УДК 311.2:004.62:330.4
JEL Classification: C80, C82, C55, O33
DOI: 10.31767/nasoa.2-2026.03

О. В. СТАВИЦЬКИЙ,
кандидат економічних наук, доцент,
завідувач кафедри інформаційних технологій,
Національна академія статистики, обліку та аудиту,
e-mail: OVStavitskiy@nasoa.edu.ua
ORCID: 0000-0002-2114-0892
Р. Л. ПАНТЄЄВ,
кандидат технічних наук,
викладач кафедри інформаційних технологій,
Національна академія статистики, обліку та аудиту,
e-mail: Pantieiev@nasoa.edu.ua
ORCID: 0000-0003-4707-4608

Використання основних джерел великих даних для формування економічної статистики

У статті розкрито досвід національних статистичних управлінь (НСУ) країн Азійсько-Тихоокеанського регіону щодо використання великих даних у виробництві економічної статистики, а також перспективи і проблеми інтеграції таких джерел інформації у систему офіційної статистики. Метою дослідження є обґрунтування доцільності інтеграції великих даних у систему економічної статистики та визначення шляхів підвищення її якості, актуальності й аналітичної цінності в сучасних умовах глобальної цифровізації.

Показано застосування великих даних для розрахунку ключових макроекономічних показників, зокрема індексів споживчих цін, індексів цін на житлову нерухомість, статистики туризму та ін. Доведено, що найефективнішими джерелами великих даних є онлайн-дані про ціни, скановані дані роздрібних мереж, інформація з мобільних телефонів, адміністративні податкові бази, фінансові транзакції та інші цифрові сліди економічної діяльності. Розкрито переваги використання великих даних (підвищення оперативності отримання статистичної інформації, можливість формування більш деталізованих показників, зниження витрат на проведення обстежень, розширення аналітичних можливостей державної статистики), а також проблеми, пов'язані з їх впровадженням (доступ до інформації, забезпечення конфіденційності персональних даних, узгодження нових джерел із традиційними методологіями та необхідність розвитку технічної інфраструктури).

Стаття опублікована на умовах відкритого доступу за ліцензією CC BY-SA 4.0
(https://creativecommons.org/licenses/by-sa/4.0/)

Дослідження показало, що НСУ Азійсько-Тихоокеанського регіону реалізують численні проєкти, пов'язані з використанням великих даних у сфері економічної статистики, але більшість із них залишається на експериментальному етапі, за винятком статистики цін, де кілька НСУ вже інтегрували дані сканування та/або онлайн-дані у розрахунок індексу споживчих цін для певних товарних категорій. У деяких випадках ці нетрадиційні джерела даних можуть повністю або частково замінити традиційні методи збирання інформації. Але великі дані найчастіше використовують додатково до традиційних джерел, забезпечуючи більш оперативну, детальну інформацію та нові аналітичні можливості.

Ключові слова: *економічна статистика, цифрова економіка, статистичний аналіз, доступність даних, великі дані, індекс споживчих цін, індекс вартості житлової нерухомості.*

O. V. STAVYTSKYI,
PhD in Economics, Associate Professor,
Head of the Department of Information Technologies,
National Academy of Statistics, Accounting and Audit,
e-mail: OVStavitskiy@nasoa.edu.ua
ORCID: 0000-0002-2114-0892
R. L. PANTYEYEV,
PhD in Engineering,
Lecturer of the Department of Information Technologies,
National Academy of Statistics, Accounting and Audit,
e-mail: Pantieiev@nasoa.edu.ua
ORCID: 0000-0003-4707-4608

Using the Main Sources of Big Data to Generate Economic Statistics

The article examines the experience of National Statistical Offices (NSOs) of the Asia-Pacific region in the use of big data for the production of economic statistics, as well as the prospects and challenges of integrating such information sources into the system of official statistics. The purpose of the study is to substantiate the feasibility of integrating big data into the system of economic statistics, and to identify ways to improve its quality, relevance, and analytical value in the context of modern global digitalization.

The study demonstrates the application of big data in estimating key macroeconomic indicators, including consumer price indices, residential property price indices, tourism statistics, and others. It is proven that the most effective sources of big data are online price data, retail scanner data, mobile phone information, administrative tax databases, financial transactions, and other digital traces of economic activity.

The advantages of using big data are highlighted, including increased timeliness of statistical information, the possibility of generating more detailed

indicators, reduced costs of conducting surveys, and expanded analytical capabilities of official statistics. At the same time, the article identifies challenges associated with their implementation, such as access to information, ensuring the confidentiality of personal data, harmonizing new data sources with traditional methodologies, and the need to develop technical infrastructure.

The study shows that the NSOs of the Asia-Pacific region are implementing numerous projects related to the use of big data in the field of economic statistics, but most of them remain at the experimental stage. An exception is price statistics, where several NSOs have already integrated scanner data and/or online data into the calculation of the consumer price index for certain product categories. In some cases, these non-traditional data sources can fully or partially replace traditional methods of data collection. However, big data are most often used in addition to traditional sources, providing more timely and detailed information as well as new analytical opportunities.

Keywords: *economic statistics, digital economy, statistical analysis, data accessibility, big data, consumer price index, residential property price index.*

Постановка проблеми. Економічна статистика відображає поточний стан розвитку національної економіки. Традиційно основний обсяг економічної інформації формується на основі вибіркових обстежень населення, домогосподарств і суб'єктів господарювання. Окремі масиви даних надходять також із адміністративних джерел, зокрема з комерційних реєстрів, податкової звітності або митних баз. Водночас розвиток цифрових технологій і зростання масштабів електронної економічної діяльності спричиняють появу значних обсягів нових даних [1, 2].

Такі інформаційні ресурси можуть ефективно доповнювати традиційні статистичні підходи, забезпечуючи вищий рівень деталізації, оперативності та відкриваючи нові аналітичні можливості. Серед прикладів подібних джерел – скановані торговельні дані та онлайн-інформація про ціни, що використовуються для розрахунку індексів споживчих цін; дані мобільного позиціювання, які застосовуються в аналізі туристичних потоків; результати дистанційного зондування Землі для оцінювання соціально-економічних умов і рівня бідності; показники інтелектуальних лічильників споживання електроенергії; масиви фінансових транзакцій, зокрема інформація про платежі в торговельних закладах для дослідження споживчої поведінки та економічних тенденцій; а також онлайн-оголошення про вакансії, що дають змогу оцінювати стан ринку праці.

Використання великих даних визначено одним із пріоритетних напрямів розвитку сучасної статистичної діяльності. Національні статистичні служби країн Азійсько-Тихоокеанського регіону активно впроваджують і тестують різні інноваційні джерела інформації з метою вдосконалення офіційної статистики, зокрема економічної. Поширення пандемії COVID-19 стало додатковим стимулом для інтенсивного вивчення потенціалу великих даних, що сприяло налагодженню нових форм співпраці між державними структурами та приватним сектором і прискорило процес створення актуальних статистичних показників [3].

Попри те, що для багатьох національних статистичних установ застосування великих даних у сфері економічної статистики все ще має експериментальний характер, окремі їх види вже повністю інтегровані у процес формування статистичних продуктів у низці країн.

У статті подано приклади практичного застосування різних джерел великих даних для формування показників економічної статистики в межах національних статистичних систем, насамперед національних статистичних служб країн Азійсько-Тихоокеанського регіону.

Розглянуті приклади охоплюють розрахунок цінових індикаторів, зокрема індексу споживчих цін та індексу вартості житлової нерухомості, з використанням онлайн-інформації про ціни та сканованих торговельних даних. Також наведено приклади підготовки статистики туристичних потоків на основі даних операторів мобільного зв'язку, оцінювання рівня бідності із застосуванням результатів дистанційного спостереження за поверхнею Землі, формування експериментальних економічних показників у період пандемії COVID-19 шляхом залучення різних масивів великих даних [3, 4].

Після висвітлення міжнародного досвіду в статті узагальнено ключові проблеми, з якими стикаються національні статистичні органи під час отримання доступу до нових інформаційних ресурсів або їх практичного використання. Окрему увагу приділено можливим шляхам подолання цих проблем і рекомендаціям щодо підвищення ефективності впровадження інноваційних джерел даних у статистичну діяльність (рис. 1).

Рис. 1. Загальна схема інтеграції великих даних у статистику

Джерело: авторська розробка

Аналіз останніх досліджень і публікацій. Національні статистичні служби (НСС) у різних країнах світу активно співпрацюють з метою дослідження потенціалу новітніх інформаційних ресурсів. Використання великих даних розглядається як важливий інструмент для доповнення традиційних методів або безпосереднього формування показників економічної статистики. У межах діяльності Комітету експертів ООН із питань великих даних і науки про дані для офіційної статистики створено низку тематичних робочих груп, які зосереджують увагу на окремих типах інформаційних джерел. Серед них – результати дистанційного зондування Землі, масиви даних мобільного зв'язку, скановані торговельні дані, а також

інформація автоматичних ідентифікаційних систем. Така співпраця дає змогу статистичним установам спільно опрацьовувати технічні, правові та методологічні аспекти використання інноваційних даних [5].

У європейському регіоні координацію досліджень у цій сфері здійснює Європейська комісія, яка реалізує проєкт європейської статистичної системи ESSnet Big Data. Ініціатива охоплює низку напрямів, пов'язаних із використанням різноманітних джерел великих даних, зокрема онлайн- вакансій, інформації про «розумне» енергоспоживання, систем відстеження морських суден, фінансових транзакцій та супутникових спостережень [6]. Важливою перевагою цього проєкту є створення відкритої платформи для обміну досвідом між країнами-учасницями, а також публікація методичних напрацювань у відкритому доступі.

На відміну від європейських країн, багато національних статистичних служб Азійсько-Тихоокеанського регіону поки що здійснюють впровадження великих даних переважно самостійно або в межах окремих міжнародних ініціатив. Деякі установи, зокрема статистичні органи Індонезії та Австралії, вже накопичили практичний досвід і активно поширюють результати своїх досліджень [7]. Водночас значна частина країн регіону лише планує інтегрувати інноваційні інформаційні джерела у власні статистичні програми та виявляє готовність до опанування нових методів роботи з даними.

Особливу увагу в регіоні приділяють застосуванню великих даних у статистиці цін. Найбільш поширеними джерелами є скановані торговельні дані та онлайн-інформація про ціни. Скановані дані формуються на основі електронної реєстрації продажів споживчих товарів у торговельних точках через системи штрихкодової ідентифікації. Вони містять відомості про характеристики продукції, її вартість і обсяги реалізації. Натомість онлайн-дані отримують шляхом автоматизованого збирання інформації з вебсайтів роздрібних продавців. Завдяки відносно простому доступу технології веб-скрейпінгу набули ширшого використання порівняно зі сканованими даними.

На глобальному рівні питання застосування великих даних для розрахунку індексу споживчих цін активно досліджують спеціалізовані міжнародні експертні групи, зокрема Оттавська група з питань статистики цін і цільова група ООН із використання сканованих даних. До їхнього складу входять представники різних держав, серед яких Австралія, Нова Зеландія, Таїланд і Туреччина [8, 9].

Іншим важливим напрямом є використання інформації мобільних операторів для підготовки статистики туризму та міграційних процесів. Для цього розрізняють дані високої деталізації, наприклад сигнальні повідомлення мережі, а також записи про телефонні з'єднання (CDR). Сигнальні масиви містять відомості про зміну місцеперебування абонента під час перемикання між базовими станціями або визначення його координат за допомогою технологій допоміжної супутникової навігації. Наприклад, оцінка кількості туристів може ґрунтуватися на підрахунку унікальних користувачів:

$$N_{tourists} = \sum_{i=1}^{N} I_i$$

де $I_i = 1$, якщо користувач відповідає критеріям туриста, і 0 – не відповідає.

Записи даних про виклики містять відомості щодо часу здійснення дзвінка, його тривалості, номера абонента-ініціатора, номера отримувача, а також приблизного місця перебування користувача під час телефонного з'єднання або іншої телекомунікаційної операції, наприклад надсилання текстового повідомлення. Така інформація фіксується мережевим обладнанням і зберігається операторами мобільного зв'язку [10].

Окрім показників, пов'язаних зі статистикою туризму та соціально- демографічними характеристиками населення, використання мобільних даних дає змогу формувати додаткові аналітичні індикатори. На їхній основі можна оцінювати рівень міської мобільності, масштаби внутрішніх міграційних процесів, а також окремі демографічні параметри.

Мета статті – розкрити досвід національних статистичних управлінь країн Азійсько-Тихоокеанського регіону щодо використання великих даних у виробництві економічної статистики, а також перспективи і проблеми інтеграції таких джерел інформації у систему офіційної статистики.

Дослідження передбачає досягнення кількох взаємопов'язаних цілей. Це насамперед аналіз сучасних практик застосування великих даних для розрахунку ключових макроекономічних показників, зокрема індексів споживчих цін, індексів цін на житлову нерухомість, статистики туризму та інших експериментальних показників соціально-економічного розвитку. Результати аналізу дають можливість оцінити рівень готовності статистичних органів до використання нових цифрових джерел інформації [11].

Важливим завданням є виявлення найбільш ефективних джерел великих даних, серед яких онлайн-дані про ціни, скановані дані роздрібних мереж, інформація з мобільних телефонів, адміністративні податкові бази, фінансові транзакції та інші цифрові сліди економічної діяльності. Це дає змогу визначити, у яких напрямах такі джерела можуть доповнювати або частково замінювати традиційні статистичні обстеження.

Окремий акцент зроблено на оцінюванні переваг використання великих даних: підвищення оперативності отримання статистичної інформації, можливість формування більш деталізованих показників, зниження витрат на проведення обстежень, а також розширення аналітичних можливостей державної статистики. Водночас розкрито й основні проблеми впровадження великих даних, зокрема ті, що стосуються доступу до інформації, забезпечення конфіденційності персональних даних, узгодження нових джерел із традиційними методологіями та необхідності розвитку технічної інфраструктури.

Важливим завданням є аналіз застосування гедоністичних моделей для розрахунку цін (особливо для техніки та нерухомості):

$$P = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon$$

де x_i – характеристики товару (наприклад, пам'ять, бренд, площа житла), β_i – вагові коефіцієнти, ε – випадкова похибка. Це дає змогу враховувати якісні зміни товарів, що особливо важливо у використанні великих даних.

Аналіз також спрямований на узагальнення практичного досвіду окремих країн та формування рекомендацій щодо подальшого розвитку співпраці між національними статистичними установами, міжнародними організаціями та приватним сектором. Це дає змогу визначити стратегічні напрями модернізації офіційної статистики в умовах цифрової економіки [12].

Отже, головна мета дослідження полягає в обґрунтуванні доцільності інтеграції великих даних у систему економічної статистики та визначенні шляхів підвищення її якості, актуальності й аналітичної цінності в сучасних умовах глобальної цифровізації.

Результати дослідження. Міжнародна цільова група ООН, що займається питаннями використання мобільних даних у статистичній діяльності, розробила спеціалізовані методичні рекомендації щодо їх застосування в офіційній статистиці. У цих матеріалах надано опис підходів до формування статистики туризму та масових заходів, обліку населення і міграції, аналізу трудових поїздок, моніторингу транспортних потоків і зайнятості населення, а також наведено приклади моделей співпраці державних установ із телекомунікаційними операторами.

Статистичні органи різних країн, зокрема Індонезії та Грузії, вже розробили власні методологічні підходи до використання мобільних даних у сфері туристичної статистики та долучилися до міжнародних напрацювань у цьому напрямі. Водночас міжнародні організації, неурядові структури та приватні компанії надають урядам експертну підтримку щодо впровадження відповідних технологій, розроблення методик аналізу та забезпечення захисту конфіденційної інформації [13].

Крім сканованих торговельних даних, онлайн-інформації про ціни та відомостей операторів мобільного зв'язку, статистичні установи Азійсько-Тихоокеанського регіону вивчають інші перспективні джерела великих даних, здатні покращити якість економічної статистики. Наприклад, результати дистанційного спостереження за Землею, отримані зі супутників, літаків або безпілотних літальних апаратів, включно із зображеннями нічного освітлення територій, можуть бути використані для оцінювання рівня економічної активності.

Дані інтелектуальних лічильників, що відображають обсяги виробництва та споживання електроенергії домогосподарствами і підприємствами, здатні доповнювати традиційні статистичні показники, зокрема у сфері енергетики, аналізі сезонних коливань туристичних потоків або дослідженні впливу господарської діяльності на довкілля.

Важливим джерелом інформації також є масиви фінансових транзакцій, які можуть надходити від центральних банків, податкових органів

чи платіжних систем. Вони дають можливість оцінювати загальні тенденції розвитку економіки та споживчої поведінки. Додатковим індикатором економічної активності є дані про переміщення морських суден, сформовані на основі автоматизованих систем ідентифікації, які дають змогу аналізувати інтенсивність морської торгівлі та інші види господарської діяльності на морі (рис. 2).

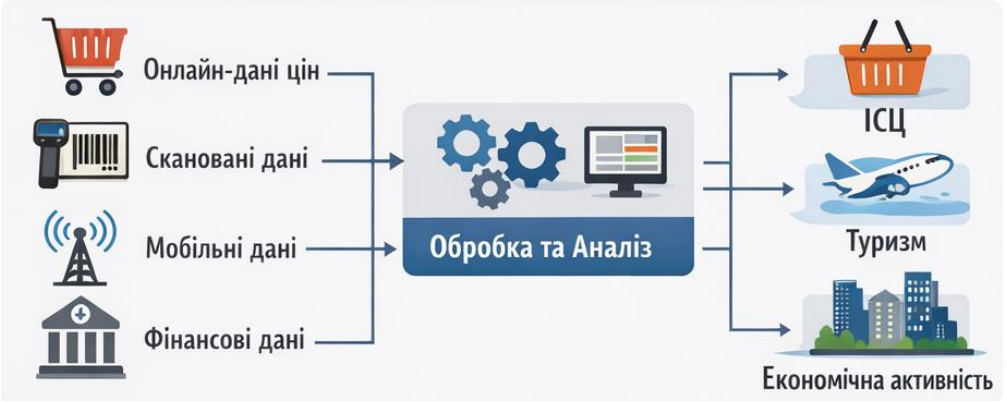

Рис. 2. Використання джерел великих даних у економічній статистиці
Джерело: [14]

Досвід національних статистичних служб Азійсько-Тихоокеанського регіону у використанні великих даних для економічної статистики. Більшість національних статистичних служб країн цього регіону застосовують інноваційні джерела інформації у сфері економічної статистики переважно в тестовому або пілотному режимі. Лише окремі установи переходять до системного впровадження великих даних у регулярний процес формування статистичних показників. Найбільш підготовленою сферою для такої інтеграції вважається статистика цін. Водночас поширення пандемії COVID-19 суттєво підвищило потребу у швидкій та детальній статистичній інформації, що стимулювало створення нових експериментальних економічних індикаторів [15].

Нами узагальнено практичний досвід дванадцяти країн – Азербайджану, Австралії, Грузії, Індонезії, Японії, Малайзії, Монголії, Нової Зеландії, Філіппін, Російської Федерації, Таїланду та В'єтнаму – щодо застосування великих даних у розрахунках індексів цін, підготовці туристичної статистики, оцінюванні рівня бідності, формуванні експериментальних економічних показників у період пандемії COVID-19, а також у дослідженні суспільних настроїв.

До основних інформаційних ресурсів, проаналізованих у цьому контексті, належать онлайн-дані про ціни, скановані торговельні масиви, відомості мобільних операторів, результати дистанційного спостереження за Землею, інформація про фінансові операції, податкові бази, дані систем нарахування заробітної плати, показники інтелектуальних лічильників і матеріали соціальних мереж [16].

Формування індексів цін (індексу споживчих цін та індексу вартості житлової нерухомості) із використанням великих даних. Одним

із центральних напрямів дослідження є використання великих даних для розрахунку індексу споживчих цін (ІСЦ). У загальному вигляді ІСЦ визначається за формулою:

$$CPI = \frac{\sum_{i=1}^{n} p_i^t \, q_i^0}{\sum_{i=1}^{n} p_i^0 \, q_i^0} \cdot 100$$

де p_i^t– ціна товару в поточному періоді, p_i^0 – ціна в базовому періоді, q_i^0 – обсяг споживання в базовому періоді.

Нами досліджено, як онлайн-дані та скановані дані можуть автоматизувати збирання p_i^t, що значно підвищує точність і оперативність розрахунків.

Обмежувальні заходи, запроваджені в багатьох країнах під час пандемії COVID-19, суттєво ускладнили традиційні способи збирання інформації для статистики цін, зокрема проведення польових опитувань і спостережень. Водночас період локдаунів сприяв активному зростанню онлайн-покупок на електронних торговельних платформах. У результаті статистичні органи отримали додатковий стимул до опрацювання онлайн- джерел цінової інформації, особливо в країнах із високим рівнем проникнення інтернет-технологій і розвитку електронної комерції.

Більшість прикладів інтеграції великих даних у систему економічної статистики пов'язані саме з обчисленням індексів цін на основі сканованих торговельних даних, інформації з інтернет-ресурсів або податкових реєстрів. Оскільки зазначені джерела мають різну природу формування та доступності, країни демонструють відмінний досвід у питаннях їх використання (рис. 3).

Рис. 3. Застосування великих даних у різних галузях

Джерело: [17]

Отримання сканованих даних зазвичай потребує тривалих переговорів і укладення угод із торговельними мережами, що може становити значні організаційні труднощі. Натомість доступ до онлайн-даних про ціни, який інколи також потребує погоджень із власниками веб-ресурсів, зазвичай є простішим для реалізації. Національні статистичні служби Австралії, Грузії,

Японії та Нової Зеландії вже мають можливість працювати як зі сканованими даними, так і з інформацією, отриманою з інтернет-платформ. Інші країни регіону перебувають на етапі вивчення потенціалу онлайн-джерел.

З огляду на поступову цифровізацію платіжних процесів і швидке зростання електронної торгівлі, дедалі більше країн розглядають нетрадиційні інформаційні ресурси як перспективний інструмент для вдосконалення розрахунку індексу споживчих цін.

Розрахунок індексу споживчих цін із використанням сканованих даних. Статистичні служби Австралії, Японії та Нової Зеландії вже інтегрували скановані торговельні масиви у процес визначення динаміки цін на окремі товари, що входять до споживчого кошика. Статистичні органи Грузії перебувають на етапі розроблення методологічної бази для впровадження цього підходу.

Статистичне бюро Японії стало одним із перших національних статистичних органів, які запровадили використання сканованих торговельних даних у процесі розрахунку ІСЦ. Ще у 2000 р. установа застосувала гедонічний підхід до оцінювання динаміки вартості персональних комп'ютерів, використовуючи інформацію з електронних систем обліку продажів. Згодом аналогічні методи було використано для аналізу змін цін на цифрові фотоапарати, а також під час експериментальних розрахунків ІСЦ для телевізорів і мобільних телефонів.

Австралійське бюро статистики почало активно досліджувати можливості залучення великих даних до формування економічних показників ще з 2011 р. Після проведення масштабних випробувань і пілотних досліджень у 2014 р. здійснено поступове впровадження сканованих торговельних масивів у процес офіційного розрахунку ІСЦ. Нині частка таких даних, отриманих від роздрібної торгівлі, перевищує чверть вагової структури ІСЦ.

Статистичне управління Нової Зеландії також розпочало роботу над методологією використання сканованих даних для оцінювання змін вартості товарів і послуг у 2011 р. Першочергову увагу приділили категорії споживчої електроніки, оскільки вона характеризується швидкими технологічними змінами, що ускладнює застосування традиційних підходів до вимірювання цін. Скановані дані щодо окремих видів електронної продукції інтегровано до офіційного ІСЦ перед початком 2014 р.

Національна статистична служба Грузії у 2020 р. уклала угоди з двома мережами супермаркетів, отримавши можливість регулярного доступу до електронних масивів даних про продажі. Пандемія COVID-19 суттєво прискорила перехід від традиційного збирання інформації про ціни у фізичних торговельних закладах до автоматизованої передачі даних у цифровому форматі. Разом із характеристиками товарів, їхньою ціною та обсягами реалізації роздрібні мережі надають і власну внутрішню класифікацію продукції. На її основі формують спеціальний конвертер між європейською системою штрихкодової ідентифікації товарів та класифікацією індивідуального споживання за цілями. Відповідний інструмент оновлюють щомісяця після оброблення нових масивів інформації.

Наразі розглядають можливість поширення співпраці й на інші торговельні мережі.

Загалом національні статистичні служби використовували різні підходи до отримання сканованих даних. Статистичні органи Австралії та Грузії отримували інформацію безпосередньо від роздрібних мереж, що потребувало встановлення партнерських відносин і гарантування конфіденційності на основі спеціальних меморандумів. Натомість у Японії та Новій Зеландії початковим джерелом таких даних стали компанії, що спеціалізуються на маркетингових дослідженнях. Водночас інші статистичні установи поки що не змогли досягти домовленостей із торговельними мережами і тому обмежують свої експериментальні дослідження використанням інформації, автоматично зібраної з інтернет- ресурсів.

Розрахунок індексу споживчих цін із використанням онлайн-даних про ціни. Статистичні служби країн Азійсько-Тихоокеанського регіону поступово впроваджують технології автоматизованого збирання інформації з інтернет-ресурсів, зокрема веб-скрейпінг, для формування ІСЦ. Такий підхід передбачає систематичне отримання цінових даних із вебсайтів інтернет-магазинів, сервісів бронювання та інших електронних торговельних платформ.

Австралійське бюро статистики з 2017 року застосовує інтернет- джерела у розрахунках окремих підіндексів ІСЦ, зокрема щодо одягу, побутової техніки, алкогольної продукції та обладнання для домашнього використання. Наприкінці 2018 р. установа розпочала додаткові дослідження щодо використання автоматизованих методів аналізу онлайн-даних у сегментах одягу та взуття. За оцінками фахівців, у перспективі такі інформаційні ресурси можуть забезпечувати до чверті показників ІСЦ. З метою збереження контролю над процесом збирання інформації статистична служба здійснює веб-моніторинг власними технічними засобами.

Статистичне управління Нової Зеландії також із 2017 р. проводить експерименти з використання онлайн-джерел. Дослідження показали, що для товарів доступно значно більше цифрових цінових даних ніж для послуг. Під час тестових проєктів створено цифровий індекс цін на продукти харчування на основі інформації з вебсайтів двох великих мереж супермаркетів. Однією з ключових проблем стала поява так званого «ланцюгового зміщення» у невиважених статистичних рядах. Надалі статистична служба почала безпосередньо взаємодіяти з торговельними мережами для отримання більш структурованих масивів даних, поєднуючи інформацію від зовнішніх постачальників із розвитком власних аналітичних інструментів.

Статистичне бюро Японії здійснило дослідження щодо визначення вартості авіаквитків на основі онлайн-інформації, отриманої від сторонніх провайдерів. Дані про рейси – включно з ціною, датою відправлення та маршрутом – щоденно збирали із сайтів бронювання провідних авіакомпаній. Для врахування різних рівнів знижок інформацію щодо одних

і тих самих перельотів фіксували в кілька часових періодів. Надалі результати автоматизованого збирання порівнювали з фактичними даними про продажі, отриманими під час інтерв'ю з представниками авіаперевізників. Основні висновки стосувалися необхідності отримання дозволів компаній на регулярний доступ до їхніх вебресурсів, а також потреби в інформації про обсяги реалізації для актуалізації вагових коефіцієнтів ІСЦ.

Департамент статистики Малайзії створив внутрішню цифрову платформу для дослідження цін із метою модернізації інструментів збирання інформації та підвищення якості розрахунку ІСЦ. У межах цієї системи впроваджено методи веб-сканування та автоматичного вилучення даних із відповідних сайтів. Також розроблено спеціальний словник відповідностей між описами товарів і національною класифікацією індивідуального споживання, сформований на основі алгоритмів текстового порівняння. Під час пандемії COVID-19 статистична служба використовувала онлайн-дані щодо вартості внутрішніх авіаперельотів і цін на основні споживчі товари, а також інформацію про орендну плату за житло в межах міжнародних статистичних зіставлень.

Окрім безпосереднього розрахунку ІСЦ малайзійські фахівці застосовують інтернет-дані для прогнозної аналітики. Вони моделюють вплив зміни вартості товарів на структуру споживчого кошика, прогнозують динаміку індексів за окремими групами товарів, регіонами та населеними пунктами, оцінюють тривалість цінових коливань і аналізують вплив погодних умов або валютного курсу на загальний рівень інфляції.

Статистичне управління Філіппін із лютого 2020 р. здійснює експериментальний моніторинг цін в окремих інтернет-магазинах. У межах цього проєкту регулярно відстежують вартість понад п'ятисот товарних позицій зі споживчого кошика столичного регіону, а отримані результати порівнюють із даними традиційних офлайн-спостережень.

Головне статистичне управління В'єтнаму також проводило пілотні дослідження щодо інтеграції онлайн-даних з електронних торговельних платформ у розрахунок ІСЦ. У співпраці з державними органами, відповідальними за розвиток електронної комерції, визначено перелік вебсайтів для збирання інформації. У результаті вдалося сформувати щоденні масиви даних щодо понад чотирьохсот споживчих товарів і послуг із десятків інтернет-ресурсів. Попри проведення наукових заходів і обговорень результатів досліджень, цей підхід поки що не інтегрований у систему офіційної статистики. Серед основних перешкод називають нормативні обмеження щодо використання нетрадиційних джерел інформації, недостатній рівень розвитку інфраструктури для зберігання великих даних, потребу у висококваліфікованих кадрах та значні початкові витрати [17].

Центральне статистичне управління Індонезії також здійснило низку експериментів із розрахунку ІСЦ на основі інформації з платформ електронної торгівлі. Було відібрано кілька ключових інтернет-майданчиків, що забезпечують найбільші обсяги онлайн-продажів. Аналіз показав, що більшість товарів зі споживчого кошика може бути врахована в розрахунках ІСЦ, однак відсутність даних про фактичні обсяги продажів

ускладнює визначення їхньої ваги. Додатковою проблемою є постійні зміни структури вебсайтів, що ускладнює автоматизоване збирання інформації. Тому статистична служба зосереджується на співпраці з обмеженим колом великих електронних платформ для забезпечення стабільного потоку даних. Наразі використання онлайн-цінової інформації залишається на стадії апробації та ще не стало частиною регулярного статистичного виробництва.

Водночас індонезійські фахівці досліджують і альтернативні джерела великих даних, зокрема можливості краудсорсингу, а також планують розширити партнерство з компаніями цифрового сектору. Згідно з національними нормативними актами, оператори електронної торгівлі зобов'язані регулярно передавати державним органам інформацію про цифрові транзакції, що створює передумови для подальшого розвитку статистичних досліджень у цій сфері.

Розрахунок індексу цін на житлову нерухомість (ІЦЖН) із використанням онлайн-даних. Ще одним напрямом застосування великих даних у статистиці цін є індекс цін на житлову нерухомість (ІЦЖН). У цьому контексті дослідження проводять національні статистичні управління Грузії та В'єтнаму, а також Статистичний департамент Банку Індонезії, які аналізують можливості використання веб-скрейпінгу для складання ІЦЖН. Геостат (Грузія) та головне статистичне управління (ГСУ) В'єтнаму отримали технічну підтримку від МВФ для включення великих даних у процес розрахунку ІЦЖН.

В'єтнам. ГСУ В'єтнаму виявило, що збирання даних через Інтернет може стати ефективним джерелом для розрахунку ІЦЖН, оскільки це зменшує витрати на оброблення інформації. Поки очікується створення адміністративного джерела даних про ціни угод із житлом, онлайн-дані дають змогу отримувати оперативну статистику.

Грузія. Геостат отримав методичну та навчальну підтримку з використання мови R для веб-скрейпінгу від МВФ, що дало змогу скласти експериментальні індекси на основі даних з двох веб-сайтів. Було підписано угоди з власниками сайтів, які дозволили автоматизоване збирання даних; воно розпочалось у 2019 р. та супроводжується належною інфраструктурою для оброблення інформації. У майбутньому Геостат планує публікувати експериментальні статистичні дані та розглядає можливість використання онлайн-даних для відстеження змін цін на вживані авто та електроніку [18].

Індонезія. Банк Індонезії також вивчав онлайн-дані як альтернативу традиційним джерелам, оскільки інформація про декларовані угоди, наприклад із кадастру чи податкових реєстрів, часто недоступна. Протягом 2016–2018 рр. зібрано дані з двох найбільших сайтів оголошень про нерухомість, а для обчислення ІЦЖН застосовано прямий гедоністичний підхід. Головною перевагою онлайн-даних є оперативність отримання цін, проте різниця між запитаною ціною та реальною ціною в угоді може впливати на точність оцінок.

Висновки. Національні статистичні управління (НСУ) Азіатсько- Тихоокеанського регіону реалізують численні проєкти, пов'язані з використанням великих даних у сфері економічної статистики. Проте

більшість таких ініціатив залишається на експериментальному етапі (рис. 4).

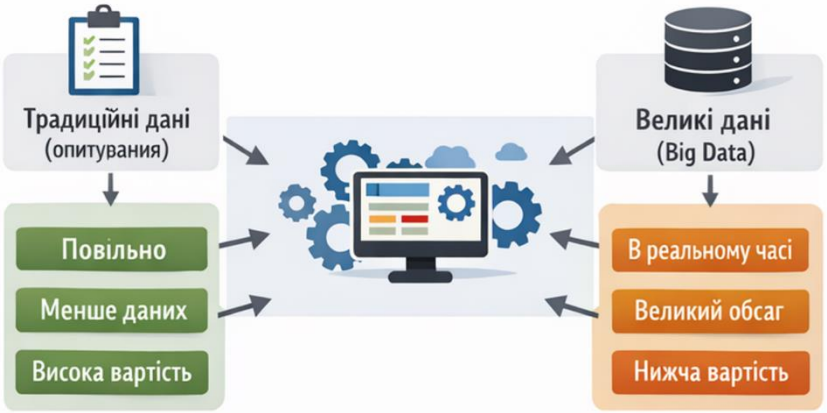

Рис. 4. Порівняння традиційних підходів у сфері економічної статистики і підходів із використанням великих даних

Джерело: авторська розробка

Винятком є статистика цін, де кілька НСУ вже інтегрували дані сканування та/або онлайн-дані у розрахунок ІСЦ для певних товарних категорій. У деяких випадках ці нетрадиційні джерела даних можуть повністю або частково замінити традиційні методи збирання інформації. Але великі дані найчастіше використовують додатково до традиційних джерел, забезпечуючи більш оперативну, детальну інформацію та нові аналітичні можливості.

З огляду на досвід НСУ, розкритий у статті, а також на зацікавленість інших статистичних установ у регіоні, найбільш перспективними джерелами великих даних для економічної статистики є дані сканування, онлайн-дані про ціни та дані мобільних телефонів.

Тому подальші регіональні зусилля мають бути спрямовані на розвиток потенціалу та компетенцій, обмін знаннями, укладання партнерств із приватним сектором заради забезпечення інтеграції великих даних у виробництво економічної статистики.

Надійшла 13.04.2026
Прорецензована 13.05.2026
Підписана до друку 28.05.2026
Оприлюднена 29.05.2026

Список використаних джерел

1. Бурашнікова О. Міжнародний досвід цифровізації в податковій сфері. *Humanities studies. Розділ «Філософія економіки та управління»*. 2023. Вип. 14 (91). С. 148– 157. DOI: 10.32782/hst-2023-14-91-17
2. Олексенко Р. І., Воронкова В. Г. Формування моделі класифікації соціальних процесів у публічному управлінні та адмініструванні: понятійно-категорійний апарат. *Теорія та практика державного управління*. 2020. № 3 (70). С. 82–90.
3. Метеленко Н. Г., Воронкова В. Г., Попова А. О., Оглобліна В. О., Сіліна І. В. Регіональна екологічна політика в умовах повномасштабної війни в контексті руйнівних наслідків для навколишнього середовища. Tabachnikov S. (Ed.) *Modern innovative strategies in education and science*. Sherman Oaks, California: GS Publishing Services, 2024. P. 5–24. DOI: 10.51587/9798-9895-14663-2024-119

4. Промисловий менеджмент: теорія і практика / За ред. Воронкової В. Г., Метеленко Н. Г. Запоріжжя: Запорізький нац. ун-т, 2020. 38 с. URL: https://dspace.znu.edu.ua/jspui/handle/12345/3282
5. Cherep A., Voronkova V., Andriukaitiene R., Nikitenko V. Концепція креативних технологій у туристичному бізнесі у умовах цифровізації. *Humanities studies. Розділ «Філософія економіки та управління».* 2020. Вип. 5 (82) С. 196–209. DOI: 10.26661/hst-2020-5-82-15
6. Цифрова трансформація промислового менеджменту: теорія і практика / За ред. Воронкової В. Г., Метеленко Н. Г. Львів–Торунь: Liha-Pres, 2023. 816 с.
7. Цифрова трансформація промислового менеджменту у контексті викликів, можливостей та змін / За ред. Воронкової В. Г., Метеленко Н. Г. Львів–Торунь: Liha-Pres, 2024. 592 с.
8. Фурсін О. О. Територіальне управління та місцеве самоврядування: проблеми, виміри, перспективи. Tabachnikov S. (Ed.). *Vectors of the development of science and education in the modern world*. Sherman Oaks, California: GS Publishing Services, 2023. С. 308–323. DOI: 10.51587/9798-9866-95976-2023-014-308-323
9. Управління сталим розвитком промислового підприємства: теорія і практика / За ред. Воронкової В. Г., Метеленко Н. Г.; МОН України, ЗНУ ІННІ. Запоріжжя: Видавничий дім «Гельветика», 2021. 586 с. URL: https://dspace.znu.edu.ua/jspui/handle/12345/5219
10. Промисловий потенціал складних соціально-економічних систем цифрового суспільства: макро-, мезо- та мікрорівень / За ред. Воронкової В. Г., Метеленко Н. Г. Запоріжжя: Видавничий дім «Гельветика». 2022. 480 с. URL: https://catalog.liha-pres.eu/index.php/liha-pres/catalog/ book/154
11. Слюсар М. Ю. Становлення і розвиток моделі мережевих платформ у Китаї та їх вплив на формування цифрової економіки. *Humanities studies. Розділ «Філософія економіки та управління»*. 2023. No. 14 (91). P. 165–175. DOI: 10.32782/hst-2023-14-91-19
12. Промисловий менеджмент: теорія і практика / За ред. Воронкової В. Г., Метеленко Н. Г. Запоріжжя: Запорізький нац. ун-т, 2020. 338 с. URL: https://dspace.znu.edu.ua/jspui/handle/12345/3282
13. Measuring the Digital Transformation: A Roadmap for the Future. Paris: OECD Publishing, 2022.
14. Самойленко Л. Б. Можливості та проблеми застосування технологій Big Data вітчизняними компаніями. *Ефективна економіка*. 2018. № 1. URL: http://www.economy.nayka.com.ua/pdf/1_2018/59.pdf
15. Handbook on the Use of Big Data for Official Statistics. United Nations Global Working Group on Big Data for Official Statistics. New York: United Nations, 2023. 218 p.
16. Орлов О. В. Великі дані у публічному управлінні: технології майбутнього. *Державне будівництво*. 2016. № 1. URL: http://nbuv.gov.ua/UJRN/DeBu_2016_1_3
17. European Statistics Code of Practice. Luxembourg: Eurostat, 2023. URL: https://ec.europa.eu/eurostat/web/quality/european-quality-standards/european-statistics-code-of-practice
18. Zikopoulos P. C., Eaton C., deRoos D., Deutsch T., Lapis G. Understanding Big Data. Analytics for Enterprise Class Hadoop and Streaming Data. Chicago: McGraw-Hill Osborne Media, 2023.

References

1. Burashnikova, O. (2023). Mizhnarodnyi dosvid tsyfrovizatsii v podatkovii sferi [International experience of digitalization in the tax field]. *Humanities studies. Rozdil "Filosofiia ekonomiky ta upravlinnia" – Humanities Studies. Section "Philosophy of Economics and Management", 14 (91),* 148–157. DOI: 10.32782/hst-2023-14-91-17 [in Ukrainian].
2. Oleksenko, R. I., & Voronkova, V. H. (2020). Formuvannia modeli klasyfikatsii sotsialnykh protsesiv u publichnomu upravlinni ta administruvanni: poniatiino-katehoriinyi aparat [Formation of a classification model of social processes in public

management and administration: Conceptual and categorical apparatus]. *Teoriia ta praktyka derzhavnoho upravlinnia – Theory and Practice of Public Administration*, *3 (70)*, 82–90 [in Ukrainian].

3. Metelenko, N. H., Voronkova, V. H., Popova, A. O., Ohloblina, V. O., & Silina, I. V. (2024). Rehionalna ekolohichna polityka v umovakh povnomasshtabnoi viiny v konteksti ruinivnykh naslidkiv dlia navkolyshnoho seredovyshcha [Regional environmental policy under conditions of full-scale war in the context of destructive environmental consequences]. Tabachnikov S. (Ed.) *Modern innovative strategies in education and science*. Sherman Oaks, California: GS Publishing Services, 2024. P. 5–24. DOI: 10.51587/9798-9895-14663-2024-119
4. Voronkova, V. H., & Metelenko, N. H. (Eds.). (2020). *Promyslovyi menedzhment: teoriia i praktyka [Industrial management: Theory and practice]*. Zaporizhzhia: Zaporizhzhia National University. Retrieved from https://dspace.znu.edu.ua/jspui/handle/12345/3282
5. Cherep, A., Voronkova, V., Andriukaitiene, R., & Nikitenko, V. (2020). Kontseptsiia kreatyvnykh tekhnolohii u turystychnomu biznesi u umovakh tsyfrovizatsii [The concept of creative digital technologies in the tourism business under digitalization]. *Humanities studies. Rozdil "Filosofiia ekonomiky ta upravlinnia" – Humanities Studies. Section "Philosophy of Economics and Management"*, *5 (82)*, 196–209. DOI: 10.26661/hst-2020-5-82-15 [in Ukrainian].
6. Voronkova, V. H., & Metelenko, N. H. (Eds.) (2023). *Tsyfrova transformatsiia promyslovoho menedzhmentu: teoriia i praktyka [Digital transformation of industrial management: Theory and practice]*. Lviv–Torun: Liha-Pres [in Ukrainian].
7. Voronkova, V. H., & Metelenko, N. H. (Eds.) (2024). *Tsyfrova transformatsiia promyslovoho menedzhmentu u konteksti vyklykiv, mozhlyvostei ta zmin [Digital transformation of industrial management in the context of challenges, opportunities and changes]*. Lviv–Torun: Liha-Pres [in Ukrainian].
8. Fursin, O. O. (2023). Terytorialne upravlinnia ta mistseve samovriaduvannia: problemy, vymiry, perspektyvy [Territorial governance and local self-government: Problems, dimensions, and prospects]. In: Tabachnikov, S. (Ed.). *Vectors of the development of science and education in the modern world* (pp. 308–323). Sherman Oaks, CA: GS Publishing Services. DOI: 10.51587/9798-9866-95976-2023-014-308-323 [in Ukrainian].
9. Voronkova, V. H., & Metelenko, N. H. (Eds.). (2021). *Upravlinnia stalym rozvytkom promyslovoho pidpryiemstva: teoriia i praktyka [Management of sustainable development of an industrial enterprise: Theory and practice]*. Zaporizhzhia: Helvetica Publishing House. Retrieved from https://dspace.znu.edu.ua/jspui/handle/12345/5219 [in Ukrainian].
10. Voronkova, V. H., & Metelenko, N. H. (Eds.). (2022). *Promyslovyi potentsial skladnykh sotsialno-ekonomichnykh system tsyfrovoho suspilstva: makro-, mezo- ta mikroriven [Industrial potential of complex socio-economic systems of the digital society: Macro-, meso- and micro-level]*. Zaporizhzhia: Helvetica Publishing House. Retrieved from https://catalog.liha-pres.eu/index.php/liha-pres/catalog/book/154 [in Ukrainian].
11. Sliusar, M. (2023). Stanovlennia i rozvytok modeli merezhevykh platform u Kytai ta yikh vplyv na formuvannia tsyfrovoi ekonomiky [Establishment and development of the network platform model in China and its impact on the formation of the digital economy]. *Humanities studies. Rozdil "Filosofiia ekonomiky ta upravlinnia" – Humanities Studies. Section "Philosophy of Economics and Management"*, *14 (91)*, 165–175. https://doi.org/10.32782/hst-2023-14-91-19 [in Ukrainian].
12. Voronkova, V. H., & Metelenko, N. H. (Eds.). (2020). *Promyslovyi menedzhment: teoriia i praktyka [Industrial management: Theory and practice]*. Zaporizhzhia: Zaporizhzhia National University [in Ukrainian].
13. OECD (2022). Measuring the Digital Transformation: A Roadmap for the Future. Paris: OECD Publishing.
14. Samoilenko, L. B. (2018). Mozhlyvosti ta problemy zastosuvannia tekhnolohii Big Data vitchyznianymy kompaniiamy [Opportunities and problems of using Big Data

technologies by domestic companies]. *Efektyvna ekonomika – Effective Economy, 1*. Retrieved from http://www.economy.nayka.com.ua/pdf/1_2018/59.pdf [in Ukrainian].

15. United Nations Global Working Group on Big Data for Official Statistics (2023). Handbook on the Use of Big Data for Official Statistics. New York: United Nations.
16. Orlov, O. V. (2016). Velyki dani u publichnomu upravlinni: tekhnolohii maibutnoho [Big Data in public administration: Technologies of the future]. *Derzhavne budivnytstvo – Public Construction*, 1. Retrieved from http://nbuv.gov.ua/UJRN/DeBu_2016_1_3 [in Ukrainian].
17. Eurostat (2023). European Statistics Code of Practice. Luxembourg: Eurostat. Retrieved from https://ec.europa.eu/eurostat/web/quality/european-quality-standards/european-statistics-code-of-practice
18. Zikopoulos, P. C., Eaton, C., DeRoos, D., Deutsch, T., & Lapis, G. (2023). *Understanding Big Data: Analytics for enterprise-class Hadoop and streaming data*. Chicago: McGraw-Hill Osborne Media.

Цитування:

Ставицький О. В., Пантєєв Р. Л. Використання основних джерел великих даних для формування економічної статистики. *Науковий вісник Національної академії статистики, обліку та аудиту: зб. наук. праць*. 2026. № 2. С. 42 – 58. DOI: 10.31767/nasoa.2-2026.03

Cite this article:

Stavytskyi, O. V., Pantyeyev, R. L. (2026). Vykorystannia osnovnykh dzherel velykykh danykh dlia formuvannia ekonomichnoi statystyky [Using the Main Sources of Big Data to Generate Economic Statistics]. *Naukovyi visnyk Natsionalnoi akademii statystyky, obliku ta audytu – Scientific Bulletin of the National Academy of Statistics, Accounting and Audit*, 2, 42 – 58. DOI: 10.31767/nasoa.2-2026.03