

Список використаних джерел

1. Онищенко І. М. Удосконалення методів обробки та зберігання даних за допомогою інструментів Big Data та Map Reduce // Економіко-математичне моделювання соціально-економічних систем. 2017. Вип. 22. С. 159–178.

2. Черняк Л. Большие Данные – новая теория и практика // Открытые системы. 2011. № 10 URL: <https://www.osp.ru/os/2011/10/13010990/> (дата обращения: 16.11.2017).

3. Big Data URL: <https://www.it.ua/ru/knowledge-base/technology-innovation/big-data-bolshie-dannye>

4. Big data: The next frontier for innovation, competition, and productivity / Manyika J. et al.; McKinsey Global Institute. June, 2011. URL: <http://www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation>

5. Research Big Data / Wikibon Inc. URL: <https://wikibon.com/research/big-data/> (viewed on November 10, 2017).

ВИКОРИСТАННЯ ДЖЕРЕЛ РІЗНИХ ДАНИХ ПІД ЧАС ПЕРЕПISУ НАСЕЛЕННЯ УКРАЇНИ В СУЧАСНИХ УМОВАХ

Єршова Ольга Леонідівна,

кандидат економічних наук,
завідувач кафедри економіко-математичних
дисциплін та інформаційних технологій;

Одноволик Вадим Іванович,

кандидат хімічних наук,
доцент кафедри економіко-математичних
дисциплін та інформаційних технологій;
Національна академія статистики, обліку та аудиту

Сучасний розвиток інформаційних технологій дозволяє організувати отримання та обробку важливих для держави даних різними способами. Рівень розвитку технологій в Україні уможливорює отримання даних з різних джерел та за різними аспектами – про країну, населення, бізнес-середовище та ін.

Сьогодні можна говорити про такі джерела та способи генерації відповідних даних:

- надання статистичної звітності підприємствами та організаціями, статистичні дослідження Державної служби статистики України;
- дані з державних реєстрів та інших державних сервісів;
- великі дані (big data), які генеруються інформаційно-комунікаційним обладнанням та обробляється відповідними компаніями.

Проведений аналіз та дослідження джерел показали, що деякі дані можна отримати з різних джерел та за різними технологіями. Ці дані будуть

достовірно та в повному обсязі описувати досліджуване явище. Використовуючи різні підходи, можна порівняти й оцінити дані з різних джерел.

Для прикладу розглянемо важливу майбутню подію – II Всеукраїнський перепис населення. Уряд прийняв рішення, що перепис відбудеться в листопаді – грудні 2020 року. Щодо методики його проведення існують досить різні думки. Уряд з метою економії коштів пропонує провести електронний перепис, а відомі вчені відповідають, що такий перепис не буде повним і не дасть відповідь на всі питання, які перед ним ставляться.

Ураховуючи сучасний стан мобільності населення України та практичну відсутність реєстрації місця проживання, проведення перепису й отримання реальних даних буде дуже непростим. Забезпечення персонального анкетування суттєво ускладнюватиметься відсутністю особи за місцем проживання та недовірою до влади, а отже, небажанням надавати достовірну інформацію.

Проведений аналіз свідчить, що отримати достовірні дані про населення можна, тільки максимально поєднавши та синхронізувавши різні джерела й методи отримання даних про сучасний стан населення України. При цьому мають бути використані такі процедури й передбачені такі вимоги:

1. Максимальне охоплення процедурою перепису всього населення країни. Дані анкет, на які кожна людина може відповісти тільки персонально, – про соціальний стан, рідну мову, освіту, національність та ін., будуть унікальними, їх неможливо отримати в інший спосіб. Для забезпечення повноти проведення перепису в сучасних умовах доцільно задіяти електронні анкети з можливістю дистанційного заповнення та облаштування ділянок перепису. Дуже важливим є проведення владою серйозної рекламної компанії з інформування про важливість участі в переписі. Необхідно також досить обережно ставитись до отриманих відповідей на запитання про доходи та власність.

2. Використання даних державних реєстрів про населення. Ідея уряду про верифікацію й об'єднання реєстрів в один реєстр громадян України є дуже слушною та вчасною. На сьогодні в країні ведуться різні реєстри громадян, що містять паспортні дані, відомості про майно, доходи, перетин кордону, про фізичних осіб – підприємців тощо. Деякі з цих реєстрів відкриті для користувачів, а їх дані можуть бути оброблені та проаналізовані. Результатом обробки окресленого масиву інформації можуть бути достовірні дані про доходи, власність, мобільність населення та ін. Робота з цим масивом даних дозволить оцінити результати, отримані під час перепису населення в масивах, які потенційно можуть бути викривлені неточними відповідями в анкетах.

3. Залучення великих компаній, які володіють великими даними (Big data) та обробляють їх для отримання додаткових даних про населення. Сьогодні поширеність та розвиток інформаційно-комунікаційних технологій в Україні дозволяє використовувати такі дані для характеристики населення.

Це можуть бути дані про мобільність, споживання, ділову активність та ін. Використання цього ресурсу потребує залучення відповідних компаній, які продукують великі дані, та законодавчого забезпечення роботи з даними такого характеру. Отримані результати роботи з такими даними можуть допомогти в обробці результатів перепису та верифікації даних у реальному режимі.

Проведена робота та аналіз дозволяють зробити висновок, що в сучасних умовах забезпечення достовірних даних про населення неможливо без використання сучасних технологій. Актуальні результати можна отримати, тільки використовуючи як традиційні, так і сучасні технології отримання даних та їх відповідної обробки.

МЕТОДОЛОГІЧНЕ ПІДГРУНТЯ ВКЛЮЧЕННЯ BIG DATA ДО ОФІЦІЙНОЇ СТАТИСТИКИ

Корепанов Олексій Сергійович,

доктор економічних наук, доцент,
професор кафедри статистики, обліку та аудиту;

Чала Тетяна Георгіївна,

кандидат економічних наук, доцент,
доцент кафедри статистики, обліку та аудиту;
Харківський національний університет імені В. Н. Каразіна

Понад п'ять років тому Євростат розпочав проект з метою «приручити» джерела великих даних (Big Data) так, щоб вони могли бути включені до офіційних статистичних систем. Проект спрямований на швидше створення статистичних даних та підвищення їх якості, що дозволить урізноманітнити аналіз даних, а в деяких випадках – зробити більш детальним, а також розробити точніші прогнози, одночасно зменшуючи навантаження на постачальників даних. Для досягнення цих цілей великі дані повинні стати невід'ємною частиною офіційного збирання статистичних даних [3]. Для вирішення проблем, з якими може зіткнутися статистик під час офіційного застосування Big Data для статистичних цілей, насамперед доцільно розглянути питання збирання даних, їх зберігання, підготовки, оброблення й аналізу, а потім визначити особливості, з якими доводиться стикатися при роботі з Big Data.

Порівняно з традиційними процедурами та методами збирання даних, щодо Big Data вони мають зовсім інший характер і функціонують на основі іншої логіки. Ця проблема включає питання інформаційних технологій та професійні дилеми. Щодо питання інформаційних технологій, то для того, щоб мати можливість керувати великими даними, збирати, зберігати,