

АКТУАЛЬНІСТЬ ТА ПРОБЛЕМНІ ПИТАННЯ ВИКОРИСТАННЯ МОДЕЛІ BIG DATA У СТАТИСТИЦІ

Какаєв Шохрат,

аспірант,

Національна академія статистики, обліку та аудиту

Прогресуючий розвиток інформаційного суспільства, впровадження в повсякденну діяльність новітніх комунікаційних і вебтехнологій зумовлюють динамічне зростання та розширення інформаційних потоків, які необхідно відповідним чином систематизувати й обробляти. Використання соціальних мереж, різноманітних смарткомунікацій для поширення й отримання необхідної інформації, вивчення поведінки споживачів та населення загалом за допомогою відеоспостереження, аудіосервісів та комп'ютерних мереж спричиняють актуалізацію використання моделі Big Data при вимірюванні й оцінці різних соціально-економічних показників в реальному часі. Необхідність зміни існуючих підходів в офіційній статистиці до вивчення соціально-економічних явищ є наслідком одночасного формування великих масивів інформації та даних, їх хаотичності та неструктурованості, відсутності певної чіткої інфраструктури та методів обробки великих даних (Big Data).

“В епоху, коли знижується частка респондентів, які беруть участь в обстеженні домашніх господарств і підприємств, великі дані можуть забезпечувати директивні органи фактологічною інформацією в масштабі реального часу в таких сферах, як ціни, зайнятість, обсяг виробництва, економічний розвиток і динаміка населення. Великі дані можуть потенційно забезпечувати підготовку більш актуальних і своєчасних статистичних даних порівняно з традиційними джерелами для офіційної статистики. Включивши джерела великих даних у процес підготовки офіційної статистики, національні, регіональні та міжнародні статистичні організації могли б зміцнити свої позиції у частині своєчасного й економічно ефективнішого отримання даних офіційної статистики за секторами економіки, соціальної сфери і навколишнього середовища з меншими витратами ресурсів” [1, с. 33].

Цифрова економіка та диджиталізація бізнес-процесів сформували підґрунтя для нових джерел виникнення Big Data для офіційної статистики в режимі реального часу. Із цих джерел надходять дані, що формуються при застосуванні програмно-технологічних засобів, пристроїв реєстрації та сенсорних мереж. Існування та вираження Big Data у цифровій формі є передумовою модернізації процесів безперервного спостереження та вимірювання соціальних явищ і процесів, специфіки й динаміки функціонування державних інститутів. Аналіз міжнародного досвіду застосування моделі Big Data у статистиці засвідчує його перспективність та прогресивність для впровадження в Україні з метою побудови системи автоматизованої обробки великих масивів інформації для вирішення

прикладних завдань бізнесу та суспільно-соціальних проблем на основі формування агрегованих статистичних показників.

Опрацювання принципово нових підходів до статистичної обробки інформації завдяки використанню моделі Big Data сприятиме розробці новітніх та сучасних інструментів виявлення невідповідностей і поглибленого аналізу даних, що дозволятимуть своєчасно виявляти ймовірні втрати або викривлення інформації. Водночас поряд з виникненням синергетичного ефекту і розширенням спектру можливостей для офіційної статистики при застосуванні моделі Big Data, постає ряд проблемних питань, які створюють перешкоди на шляху оперативного впровадження моделі та її використання у національній практиці статистичних досліджень.

Застосування великих даних в офіційній статистиці породжує багато проблем, серед яких:

- юридичні, тобто пов'язані з доступом до даних і їх використанням;
- пов'язані з недоторканністю приватного життя, тобто з використанням громадської довіри й отриманням згоди на вторинне використання даних і їх ув'язку з іншими джерелами;
- фінансові, пов'язані з потенційними витратами на вилучення даних із джерела порівняно з отримуваними вигодами;
- управлінські, пов'язані, наприклад, із політикою та директивами з питань управління даними і забезпечення їх захисту;
- методологічні, пов'язані з якістю даних і придатністю статистичних методів;
- технологічні, пов'язані з інформаційними технологіями [1, с. 35].

В. Саріогло вважає, що “основними проблемами використання великих даних є труднощі у встановленні реального ступеня охоплення цільових сукупностей, можливість суттєвих зміщень оцінок показників на основі “великих даних”, доступність останніх, необхідність збереження й обробки дуже великих обсягів даних, ризики, пов'язані з можливістю маніпуляцій з даними та ін. Можливість використання “великих даних” в офіційній статистиці суттєво залежить від кваліфікації працівників. Сучасний статистик або статистики найближчого майбутнього мають володіти певними специфічними знаннями та навичками з інформаційних технологій, а саме: знати особливості побудови та функціонування web-сайтів, он-лайн сервісів, систем мобільного зв'язку тощо; використовувати інструменти автоматизованого збирання даних з сайтів, обробки значних обсягів інформації; вміти виконувати розрахунки та зберігати дані на основі хмарних технологій тощо” [3, с. 17].

Окремі дослідники звертають увагу, що великі дані як інформаційна категорія мають одну особливість на відміну від матеріальних ресурсів: для їх застосування необхідний по-справжньому високий рівень організації бізнес-процесів компанії. Без такого рівня підготовки, без наявності певної кваліфікації у бізнесу покупка (або збирання) великих даних буде характеризуватися настільки низькою ефективністю, що не виправдає вкладені кошти [2, с. 895].

Разом з тим переваги, які отримує суспільство, бізнес та офіційна статистика від використання моделі Big Data, перевищують та нівелюють вплив вищенаведених недоліків. Упровадження моделі Big Data забезпечує широку інтеграцію даних та необхідність реінжинірингу бізнес-процесів, що в подальшому забезпечить отримання синергетичного ефекту й сприятиме формуванню потенціалу для подальшого розвитку. Удосконалення інформаційної інфраструктури та статистична обробка великих даних забезпечать у майбутньому визначення точних і якісних характеристик та опису суспільних явищ, функціонування економіки в цілому.

Адаптація статистики України до особливостей застосування моделі Big Data повинна базуватися на перекваліфікації та підвищенні комп'ютерної грамотності працівників статистики, опанування ними навичок роботи зі спеціальними інструментами, комп'ютерними програмами автоматизованого збирання, накопичення й обробки даних, розробки методології та процедур статистичних досліджень в умовах використання великих даних, а також поступового усунення вищенаведених недоліків застосування аналізованої моделі. Впровадження міжнародного досвіду використання моделі Big Data в статистиці сприятиме раціональному та ефективному її впровадженню в Україні.

Список використаних джерел

1. Єршова О. Л., Томашевська Т. В. Великі дані (Big Data) і модернізація національних систем офіційної статистики // Прикладна статистика: проблеми теорії та практики. 2016. Вип. 18–19. С. 32–38.
2. Мінакова В. П., Шіковець К. О. Актуальність використання моделі Big Data в бізнес-процесах // Економіка і суспільство. 2017. Вип. 10. С. 892–896.
3. Саріогло В.Г. “Великі дані” як джерело інформації та інструментарій для офіційної статистики: потенціал, проблеми, перспективи // Статистика України. 2016. № 4. С. 12–19.

ДОСТОВІРНІСТЬ БУХГАЛТЕРСЬКОГО ОБЛІКУ У СИСТЕМІ ФОРМУВАННЯ ЯКІСНИХ СТАТИСТИЧНИХ ДАНИХ

Корінько Микола Данилович,
доктор економічних наук, професор,
завідувач кафедри аудиту та підприємництва,
Національна академія статистики, обліку та аудиту

Термін «Big Data» з'явився в 2008 році. Першим його вжив Кліффорд Лінч, редактор журналу Nature. Він розповідав про вибухове зростання обсягів світової інформації та відзначав, що освоїти їх допоможуть нові інструменти і більш розвинені технології. Одним із визначень поняття «Big Data» є таке: величезний масив даних, який при правильній обробці дозволяє виявити приховані закономірності та використовувати їх для підвищення