

О. О. Горобець,
*кандидат економічних наук,
Національна академія статистики, обліку та аудиту, м. Київ*

СИЛА ВЕЛИКИХ ДАНИХ У ПОДОЛАННІ ГЛОБАЛЬНИХ КРИЗ

Генеральний секретар ООН – Гутерреш А., під час доповіді «Наш спільний порядок денний» зазначив, що «...зараз світ переживає найбільше спільне випробування з часів Другої світової війни» [1], маючи на увазі пандемію Covid-19. Окрім того, у 2019 р. за матеріалами Всесвітнього економічного форуму було опубліковано «Звіт про глобальні ризики 2019» у якому виокремлено тринадцять основних ризиків, серед яких: кліматичні зміни, зміна ландшафту міжнародного управління, зростання хронічних захворювань у населення, старіння населення, зміщення влади, нерівність доходів, зростання географічної мобільності, посилення поляризації суспільства, зростання кіберзалежності та ін. Зазначене підлягало обговоренню й під час проведення Всесвітнього економічного форуму у 2022 р. [2], у рамках якого було зосереджено увагу на використанні великих даних з метою збільшення рівня обізнаності світової спільноти про глобальні загрози, вирішення питання про заповнення пробілів та дір у даних, у результаті чого наблизити перспективу вирішення питань економічних, екологічних, геополітичних, соціальних та технологічних глобальних криз.

З огляду на зазначене, засновники Future of Sustainable Data Alliance у 2020 р. опублікували початкові рекомендації щодо найкращих методів обробки даних для глобального майбутнього переходу до стійкої економіки з нульовими викидами вуглецю. Загалом, рекомендації містять три частини: заклик щодо визначення та створення шляхів до заповнення дір та пробілів у даних навколишнього середовища та управління (ESG); пропозиція зіставлення даних з таксономією та політикою сталого розвитку; потреба у розвитку талантів ESG у глобальному масштабі. Зі слів засновника IcebreakerOne Старкса Г. «...дані відіграють вирішальну роль у розвитку майбутнього. Ми закладаємо основи для надійної інфраструктури даних, яка дозволить реалізувати транскордонний та регульований обмін даними, який допоможе закрити прогалини та діри в даних і дозволить кожному долучитися до вирішення надзвичайної кліматичної ситуації» [3].

Усвідомлюючи важливе значення великих даних, досі немає єдиного алгоритму їхнього збирання, групування, очищення та аналізу, що призводить до спотворення отриманих результатів та формування хибних висновків. Враховуючи обсяг, швидкість, різноманітність, рівень достовірності, варіативності та наочність потоку великих даних важливо зауважити ще й на його непростій екосистемі. Виходячи з цього, доцільно додати ще одну характеристику великих даних – ієрархічну складність, викликану прямою залежністю від ІКТ, які своїм розвитком стимулюють появу та розгалуження унікальних даних.

Відомо, що вироблення даних – це безперервний процес, під час якого утворюється єдиний масив первинних (необроблених) даних, які умовно можна поділити на три ланки кожна з яких міститиме декілька взаємопов'язаних груп, тим самим формуючи екосистему даних. Так, перша ланка, містить групи даних від загальних адміністративних до даних індивідуальної та суспільної думки. Друга ланка, є більш гранульованою та персоналізованою і містить, зокрема, конфіденційні дані, які синергують між собою. Третя ланка в екосистемі виконує функцію своєрідного фільтру для даних та містить уже точні та гармонізовані дані, які сформовані у чотири глобальні групи – дрібні дані (Shallow data), глибокі дані (Deep data), мікро-дані (Micro-data) та нано-дані (Nano-Data). Варто додати, що будь-який вид даних має власні так звані «темні дані» (Dark data), які залишаються після майнінгування. Зазвичай темні дані повторно не використовуються, адже їхнє обслуговування є досить вартісним, що в свою чергу піднімає питання доцільності зберігання таких даних [4].

На підставі вищевикладеного, доцільно зауважити, що поява додаткових характеристик великих даних та виникнення нових видів даних є закономірною. У цьому контексті зазначимо, що наразі аналітики усе частіше звертають увагу на появу в екосистемі структурованих, неструктурованих та напівструктурованих даних нового виду – даних досвіду, які акумулюють у собі дані навчання, діяльності, поведінки та результативності [2]. Так, дані діяльності характеризують в першу чергу особливості повсякденного життя людини та її щоденні взаємодії; дані навчання містять свідчення про формальну та неформальну освіту, здобуті навички, компетенції та ін.; дані поведінки інформують про ефективність та активність людини у соціумі, допомагають виявити її потенційні можливості та загрози; дані результативності засвідчують

досягнення на робочому місці (джерелами таких даних виступають рейтинги, результати опитувань та ін.).

Хоча вивчення даних досвіду знаходиться на початковому етапі, однак уже зараз зрозуміло їхнє виняткове значення, яке полягає в організації проведення аналізу усього накопиченого людством досвіду, формуванні логічних висновків та імплементації отриманих результатів у практичну діяльність з метою уникнення або послаблення впливу глобальних криз.

Список використаних джерел:

1. United Nation. Our Common Agenda. Report of the Secretary-General. URL: https://www.un.org/en/content/common-agenda-report/assets/pdf/Common_Agenda_Report_English.pdf
2. Henriques A. Mind the gaps: how experience data can help fight climate change // World Economic Forum. 2022. URL: <https://www.weforum.org/agenda/2022/02/climate-change-experience-data/>
3. ESG Data holes and empty talent pools – Future of Sustainable Data Alliance publishes key initial recommendations // Future of Sustainable Data Alliance. URL: <https://futureofsustainabledata.com/esg-data-holes-and-empty-talent-pools-future-of-sustainable-data-alliance-publishes-key-initial-recommendations/>
4. Gartner. IT-glossary. Dark data. URL: <https://www.gartner.com/it-glossary/dark-data>