передбачена в статистичному сенсі з інших аспектів їхнього життя, які вони розкрили. Наприклад, велика роздрібна мережа США використовувала записи покупців, щоб передбачити вагітність своїх клієнтів-жінок та надсилати їм чіткі та цілеспрямовані пропозиції. У деяких контекстах несподіваний приплив ваучерів на внутрішньоутробні вітаміни та одяг для материнства може бути бажаним, але це також може призвести до трагічного результату, наприклад шляхом виявлення (або неправильного припущення) щодо вагітності незаміжної жінки в культурі, де це вважається неприпустимим [19]. Як показано на цьому прикладі, прогнозування особистої інформації для покращення маркетингових заходів може мати великі негативні наслідки.

А починалося все дуже невинно, з інтелектуалізації інформаційних технологій, тобто із впровадження елементів штучного інтелекту, зокрема певних методів інтелектуального аналізу даних, в наші «нові медіа» [4], з розвитку так званого інтелектуального інтернету (*intelligent internet*), специфіка якого визначається інтелектуалізацією вже існуючих сервісів, а також широким застосуванням інтелектуальних агентів та інтелектуального аналізу даних для більшого комфорту та задоволення потреб користувачів [20]. Проте англійське слово «*intelligent*» має два значення: інтелект та розвідка. Тобто «інтелектуальний інтернет» (*intelligent internet*), «інтелектуальні медіа» (*media intelligence*) є засобами розвідки, вивідування у пересічних користувачів даних про них самих, їхнє оточення та соціальні взаємодії, що є основою для застосування методів прогнозної аналітики з метою «майнінгу життя».

У середовищі розробників штучного інтелекту (ШІ) все частіше лунають попередження стосовно того, що не треба поспішати з його повсюдним впровадженням, бо існує безліч не повністю досягнутих наслідків, які можуть перекреслити всі ймовірні «блага» та стати загрозою існування людської цивілізації [8; 30]. Зокрема, деякі інтелектуальні системи можуть викликати етичні та юридичні питання, пов'язані з потенційно упередженим прийняттям рішень. Якщо дані перекошені та / або конструкція системи містить упередженість, рішення, рекомендовані такими системами, можуть бути дискримінаційними щодо певних категорій або груп. «Коли ШІ самостійно вчиться, як ми можемо це заблокувати? Фіксовані правила етики занадто жорсткі і легко виходять з ладу. Але якщо ми просто дозволимо машинам вивчати етику, спостерігаючи і наслідуючи нас, вони навчаться робити багато неетичних речей. Тож, можливо, ШІ змусить нас протистояти тому, що ми насправді маємо на увазі під етикою, перш ніж ми зможемо вирішити, яким чином ми бачимо етичність інтелектуальних систем» [30, р.44-45].

Загрози «майнінгу життя». Дослідники, які є прихильниками широкого впровадження даніфікації та застосування методів «майнінгу життя» з метою досягнення «соціального блага», виходять з припущення що цифрові сліди отримуються нейтральними посередниками – соціальними медіа та інтернет-платформами. Проте всі ці «посередники» в дійсності не є нейтральними. Платформи соціальних медіа вимірюють та одночасно монетизують поведінку людей в інтернеті. Алгоритми, що використовують Google, Facebook, Twitter та ін., по суті своїй селективні і маніпулятивні. Як користувачі, так і власники мають можливість варіювати налаштуваннями, обираючи власні «правила гри». В маркетингу прогнозування потреб майбутніх клієнтів схоже на маніпулювання бажанням: виявлення конкретних закономірностей у споживчих звичках зазвичай ведеться з метою штучного створення попиту - маркетингової стратегії, яка успішно монетизується за допомогою алгоритмів краудмаркетингу, який полягає у створенні необхідних відгуків (рекомендацій) нібито незалежних споживачів. При цьому інтернет-пошук та зміст соціальних медіа підлягає персоналізації та налаштуванню, що дозволяє

приспосовувати повідомлення до конкретних аудиторій або окремих людей з метою спровокувати певну поведінку – покупку товару чи голосування за певного кандидата. Отже, популяризація ідеї, що цифрові сліди людської поведінки є повністю об'єктивними даними, суперечить відомим практикам фільтрації даних та алгоритмічному маніпулюванню з комерційних чи інших причин.

Таким чином, існування алгоритмів «майнінгу життя» та даних, з яких ці алгоритми видобувають інформацію про нас, провокує масштабні трансформації: соціальні, економічні та правові. *Відбувається масове привласнення соціального життя шляхом вилучення закономірностей з великих даних.* Це дозволило Ніку Коулдрі та Улісесу Мехіасу говорити про нову форму колоніалізму – *колоніалізм даних* (data colonialism) [12]. Вони стверджують, що процеси, в яких задіяні великі дані, нерозривно пов'язані з розвитком капіталізму, що сьогоднішня кількісна оцінка соціального, відома як даніфікація, є першим кроком привласнення такого ресурсу, як життя людини. Хто отримує прибуток? Це, безумовно, великі компанії, які займаються передачею даних. І не тільки вони. Зараз виникає цілий «сектор соціальної кількісної оцінки» (social quantification sector) [11], до якого входять не тільки інтернет-компанії та виробники апаратних засобів, а й брокери даних, що займаються їхнім аналізом та продають третім особам.

Н. Коулдрі та У. Мехіас вважають, що дані колонізують людське життя. Ми платимо своїми грошима телекомунікаційним провайдерам за те, щоб постійно бути «на зв'язку». Але, крім того, ми віддаємо (часто не розуміючи цього) свої персональні дані та особисту інформацію, інформацію про себе та своє особисте життя компаніям, які обробляють її з великою вигодою для себе. Таким чином, наше соціальне життя стає прямим внеском у капіталістичне виробництво, а дані (зібрані та оброблені) є засобом для цього перетворення. Незважаючи на те, що змінилися режими, інтенсивність, масштаби та контексти відчуження, основна мета сучасного колоніалізму – колоніалізму даних – залишається незмінною, а саме: придбати «територію» та ресурси, які здатні принести прибуток.

Сьогодні організація капіталу та конфігурація влади різко змінюються через нову форму суспільних відносин. Колоніалізм даних виправдовує те, що він робить як прогрес наукових знань, персоналізований маркетинг або раціональне управління, подібно до того, як історичний колоніалізм претендував на цивілізаційну місію.

Колоніалізм даних прокладає шлях до нового етапу капіталізму, контури якого ми лише частково бачимо: капіталізація життя без обмежень. Не буде жодної частини людського життя, жодного шару досвіду, з яких не можна буде отримати вигоду (прибуток). «Цей процес капіталізації стане основою для дуже нерівного нового соціального устрою, нового соціального порядку, глибоко несумісного з людською свободою та самостійністю» [12].

Але опір все-таки можливий. І перший крок полягає в тому, щоб назвати своїми іменами те, що відбувається.

А відбувається формування феномену, який зараз називають по-різному: *диктатура даних* (data dictatorship) [5; 23], *цифрова диктатура* (digital dictatorship) [6; 7], *колоніалізм даних* (data colonialism) [11; 12], *наглядний капіталізм* (surveillance capitalism) [31; 32], *датаїзм* (dataism) [7; 9; 25].

Нові технології спокушають деякі уряди створити безпрецедентні режими стеження. Вони прагнуть спостерігати за всіма і контролюватимуть усіх увесь час. У результаті може бути створена система, яка розуміє нас (людей) краще, ніж ми розуміємо самі себе. Ця система зможе передбачати не тільки наші уподобання та дії, а навіть відчуття, що призведе до повного контролю. «Ми говоримо про алгоритми машинного навчання, які можуть зламати розум людей. Вони випускають повідомлення, бачать реакції людей, вдосконалюють повідомлення, а потім переробляють його ще раз, поки алгоритм не зможе досягти будь-якої місії, яку він має» [23].

Проте остаточний курс поки не визначений. Дискусія тільки починається (див, напр., [1; 13; 21; 24]), і до неї долучаються фахівці у сфері штучного інтелекту та великих даних, які мають кардинально протилежні точки зору: від застережливого ставлення до своїх творінь [30] до повного їхнього виправдування [4] на кшталт такого: «У сьогоднішньому світі, де розповсюджуються різномасштабні конфлікти та війни нового типу, питання вивчення поведінки, що призводить до насильства та тероризму, є критично важливим. І забалакати це питання істериками про «тотальний цифровий контроль» і застарілою демагогією про «безпеку персональних даних» сьогодні є просто небезпечним – і для суспільства і для держави» [4].

Зараз ми спостерігаємо виникнення дедалі більше складних режимів цифрового спостереження у всьому світі. Мова йде не тільки про авторитарні режими, а й про демократичні уряди. «Тепер важлива не та хмара, на якій сидить Бог. Важливі тільки хмари IBM і Microsoft, де зберігаються дані. Це нова рушійна сила еволюції. Тому володіти даними так важливо. Якщо це не регулювати, невелика кількість людей, які володіють більшістю даних, буде контролювати майбутнє людського суспільства та навіть може спровокувати виникнення нових форм життя» [7].

В якості висновку зазначимо, що ми вважаємо нагальною потребою залучення соціологів до дискусій відносно соціальних наслідків великих даних, а також до прогнозування потенційних шляхів розвитку нашого суспільства, оскільки без цього неможливий опір тим, хто прагне встановити диктатуру даних та заволодіти світом; ми дійсно можемо опинитися в суспільстві безпрецедентного нагляду та контролю.

Література:

1. Денисюк С. Г., Косюк Б. О. Технологія Big Data: переваги та ризики. *Матеріали XLVIII науково-технічної конференції підрозділів ВНТУ, Вінниця, 13-15 березня 2019 р.* URL: <https://conferences.vntu.edu.ua/index.php/all-hum/all-hum-2019/paper/view/6730>.
2. Кислова О. М., Бердник К. О. Нові медіа як комунікативні технології XXI століття: наслідки мережевізації та інтелектуалізації комунікацій. *Соціальні технології: заради чого? Яким чином? З яким результатом?* Одеса: Одеський національний університет імені І. І. Мечнікова, 2015. С. 277-288.
3. Кислова О. М. Великі дані в контексті дослідження проблем сучасного суспільства. *Вісник Харківського національного університету імені В. Н. Каразіна. Серія «Соціологічні дослідження сучасного суспільства: методологія, теорія, методи»*. Випуск 42, 2019. С. 59-68. URL: <https://periodicals.karazin.ua/ssms/article/view/14869>.
4. Костюченко Ю. Не обмежувати, а змінюватись: виклики еволюції у світі загальнодоступних Big Data. *InformNapalm*. 2018. URL: <https://informnapalm.org/ua/vyklyky-evolyutsiyi-u-sviti-zagalnodostupnyh-big-data/>.
5. Майер-Шенбергер В., Кукьєр К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / пер. с англ. Инны Гайдюк. М.: Манн, Иванов и Фербер, 2014. URL: <https://rutlib5.com/book/7832/p/6>.
6. Некрасов В., Корнієнко Я. Як запобігти появі класу «непотрібних людей» і цифровій диктатурі. *Економічна правда*. 16 вересня 2019. URL: <https://www.epravda.com.ua/publications/2019/09/16/651670/>.
7. Харари Ю. Н. Homo Deus. Краткая история будущего / переклад з англ. А. Андреев. М.: Синбад, 2018. URL: <https://reader.bookmate.com/vKF6v9bb>.
8. Abiteboul S., Stoyanovich J. Transparency, Fairness, Data Protection, Neutrality: Data Management Challenges in the Face of New Regulation. *Journal of Data and Information Quality (JDIQ)*. Vol. 11(3), 2019. P 15. URL: <https://hal.inria.fr/hal-02066516/file/1903.03683.pdf>.
9. Brooks D. The philosophy of data. *New York Times*. 02.05.2013. URL: <http://www.nytimes.com/2013/02/05/opinion/brooks-the-philosophy-of-data.html>.
10. Chakravorti B., Bhalla A., Chaturvedi R. S. Which Countries Are Leading the Data Economy? *Harvard Business Review*. 2019. URL: <https://hbr.org/2019/01/which-countries-are-leading-the-data-economy>
11. Couldry, N., Mejias, U. A. Datafication. *Internet Policy Review*. Vol. 8 (4), 2019. URL: <https://policyreview.info/concepts/datafication>
12. Couldry N., Mejias U. A. Making data colonialism liveable: how might data's social order be regulated? *Internet Policy Review*. Vol. 8 (2), 2019. URL: <https://policyreview.info/articles/analysis/making-data-colonialism-liveable-how-might-datas-social-order-be-regulated>.
13. Cuquet M., Vega-Gorgojo G., Lammerant H., Finn R. *Societal impacts of big data: challenges and opportunities in Europe*. 2017. URL: <https://arxiv.org/ftp/arxiv/papers/1704/1704.03361.pdf>.
14. Dencik L. Understanding the Datafied Society by Decentring Data. *iCS Symposium*. 2018. URL: <https://snurb.info/node/2424>.
15. Hintz A., Dencik L., Wahl-Jorgensen K. (). *Digital citizenship in a datafied society*. Medford: Polity Press, 2018.
16. Katzenbach C., Ulbricht L. Algorithmic governance. *Internet Policy Review*. Vol. 8 (4), 2019. URL: <https://policyreview.info/concepts/algorithmic-governance>
17. Kitchin R. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. Thousand Oaks, CA: SAGE Publications, 2014.
18. Kosinski M., Stillwell D., Graepel T. Private traits and attributes are predictable from digital records of human behavior. *PNAS*. 2013. URL: <http://www.pnas.org/content/early/2013/03/06/1218772110.full.pdf>.
19. Kosinski M., Wang Y., Lakkaraju H., Leskovec J. Mining big data to extract patterns and predict real-life outcomes. *Psychological methods*. Vol. 21(4), 2016. P. 493–506. URL: <https://psycnet.apa.org/fulltext/2016-57141-003.html>.
20. Kyslova, O. N., Berdnyk, E. A. New media as a formation factor for digital sociology: the consequences of the networking in the society and the intellectualization of the communications. *Study of Changing Societies: Study of Changing Societies: Comparative and Interdisciplinary Focus*. Vol. 3 (7), 2013. P. 67-106. URL: <https://www.degruyter.com/downloadpdf/j/scs.2013.2013.issue-3/scs-2014-0154/scs-2014-0154.pdf>.
21. Lyon D. Surveillance, Snowden, and big data: Capacities, consequences, critique. *Big Data & Society*. Vol. 1(2), 2014. URL: <https://journals.sagepub.com/doi/full/10.1177/2053951714541861>.
22. Mayer-Schönberger V. Big data's bright and dark sides. *Skoll World Forum*. 2013. URL: <http://archive.skoll.org/2013/03/19/big-datas-bright-and-dark-sides/?series=how-can-big-data-have-a-social-impact>.
23. Moya B. *Today politics is about the struggle of controlling the flows of data, and that might give birth to Data Dictatorships*. 2019. URL: <https://medium.com/privateid/modern-dictatorships-and-the-fall-of-liberal-democracy-a-global-system-ruled-by-data-oligarchs-fc0988739588>.
24. Schäfer M. T., Van Es K. F. The datafied society: Studying culture through data. Amsterdam: Amsterdam University Press, 2017.
25. van Dijck J. Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology. *Surveillance & Society*, Vol. 12 (2), 2014. P. 197–208. URL: https://www.researchgate.net/publication/283985880_Datafication_Dataism_and_Dataveillance_Big_Data_Between_Scientific_Paradigm_and_Ideology

26. van Dijck J. (2017). In data we trust? The implications of datafication for social monitoring. *MATRIZES*. Vol. 11 (1). P. 39-59. URL: https://www.researchgate.net/publication/317255516_In_data_we_trust_The_implications_of_datafication_for_social_monitoring.
27. Weerkamp W., De Rijke M. Activity prediction: A twitter-based exploration. In *SIGIR workshop on time-aware information access*. 2012. URL: https://pure.uva.nl/ws/files/2475778/163223_tai2012_activities.pdf.
28. Wigan M. R., Clarke R. Big data's big unintended consequences. *Computer*. Vol. 46(6), 2013. P. 46-53 URL: <http://www.rogerclarke.com/DV/BigData-1303.html>.
29. Yeung K. Algorithmic regulation: A critical interrogation. *Regulation & Governance*. Vol. 12 (4), 2018. P. 505-523. URL: https://www.researchgate.net/profile/Karen_Yeung/publication/318820255_Algorithmic_regulation_A_critical_interrogation/inks/5a7d99cd0f7e9b9da8d8e1e2/Algorithmic-regulation-A-critical-interrogation.pdf.
30. Zicari R. V. Big Data and Artificial Intelligence: Ethical and Societal Implications. *Whither Artificial Intelligence? Debating the Policy Challenges of the Upcoming Transformation*. Science Policy Paper No. 3. Frankfurt am Main: Mercator Science-Policy Fellowship-Programme. 2018. P. 45-51.
31. Zuboff, S. Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, Vol. 30(1), 2015. P. 75-89. URL: <https://cryptome.org/2015/07/big-other.pdf>.
32. Zuboff, S. The age of surveillance capitalism: The fight for a human future at the new frontier of power. Profile Books, 2019. URL: https://books.google.com.ua/books?hl=uk&lr=&id=W7ZEDgAAQBAJ&oi=fnd&pg=PT14&dq=Shoshana+Zuboff&ots=dor5EXRAu_&sig=_CJOLFKZMwZCFAYyz7bXS1CeCwc&redir_esc=y#v=onepage&q=Shoshana%20Zuboff&f=false.

Социальные последствия big data: данификация социальной жизни

Статья посвящена анализу социальных последствий стремительного распространения больших данных, которые в совокупности постепенно приводят к формированию новой фазы эволюционного развития цивилизации – общества, основанного на данных (datafied society). Показано, что главным фактором социальных трансформаций является данификация (datafication), которая предполагает перевод в количественную форму (в «данные») всех явлений и аспектов окружающего мира, даже таких, которые ранее мыслились исключительно «качественными», количественно не оцениваемыми: человеческих действий, отношений, желаний, настроений и т.п.. В этом контексте меняется интерпретация «данных», они перестают быть тем, что дано, что существует независимо от чего бы то ни было. Сегодня данные «производятся», «добываются» из социальной жизни. При этом их производство обусловлено двумя факторами: 1) внешней инфраструктурой, с помощью которой они собираются, обрабатываются и хранятся; 2) спецификой генерирования их стоимости (ценности), включающий прежде всего монетизацию, а кроме того, средства государственного контроля, культурное производство, гражданские возможности. В современных условиях все чаще говорят о развитии нового направления интеллектуального анализа данных, получившего название «майнинг жизни» («life mining»), под которым понимают извлечение шаблонов из комбинированных цифровых следов, оставленных людьми в интернете, и прогнозирование на их основе реальных результатов: будущих действий и эмоциональных состояний; восстановление недостающей информации, например личной информации, которую человек не хотел бы сделать достоянием общественности. Наличие инструментов «майнинга жизни» и все более широкое их применение приводит к массовому присвоению социальной жизни теми, кто обладает большими данными, к формированию феномена диктатуры данных. В связи с этим возникает насущная необходимость привлечения социологов к дискуссиям относительно социальных последствий больших данных, а также к прогнозированию возможных путей развития общества, поскольку без этого невозможно сопротивление тем, кто хочет установить диктатуру данных и создать общество беспрецедентного надзора и контроля.

Ключевые слова: большие данные, данификация, майнинг жизни, диктатура данных.

References

1. Denisyuk, S. G., & Kosyuk, B. A. (2019) 'Big Data technology: advantages and risks'. *Proceedings of the XLVIII Scientific and Technical Conference of the Universities of VNTU*, Vinnytsia, March 13-15, 2019 [online]. Available at: <https://conferences.vntu.edu.ua/index.php/all-hum/all-hum-2019/paper/view/6730> [in Ukrainian].
2. Kyslova, O. M., & Berdnik, K. A. (2015) 'New Media as Communication Technologies of the 21st Century: Implications of Networking and Intellectualization of Communications'. *Social technologies: why? How? What is the result?* [online]. Available at: <http://dspace.univer.kharkov.ua/handle/123456789/11351> [in Ukrainian].
3. Kyslova O. 'Big Data in the Context of Studying Problems of Modern Society'. *Visnyk of V. N. Karazin Kharkiv National University. Series "Sociological studies of contemporary society: methodology, theory, methods"*. Vol. 42. P. 59-68 [online]. Available at: <https://periodicals.karazin.ua/ssms/article/view/14869> [in Ukrainian].
4. Kostyuchenko, Y. (2018) 'Not to limit, but to change: the challenges of evolution in the Big Data public world'. *InformNapalm [online]*. Available at: <https://informnapalm.org/ua/vyklyky-evolyutsiyi-u-sviti-zagalnodostupnyh-big-data/> [in Ukrainian].
5. Mayer-Schönberger V., & Cukier K. (2014) *Big data: A revolution that will transform how we live, work, and think*. Trans. from English Inna Gaydyuk. Moscow: Mann, Ivanov and Ferber [online]. Available at: <https://rutlib5.com/book/7832/p/6> [in Russian].
6. Nekrasov, V., Komienko, Y. (September 16, 2019) 'How to prevent the emergence of a "needless people" class and digital dictatorship' *Economic truth* [online]. Available at: <https://www.epravda.com.ua/publications/2019/09/16/651670/> [in Ukrainian].

7. Harari, Y. N. (2018). *Homo Deus: A brief history of tomorrow*. Trans. from English A. Andreev [online]. Available at: <https://reader.bookmate.com/vKF6v9bb> [in Russian].
8. Abiteboul, S., & Stoyanovich, J. (2019) 'Transparency, Fairness, Data Protection, Neutrality: Data Management Challenges in the Face of New Regulation'. *Journal of Data and Information Quality (JDIQ)*. Vol. 11(3) [online]. Available at: <https://hal.inria.fr/hal-02066516/file/1903.03683.pdf>.
9. Brooks, D. (2013) 'The philosophy of data. *New York Times*' [online]. Available at: <http://www.nytimes.com/2013/02/05/opinion/brooks-the-philosophy-of-data.html>.
10. Chakravorti, B., Bhalla, A., & Chaturvedi, R. S. (2019) 'Which Countries Are Leading the Data Economy?' *Harvard Business Review* [online]. Available at: <https://hbr.org/2019/01/which-countries-are-leading-the-data-economy>
11. Couldry, N., & Mejias, U. A. (2019) 'Datafication'. *Internet Policy Review*. Vol. 8 (4) [online]. Available at: <https://policyreview.info/concepts/datafication>
12. Couldry, N., & Mejias, U. A. (2019) 'Making data colonialism liveable: how might data's social order be regulated?' *Internet Policy Review*. Vol. 8 (2) [online]. Available at: <https://policyreview.info/articles/analysis/making-data-colonialism-liveable-how-might-datas-social-order-be-regulated>.
13. Cuquet, M., Vega-Gorgojo, G., Lammerant, H., & Finn, R. (2017) *Societal impacts of big data: challenges and opportunities in Europe* [online]. Available at: <https://arxiv.org/ftp/arxiv/papers/1704/1704.03361.pdf>.
14. Dencik, L. (2018). 'Understanding the Datafied Society by Decentering Data'. *iCS Symposium* [online]. Available at: <https://snurb.info/node/2424>.
15. Hintz, A., Dencik, L., & Wahl-Jorgensen, K. (2018). *Digital citizenship in a datafied society*. Medford: Polity Press.
16. Katzenbach, C., & Ulbricht, L. (2019) 'Algorithmic governance'. *Internet Policy Review*. Vol. 8 (4) [online]. Available at: <https://policyreview.info/concepts/algorithmic-governance>
17. Kitchin, R. (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. Thousand Oaks, CA: SAGE Publications.
18. Kosinski, M., Stillwell, D., & Graepel, T. (2013) 'Private traits and attributes are predictable from digital records of human behavior'. *PNAS* [online]. Available at: <http://www.pnas.org/content/early/2013/03/06/1218772110.full.pdf>.
19. Kosinski, M., Wang, Y., Lakkaraju, H., & Leskovec, J. (2016). 'Mining big data to extract patterns and predict real-life outcomes'. *Psychological methods*. Vol. 21(4). P. 493–506 [online]. Available at: <https://psycnet.apa.org/fulltext/2016-57141-003.html>.
20. Kyslova, O., & Berdnyk, E. (2013) 'New media as a formation factor for digital sociology: the consequences of the networking in the society and the intellectualization of the communications'. *Study of Changing Societies: Study of Changing Societies: Comparative and Interdisciplinary Focus*. Vol. 3 (7). P. 67-106 [online]. Available at: <https://www.degruyter.com/downloadpdf/j/scs.2013.2013.issue-3/scs-2014-0154/scs-2014-0154.pdf>.
21. Lyon, D. (2014) 'Surveillance, Snowden, and big data: Capacities, consequences, critique'. *Big Data & Society*. Vol. 1(2) [online]. Available at: <https://journals.sagepub.com/doi/full/10.1177/2053951714541861>.
22. Mayer-Schönberger, V. (2013) 'Big data's bright and dark sides' *Skoll World Forum* [online]. Available at: <http://archive.skoll.org/2013/03/19/big-datas-bright-and-dark-sides/?series=how-can-big-data-have-a-social-impact>
23. Moya, B. (2019) *Today politics is about the struggle of controlling the flows of data, and that might give birth to Data Dictatorships* [online]. Available at: <https://medium.com/privateid/modern-dictatorships-and-the-fall-of-liberal-democracy-a-global-system-ruled-by-data-oligarchs-fc0988739588>.
24. Schäfer, M. T., Van Es, K. F. (2017). *The datafied society: Studying culture through data*. Amsterdam: Amsterdam University Press.
25. van Dijck, J. (2014). 'Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology'. *Surveillance & Society*, Vol. 12 (2). P. 197–208 [online]. Available at: https://www.researchgate.net/publication/283985880_Datafication_Dataism_and_Dataveillance_Big_Data_Between_Scientific_Paradigm_and_Ideology
26. van Dijck, J. (2017). 'In data we trust? The implications of datafication for social monitoring'. *MATRIZES*. Vol. 11(1). P. 39-59 [online]. Available at: https://www.researchgate.net/publication/317255516_In_data_we_trust_The_implications_of_datafication_for_social_monitoring.
27. Weerkamp, W., & De Rijke, M. (2012) 'Activity prediction: A twitter-based exploration'. In *SIGIR workshop on time-aware information access* [online]. Available at: https://pure.uva.nl/ws/files/2475778/163223_tai2012_activities.pdf.
28. Wigan, M. R., & Clarke, R. (2013). 'Big data's big unintended consequences'. *Computer*. Vol. 46 (6). P. 46-53 [online]. Available at: <http://www.rogerclarke.com/DV/BigData-1303.html>.
29. Yeung, K. (2018). 'Algorithmic regulation: A critical interrogation'. *Regulation & Governance*. Vol. 12 (4). P. 505-523 [online]. Available at: https://www.researchgate.net/profile/Karen_Yeung/publication/318820255_Algorithmic_regulation_A_critical_interrogation/links/5a7d99cd0f7e9b9da8d8e1e2/Algorithmic-regulation-A-critical-interrogation.pdf.
30. Zicari, R. V. (2018) 'Big Data and Artificial Intelligence: Ethical and Societal Implications'. *Whither Artificial Intelligence? Debating the Policy Challenges of the Upcoming Transformation*. Science Policy Paper No. 3. Frankfurt am Main: Mercator Science-Policy Fellowship-Programme. P. 45-51.
31. Zuboff, S. (2015). 'Big other: surveillance capitalism and the prospects of an information civilization'. *Journal of Information Technology*, Vol. 30 (1). P. 75–89 [online]. Available at: <https://cryptome.org/2015/07/big-other.pdf>.
32. Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books [online]. Available at: https://books.google.com.ua/books?hl=uk&lr=&id=W7ZEDgAAQBAJ&oi=fnd&pg=PT14&dq=Shoshana+Zuboff&ots=dor5EXRAu_&sig=_CJOLFKZMwZCFAYyz7bXS1CeCwc&redir_esc=y#v=onepage&q=Shoshana%20Zuboff&f=false.