



Doi: <https://doi.org/10.15407/dse2019.01.196>

УДК 303.436

JEL CLASSIFICATION: C55

А.В. ВЕРЕМЧУК

голов. економіст

Інститут демографії та соціальних досліджень
ім. М.В. Птухи НАН України
01032, Україна, м. Київ, бул. Шевченка, 60
E-mail: anne.veremtchouk@gmail.com

М.А. РОЗБИЦЬКИЙ

аспірант

Інститут демографії та соціальних досліджень
ім. М.В. Птухи НАН України
01032, Україна, м. Київ, бул. Шевченка, 60
E-mail: mishanyarozbitskiy@gmail.com

ОЦІНКА ПОТЕНЦІАЛУ «ВЕЛИКИХ ДАНИХ» ДЛЯ ДОСЛІДЖЕННЯ ТРУДОВОЇ МІГРАЦІЇ

Кількість українців, які працюють за кордоном, зростає, а наявні статистичні, соціологічні та адміністративні дані не можуть задовольнити потреби дослідників, експертів і фахівців органів державного управління у своєчасних та невикривлених даних. У результаті у публікаціях представлені суттєво відмінні оцінки масштабів і тенденцій міграційних процесів. У такій ситуації дослідники все частіше намагаються використати альтернативні джерела інформації, до яких належать «великі дані», які не тільки вирішують питання своєчасності отримання інформації, але і дають можливість отримати інформацію з питань, щодо яких раніше свідчення були «точковими» (наприклад, маршрути нелегальних мігрантів). Описані основні типи «великих даних», які можуть бути використані для вивчення міграції, їхня доступність для дослідника і основні недоліки. Для оцінки потенціалу використання «великих даних» виконано пілотне дослідження з використанням інформації з відкритих джерел, а саме Google Trends та Інтернет-сайтів для пошуку роботи. Показано, що дані Google Trends відображають зміну напрямів трудової міграції з Російської Федерації на користь Польщі після 2014 р. (коли відбулась окупація АР Крим і почався конфлікт на сході України). У той самий час, структура вакансій із Інтернет-сайтів для пошуку роботи не відповідає професійній структурі трудових мігрантів з України. Зроблено висновок, що на даний момент в Україні «великі дані» відкритого доступу можуть допомогти лише поверхово охарактеризувати тенденції міграційних процесів, але не можуть показати чисельність трудових мігрантів, тобто можуть слугувати доповненням до традиційних статистичних даних. До основних бар'єрів для формування рішень на основі «великих даних» у вивченні соціальних процесів належать низький рівень використання Інтернету в Україні (у порівнянні з розвинутими країнами), відсутність законодавчого регулювання питання доступу дослідників до даних приватних компаній і брак спеціалістів необхідної кваліфікації.

Ключові слова: трудова міграція, джерела даних, «великі дані», Google Trends, Інтернет, джерело даних.

© ВЕРЕМЧУК А.В., РОЗБИЦЬКИЙ М.А., 2019

А.В. Веремчук

главн. экономист
Институт демографии и социальных исследований
им. М.В. Птухи НАН Украины
01032, Украина, г. Киев, бул. Шевченко, 60
E-mail: anne.veremtchouk@gmail.com

М.А. Розбицкий

аспирант
Институт демографии и социальных исследований
им. М.В. Птухи НАН Украины
01032, Украина, г. Киев, бул. Шевченко, 60
E-mail: mishanyarozbitskiy@gmail.com

ОЦЕНКА ПОТЕНЦИАЛА «БОЛЬШИХ ДАННЫХ» ДЛЯ ИССЛЕДОВАНИЯ ТРУДОВОЙ МИГРАЦИИ

Количество украинцев, работающих за рубежом, растет, а имеющиеся статистические, социологические и административные данные не могут удовлетворить потребности исследователей, экспертов и специалистов органов государственного управления в своевременных и неискаженных сведениях. В результате в публикациях представлены существенно различные оценки масштабов и тенденций миграционных процессов. В такой ситуации исследователи все чаще совершают попытки использовать альтернативные источники информации, к которым относятся «большие данные», не только решающие вопрос о своевременности получения информации, но и дающие возможность получить информацию по сферам, в которых данные ранее были «точечными» (например, маршруты нелегальных мигрантов). Описаны основные типы «больших данных», которые могут быть использованы для изучения миграции, их доступность для исследователей и основные недостатки. Для оценки потенциала использования «больших данных» выполнено пилотное исследование с использованием информации из открытых источников, а именно Google Trends и Интернет-сайтов для поиска работы. Показано, что данные Google Trends отражают смену направлений трудовой миграции с Российской Федерации в пользу Польши после 2014 г. (когда произошла оккупация АР Крым и начался конфликт на востоке Украины). Однако структура вакансий на Интернет-сайтах для поиска работы не соответствует профессиональной структуре трудовых мигрантов из Украины. Сделан вывод, что на данный момент в Украине «большие данные» открытого доступа могут только поверхностно охарактеризовать тенденции миграционных процессов, но не могут указать численность трудовых мигрантов, они могут служить дополнением к традиционным статистическим данным. К основным барьерам для принятия решений на основании «больших данных» в изучении социальных процессов относятся низкий уровень использования Интернета в Украине (по сравнению с развитыми странами), отсутствие законодательного регулирования вопроса доступа исследователей к данным частных компаний и нехватка специалистов необходимой квалификации.

Ключевые слова: трудовая миграция, источники данных, «большие данные», Google Trends, Интернет, источник данных.

A.V. Veremchuk

Economist
Ptoukha Institute for Demography and Social Studies
of the National Academy of Sciences of Ukraine
01032, Ukraine, Kyiv, Taras Shevchenko Blvd, 60
E-mail: anne.veremtchouk@gmail.com

M.A. Rozbyckiy

Ph.D. Student
Ptoukha Institute for Demography and Social Studies
of the National Academy of Sciences of Ukraine
01032, Ukraine, Kyiv, Taras Shevchenko Blvd, 60
E-mail: mishanyarozbitskiy@gmail.com

ESTIMATION OF BIG DATA POTENTIAL IN LABOUR MIGRATION RESEARCH

In recent years the number of Ukrainians working abroad has been constantly increasing making labour migration one of the topical issues. Notwithstanding, the available statistics on labour migrants derived from statistical, social and administrative data cannot satisfy the needs of researchers, experts and public officials in timely and unbiased data. As a result, various estimates of labour migration scales and future trends are presented in articles. In this situation researches try to use alternative information sources, including Big Data. Big Data does not only solve the problem of timeliness in receiving information, but also allows obtaining accurate data in areas where earlier evidences were partial (for example, routes of illegal migrants). In the article the main types of “big data” which can be used for studying migration are described as well as their availability for researchers and main shortcomings. To estimate potential of “big data” usage, pilot research was conducted based on the information from open sources, namely Google Trends and job search websites. The results showed that data retrieved from Google Trends reflects the change of the labor migration directions from the Russian Federation to Poland after 2014 (when annexation of Crimea took place and the armed conflict in Eastern Ukraine started). At the same time, the structure of vacancies (coded according to Ukrainian analogue of ISCO) from the job search websites does not correspond to professional structure of labor migrants from Ukraine. It was concluded, that “big data” (from open sources) can only be used for superficial description of trends in migration processes, but cannot help researches in identifying scales of migration. Thus, today in Ukraine “big data” cannot replace traditional statistics in migration research. The main barriers for implementation of “big data”-driven decisions are (1) low level of Internet usage in Ukraine (compared to developed countries); (2) absence of legislative regulation of researchers’ access to data retained by private companies and (3) shortage of specialists with necessary qualifications.

Keywords: labor migration, data sources, Big Data, Google Trends, Internet, data source.

Постановка проблеми та актуальність. Складна економічна ситуація в Україні обумовила нині значну актуальність питань трудової міграції. На даний момент точні дані про кількість українців, які працюють за кордоном, відсутні, Державна служба статистики України вказує 1,3 млн осіб (за обстеженням «Зовнішня трудова міграція населення України» 2015–2017 рр.), а за інформацією з інших джерел кількість може перевищувати 2 млн осіб [1].

Основна причина такої різниці в оцінках обумовлена відсутністю єдиної методики дослідження процесів трудової міграції в Україні. Джерелами даних про трудову міграцію у різних країнах слугують реєстри населення (адміністративні дані), дані про посвідки на проживання, перетин кордону та обстеження (опитування)¹ [1]. В Україні Держстат надає оцінки про кількість трудових мігрантів на основі обстеження «Зовнішня трудова міграція населення України», яке мало три раунди 2005–2008, 2010–2012, 2015–2017 рр. Основними недоліками цього джерела даних є зростання кількості осіб, які відмовляються брати участь в обстеженні; надання респондентами некоректної інформації (особливо, якщо перебування за кордоном було нелегальним); зміщення оцінок через неохоплення осіб із доходом вище від середнього; несвоєчасність даних через значний часовий проміжок між збором інформації та оприлюдненням результатів дослідження. Крім цього, точність оцінок кількості трудових мігрантів викликає сумніви через відсутність даних щодо чисельності всього населення – останній перепис відбувся в грудні 2001 року. У даному випадку проблема визначення точної кількості осіб, які працюють за кордоном, посилюється через неврегулювання питання щодо статусу осіб із подвійним громадянством. Законом України заборонено мати подвійне громадянство, однак частина населення, живучи більшість часу на території України і маючи паспорт громадянина України, отримує паспорти інших держав (наприклад, Російської Федерації, Угорщини, Болгарії) або має документи, які дозволяють працювати за кордоном протягом певного періоду

¹ Це традиційні джерела даних, які вже добре описані в літературі і не є предметом статті, тому детальна інформація про їхні переваги / недоліки відсутня у тексті.

без додаткових дозволів / віз (наприклад, карта поляка або біометричний паспорт в Угорщині).

У результаті виникає необхідність залучення нових джерел інформації для здійснення оцінки трудової міграції. Таким джерелом можуть бути дані країн-реципієнтів трудових мігрантів. Наприклад, інформацію про видані дозволи на проживання за країною громадянства публікує Євростат. Істотним недоліком таких даних є те, що вони враховують тільки тих, хто легально перебуває в країні. Крім цього, в останній час здійснюються спроби адаптувати до потреб статистики інше альтернативне джерело – «великі дані» (*Big Data*). Оскільки «великі дані» – це масиви інформації, яка оновлюється в режимі реального часу, існує можливість оперативно здійснювати оцінку того чи іншого процесу та відповідно коригувати прогнози розвитку явища наукастингу (від англ. *nowcasting*) – прогнозування теперішнього та найближчого майбутнього [2].

Аналіз наявних досліджень. У західній літературі вже є розробки щодо використання «великих даних» у дослідженні міграційних процесів². Мова йде, наприклад, про роботи Е. Загхені і ін. (E. Zagheni et al.) [3], Б. Стате та ін. (B. State et al.) [4], Дж.Е. Блуменстока (J.E. Blumenstock) [5], Я. Новака і ін. (J. Novak et al.) [6], Д.К. Владики (D.K. Wladyka) [7], М. Бохме (M. Böhme) [8]. Українські публікації про використання «великих даних» у вивченні міграції під час аналізу літератури ми не знайшли. Однак у сфері соціальних наук опубліковано низку тематично розрізаних робіт, в яких охарактеризовано вплив «великих даних» або проілюстровано їх використання. Це статті О. Кислової (O. Kyslova) [9], В. Сариогло (V. Sarioglo) [10], О. Хмелевської (O. Khmelevska) [11], В. Кучерук та ін. (V. Kucheruk et al.) [12].

Варто зауважити, що основним недоліком широкого впровадження рішень на основі «великих даних» в Україні є, порівняно з іншими країнами, низький рівень penetрації Інтернету. За даними Світового банку [13], в Україні у 2016 році відсоток користувачів Інтернету серед всього населення становив 53 %; для порівняння: у країнах Вишеградської групи найнижчий відсоток у Польщі – 73,3 %, серед держав Прибалтики – у Литві (74,4 %). Такий низький рівень проникнення Інтернету в Україні ставить під сумнів можливості використання «великих даних» для аналізу.

Метою статті є оцінка можливості використання «великих даних» окремо або разом із даними з традиційних джерел для оцінювання різних аспектів трудової міграції в Україні.

Наукова новизна полягає у здійсненні пілотного дослідження з порівняння даних офіційної української статистики щодо трудової міграції з даними, отриманими з Інтернет-джерел («великі дані»).

Методи і матеріали. Інформаційною базою дослідження були дані з платформи *Google Trends*; дані про вакансії з Інтернет-сайтів пошуку роботи; результати дослідження «Зовнішня трудова міграція населення України», виконаного Державною службою статистики України. Результати аналізу даних, отриманих з Інтернету, було зіставлено з результатами дослідження «Зовнішня трудова міграція населення України» (на рівні порівняння відсотків) для визначення доцільності використання даних з Інтернету для вивчення трудової міграції.

Виклад основного матеріалу. У цій статті використано визначення «великих даних», яке подає ООН [14, с. 13]: «великі дані» – це великий за обсягом масив структурованої та неструктурованої інформації, обробка якої вимагає використання

² Міграція інтерпретується у даних дослідження як просторова мобільність. Далі у статті проблему такого тлумачення розглянуто детальніше.

новітніх технік. Виділяють [14, с. 13] три головні характеристики «великих даних»: 1) великий обсяг (розмір бази даних); 2) високу різноманітність (обробляються дані різних джерел і різних типів); 3) високий рівень швидкості накопичення даних та їх первинної обробки. Ці три характеристики називають «три V» від англійських слів *volume* (обсяг), *variety* (різноманітність) та *velocity* (швидкість). Для вивчення міграційних процесів, у тому числі трудової міграції, використовують дані з різноманітних джерел (табл. 1³). Основною перевагою використання «великих даних» є можливість отримати дані своєчасно (з мінімальним проміжком часу між збором даних та їх доступністю для обробки дослідником або у режимі реального часу), що полегшує прогнозування міграційних потоків, які раніше вважали «неочікуваними» (наприклад, у результаті війни / екологічної катастрофи і ін.); ці дані охоплюють значні за чисельністю групи людей; умови доступу до джерела даних, а саме відсутність фінансових витрат на їх збирання.

Основною проблемою у вивченні міграції на основі «великих даних» є використання дослідниками геотегів (наприклад, із повідомлень у соціальних мережах або дані мобільних операторів). Таким чином відстежується виключно просторове переміщення, а інформація щодо мети переміщення (туризм, навчання, робота тощо) залишається поза полем зору дослідників [16, с. 51–52]. У разі вимірювання мобільності робота — дім можна на основі часових меж (день, ніч, час доби) встановити місцезнаходження дому і місця роботи [6]. У випадку вивчення міжнародної міграції ситуація ускладнюється неможливістю відділити туристів від трудових мігрантів і мігрантів, які переїжджають назавжди; встановити громадянство особи.

Іншим джерелом, яке може спричинити зсув (некоректність) оцінок на основі «великих даних», є відсутність достовірної інформації про користувачів (абонентів). Так, користувачі соціальних мереж можуть подавати про себе викривлену інформацію, а дослідники не мають змогу її перевірити. У випадку даних мобільних операторів, навіть якщо придбання номеру здійснюється за паспортом, особа, яка купила номер, може передати номер у користування іншій особі. Хоча навряд чи таке викривлення інформації значуще, передачу номерів практикує не значна кількість людей, але це підкреслює проблему репрезентативності у подібних дослідженнях. Через відсутність точних даних про соціально-демографічні характеристики сукупності, від якої була отримана інформація, дослідники нічого не можуть сказати щодо того, чи отримані дані є репрезентативними стосовно всього населення чи певних груп.

Для нашого подальшого дослідження було використано дані *Google Trends* [17] щодо запитів, які стосуються роботи за кордоном. Щоб визначити, чи можуть дані з цього джерела бути використані для визначення основних тенденцій у трудовій міграції в Україні, їх було зіставлено з даними результатів дослідження «Зовнішня трудова міграція населення України». Дане обстеження здійснює Державна служба статистики України, охоплюючи осіб віком від 15 до 70 років, які працювали за кордоном упродовж відповідного періоду: для 2010–2012 рр. — з 1 січня 2010 р. до 17 червня 2012 р. (включає АР Крим), для 2015–2018 рр. — з 1 січня 2015 р. до 18 червня 2017 р. (не включає АР Крим). Методологія обстеження передбачає, що опитати можуть: осіб, які працювали за кордоном у визначений період і повернулися в Україну на момент проведення опитування; для осіб, які на момент опитування все ще працювали за кордоном — можна було опитати членів їх домогосподарств. Дані *Google Trends* (включають

³ У табл. 1 представлені основні джерела, які можуть бути застосовані для вивчення міграції. Цей перелік не є вичерпним і постійно доповнюється. Наприклад, є спроби використання даних супутникових знімків для вивчення міграції, що дає можливість відстежити шляхи біженців / мігрантів [15].

Таблиця 1. Основні переваги і недоліки використанні основних джерел «великих даних» для вивчення трудової міграції

Джерело	Потенціал використання у дослідженні міграції	Складнощі, які виникають у роботі з даними	Доступність для дослідника
Соціальні мережі	Відстеження основних трендів у переміщеннях на основі геотегів за певні часові інтервали	Відсутність достовірних даних про користувачів	Недоступні / доступні за запитом (доступна інформація з відкритих профілів)
Мобільні оператори	Відстеження переміщення на основі місцезнаходження абонента (абонент не може викривити дані, як у випадку соціальних мереж)	Необхідність використання даних декількох операторів, якщо ринок не монополізований; проблема репрезентативності за низького рівня користування мобільним зв'язком; відсутність достовірних даних про користувачів; геолокація фіксується тільки коли мобільний телефон ввімкнений	Недоступні / доступні за запитом
Сайти, на яких розміщуються об'яви про вакансії за кордоном	Оцінка переліку, структури та динаміки зміни професій, попит на які потенційно може бути задоволений у країні-реципієнті за рахунок мігрантів	Кодування професій за класифікатором професій; відсутність даних про закриття вакансій, розміщених на сайті; можливість дублювання вакансій на різних сайтах	Доступні для дослідника
Google Trends дані (індекс) ⁴	Запити користувачів Інтернету можуть бути використані для відслідковування тенденцій у міграції (наприклад, зміна напрямів міграції) та у прогнозуванні міграції	Індекс може варіюватися через удосконалення технології його розрахунку; відсутність достовірних даних про користувачів; запити можуть бути зумовлені цікавістю до теми, а не міграційними настроями	Доступні для дослідника

Джерело: створено авторами на основі [3–8, 16].

АР Крим як територію України) розглянуто за період 01.01.2013–31.12.2018. Дані до 2013 р. не розглянуті через низький рівень користування Інтернетом у цей період в Україні (у 2013 р. – 40,9 % населення, у 2012 р. – 35,2 % [13]). Безумовно 40,9 % не є достатньою часткою для визначення тенденцій на рівні України, але включення даних від 2013 р. дасть змогу відстежити зміну напрямів міграції після початку конфлікту на сході України та окупації Криму.

⁴ Google trends – платформа, на якій користувач отримує доступ до даних щодо відносної частоти пошуку (індекс) будь-яких термінів (пошукових запитів) [17]. Користувач сам уводить термін, визначає регіон (наприклад, Україна або Київська область) та період часу і отримує дані щодо популярності терміну у вказаному регіоні за вказаний період часу. Індекс розраховується як відношення обсягу цих запитів до всіх запитів у Google. Отримане число нормується на шкалі від 0 до 100, відповідно 100 показує, коли саме пошуковий запит був «найбільш популярним».

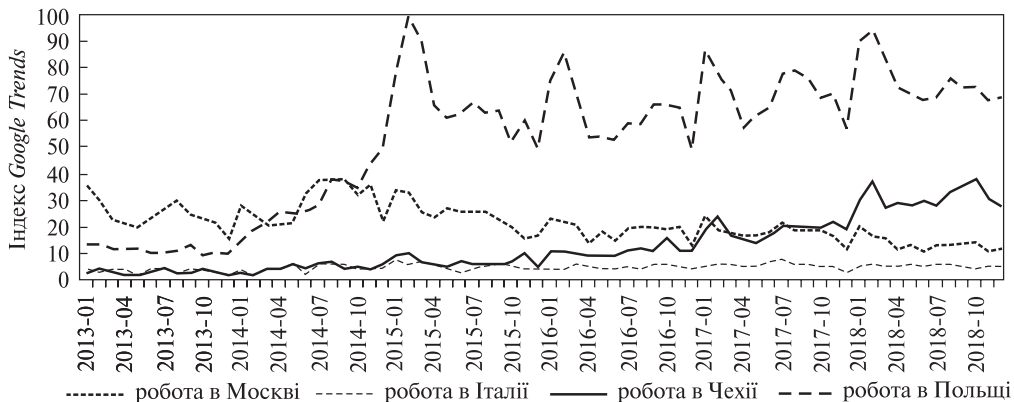


Рис. 1. Динаміка пошукових запитів щодо роботи за межами України, період 01.01.2013–31.12.2018

Джерело: побудовано авторами на основі даних Google Trends [17].

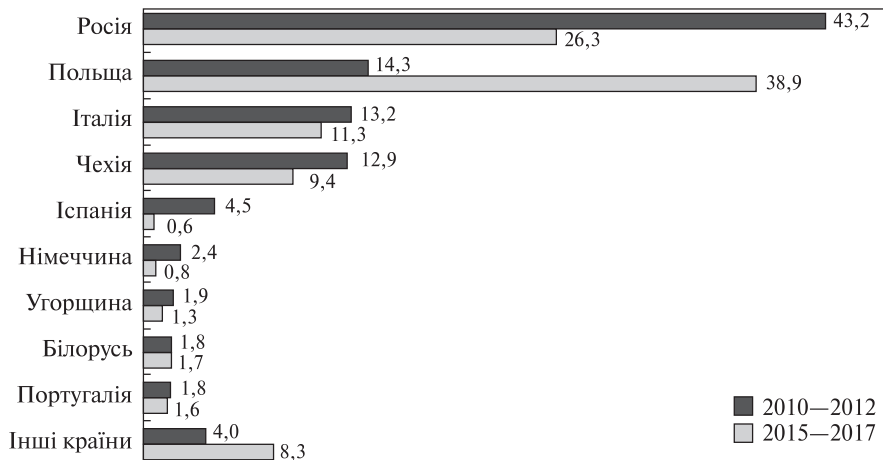


Рис. 2. Порівняння часток трудових мігрантів за країнами, в яких вони працювали, %

Джерело: для 2010–2012 рр. дані були взяті з [18], для 2015–2017 рр. – розрахунки авторів на основі масиву даних дослідження «Зовнішня трудова міграція населення України» за 2015–2017 рр.

На основі дослідження «Зовнішня трудова міграція населення України» були виділені країни, куди прямує найбільше трудових мігрантів: Італія, Польща, Російська Федерація та Чехія. Далі для кожної країни був створений перелік пошукових запитів, які містили слово робота чи вакансія (наприклад, «робота в Італії», «робота Італія», «робота Рим» і ін.). У пошукових запитах використано також назви країни або столиці. Щодо Російської Федерації у запитах враховано і місто Санкт-Петербург, Італії – Мілан. На етапі відбору пошукового запиту використано запити українською, російською, англійською мовою та національною мовою країни (наприклад, для Італії – італійською). Далі зі всього переліку запитів для одної країни відбирали один запит, який мав найвищу середню частоту за досліджуваний період. Для пошуку роботи в Італії таким запитом був «робота в Італії», для Польщі – «робота в Польщі», для Росії – «робота в Москві», в Чехії – «робота в Чехії» (усі запити російською мовою). Результати зіставлення запитів ілюструє рис. 1.

За даними *Google Trends* (рис. 1) простежується зміна основних напрямів пошуку роботи за кордоном: якщо до 2014 р. (коли почався збройний конфлікт і відбулась окупація АР Крим) роботу в основному шукали у Російській Федерації, то після 2014 р. відбулося зниження кількості пошукових запитів щодо роботи у РФ та значне збільшення кількості запитів щодо роботи у Польщі. Це знаходить відображення і у даних обстеження «Зовнішня трудова міграція населення України» (рис. 2), які показують переорієнтацію напрямів трудової міграції з Росії на Польщу. Крім цього, за даними *Google Trends* можна побачити, що «пікова» популярність роботи у Польщі припадає на лютий 2015 року, що можна зіставити з подіями на сході країни (у січні 2015 р. активізувались бої за Донецький аеропорт) і з девальвацією національної валюти⁵, що значно могло вплинути на міграційні настрої населення в Україні через доларизацію економіки. Зростання індексу *Google Trends* щодо запитів про роботу у Польщі відбувається в основному у січні – березні кожного року. Найбільш імовірно, в дані періоди користувачі Інтернету шукають інформацію щодо необхідних документів для виїзду на сезонні роботи. За даними Євростату, у Польщі у 2017 році було видано близько 545 тис. дозволів на проживання громадянам України (у зв'язку з працевлаштуванням, уперше видані дозволи [19]), з них – 91 % на сезонні роботи.

Результати опитування домогосподарств показують, що частки осіб, які мали роботу в Італії та Чехії, доволі близькі та значних змін до / після 2014 р. не спостерігаємо. Це в принципі підтверджується на даних *Google Trends*, але крива запитів щодо роботи у Чехії з другої половини 2016 р. поступово зростає. Однак, чи відбулась зміна напрямів міграції (з Італії на Чехію), можна буде визначити тільки після отримання нових даних обстеження.

Ще одним джерелом даних щодо роботи за межами України можуть бути сайти, на яких розміщуються об'яви про наявні вакансії за кордоном. Наприклад, нижче наведено результати, отримані на основі аналізу об'яв про вакансії з двох сайтів – *hh.ua* та *flagma.pl*. Сайт *hh.ua* входить у п'ятірку сайтів для пошуку роботи в Україні (вересень 2018 р., за даними *Factum Group* [20], – за рівнем охоплення аудиторії 1+). На цьому сайті розміщено оголошення про вакансії в Україні та за її межами. З цього сайту були відібрані лише пропозиції у РФ. Сайт *flagma.pl* має найбільший трафік із України (станом на 11.2018 – 30,5 % [21]) серед польських сайтів для пошуку роботи (проаналізовано трафік за даними сайтів *praca.pl*, *pl.jooble.org*, *infopraca.pl*, *pracuj.pl*), оскільки оголошення на сайті представлені в основному російською та українською мовами. З цього сайту було відібрано вакансії у Польщі.

Вивантаження оголошень із сайту виконано за допомогою програми з платформи *dexi.io* [22]. Однією з переваг цієї платформи є те, що в ній є візуальний редактор веб-механізмів, що дає змогу працювати з програмою людям, не знайомим із кодуванням. Ще одна перевага полягає в тому, що всі створені та використані користувачем «інструменти» (алгоритми програми) контролюються самою платформою, тому не потрібно налаштовувати сервер та внутрішні бази даних. Завдяки широкому спектру функцій у наборі інструментів було створено інструмент, який сканував веб-сторінки сайтів для пошуку роботи, «витягував» та трансформував веб-дані у таблиці (рис. 3). Швидкість сканування становить близько 30 вакансій за хвилину (враховуючи, що середня кількість вакансій на сторінці – 20), з цього впливає, що за повний день роботи створюється база вакансій із понад 40 000 позицій, які можуть бути структуровані по містах, професіях тощо.

⁵ У 2014 р. 100 доларів США коштували 1188 грн 67 коп. (середнє за період), у січні 2015 р. – 1581 грн 27 коп., у лютому 2015 р. – 2447 грн 99 коп.; 100 євро у 2014 році відповідали 1571 грн 59 коп. (середнє за період), у січні 2015 р. – 1853 грн 45 коп., у лютому 2015 р. – 2779 грн 68 коп.

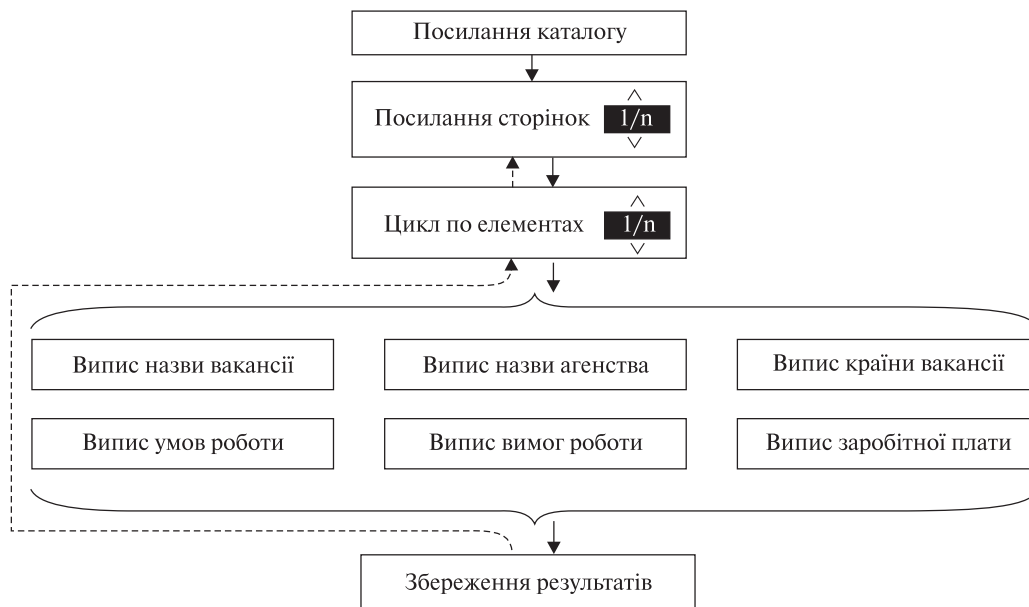


Рис. 3. Алгоритм дій робота для сканування сайту

Джерело: розроблено авторами на основі інформації з сайту [22].

За назвою роботи в оголошенні усі вакансії були закодовані за Класифікатором професій (версія від 2010 р.; на рівні коду з першого знаку; на основі порівняння назви роботи у вакансії та назви професії). Істотним недоліком такого кодування є те, що, наприклад, всі вакансії менеджерів були віднесені до категорії «Законодавці, вищі державні службовці, керівники, менеджери». На наступному етапі ми вирішили перевірити, чи відповідає структура вакансій із сайтів пошуку роботи структурі професій трудових мігрантів з України у Польщі та РФ. Дані у табл. 2 свідчать, що структура вакансій із сайтів не відповідає структурі професій трудових мігрантів, за винятком групи «найпростіші професії» у Польщі. Також за вакансіями з сайтів видно зсув вакансій у Польщі у бік найпростіших професій, а у РФ – у бік професій з кодами 1–4 (за результатами дослідження цей зсув не настільки значний: за професіями з кодами 1–4 у РФ працювало 13,3 % мігрантів, а у Польщі – 2,4 %). Цей зсув може бути зумовлений різним типом сайтів: на *flagma.pl* роботодавці можуть розмістити об’яви безкоштовно, а на *hh.ua* за розміщення оголошення необхідно сплатити, що відразу «відфільтровує» значну частину роботодавців, які шукають працівників для роботи на заводах або у сільському господарстві.

Відмінність між даними з сайту пошуку роботи та даними статистичного обстеження може бути результатом того, що не були розглянуті всі сайти для пошуку роботи: наприклад, як зазначено вище, на сайті *flagma.pl* оголошення в основному представлені російською і українською мовами. Відповідно, професії, які вимагають знання мови (коди 1–5) будуть розміщені роботодавцем на інших сайтах. Також є обмеженням, що дані з сайту для пошуку роботи було взято станом на одну дату. Вірогідно, що за умови відстеження довгих часових проміжків, результати були б точнішими.

Таблиця 2. Структура вакансій із сайтів пошуку роботи і структура професій за дослідженням «Зовнішня трудова міграція населення України», %

Назва професійної групи (за Класифікатором професій 2010)	За даними з Інтернет-сайтів пошуку роботи		За даними дослідження «Зовнішня трудова міграція»	
	<i>flagma.pl</i> (Польща)	<i>hh.ua</i> (РФ)	Польща	РФ
1. Законодавці, вищі державні службовці, керівники, менеджери	0,4	27,8	0,0	1,3
2. Професіонали	0,4	15,6	0,0	0,5
3. Фахівці	1,4	9,6	2,4	11,4
4. Технічні службовці	0,9	3,4	0,0	0,0
5. Працівники сфери торгівлі та послуг	4,4	32,1	17,0	7,6
6. Кваліфіковані робітники сільського та лісового господарств, риборозведення та рибальства	1,7	0,0	0,7	4,5
7. Кваліфіковані робітники з інструментом	26,2	2,3	23,7	37,0
8. Робітники з обслуговування, експлуатації та контролювання за роботою технологічного устаткування, складання устаткування та машин	12,4	6,5	4,4	10,5
9. Найпростіші професії	52,2	2,7	51,7	27,0
Всього (у тис.; вакансій / осіб)	20,0	1,0	493,5	341,0

Джерело: розраховано авторами на основі даних *hh.ua*, *flagma.pl* та масиву «Зовнішня трудова міграція населення України» 2015–2017 рр.

Потенційно дані з сайтів пошуку роботи мають вказувати на основні тенденції щодо того, попит на яких працівників задовольняється за рахунок робочої сили з України. З української сторони єдиним джерелом даних щодо мігрантів є обстеження «Зовнішня трудова міграція населення України», але воно не враховує громадян, які виїхали родинами. Це ускладнює зіставлення даних сайтів і обстеження, враховуючи, що політика, наприклад, РФ, заохочує переселення на територію країни саме цілих родин, у тому числі через програми, які надають фінансову допомогу кожному члену родини [23]. Таким чином, попит на представників певних професій задовольняється у країні-реципієнті, що не знаходить відображення у даних української статистики.

Висновки. Здійснено спробу визначити потенціал використання «великих даних» для оцінки певних аспектів трудової міграції в Україні. Отримані результати свідчать, що динаміка пошукових запитів (на прикладі *Google Trends*) відображає основну переорієнтацію з роботи у Росії на роботу у Польщі, що мала місце після 2014 р. Водночас, можливість зіставлення структури вакансій із сайтів пошуку роботи і структури професій трудових мігрантів, оцінених за результатами дослідження «Зовнішня трудова міграція населення України», є проблематичною і потребує подальших досліджень. В цілому, дані, отримані з *Google Trends*, можуть бути використані для наближеної, якісної характеристики процесів, тоді як проблема масштабів трудової міграції залишається невирішеною. З урахуванням низького рівня користування Інтернетом в

Україні видається, що, принаймні у найближчі роки, дослідники мають поєднувати використання даних Інтернет із даними з традиційних джерел.

Серйозними бар'єрами для використання «великих даних» є: 1) відсутність законодавчого регулювання, яке б забезпечувало фахівцям органів державної влади, експертам та дослідникам доступ до неперсоніфікованих даних приватних компаній; 2) брак фахівців з аналізу «великих даних»; 3) брак доступного методичного та програмного забезпечення для роботи з «великими даними».

За результатами виконаного дослідження можна дійти висновку, що традиційні методи дослідження процесів трудової міграції на основі статистичних та адміністративних даних як основи для обґрунтованих управлінських рішень в Україні навряд чи поступляться «великим даним» у найближчій часовій перспективі.

ЛІТЕРАТУРА

1. Тишук Т. Велика міграція: чому ніхто в Україні не знає, скільки наших громадян виїхало за кордон [Електронний ресурс]. – 2018. – Режим доступу: <https://voxukraine.org/uk/velika-migratsiya-chomu-nihtu-v-ukrayini-ne-znae-skilki-nashih-gromadyan-viyihalo-za-kordon> (дата звернення: 18.01.2019).
2. Nowcasting (economics) [Електронний ресурс]. – Режим доступу: [https://en.wikipedia.org/wiki/Nowcasting_\(economics\)](https://en.wikipedia.org/wiki/Nowcasting_(economics)) (дата звернення: 18.01.2019).
3. Zagheni E., Garimella V., Weber I., State B. Inferring International and Internal Migration Patterns from Twitter Data [Електронний ресурс] // Proceedings of the 23 International Conference on World Wide Web. – 2014. – Режим доступу: <https://ingmarweber.de/wp-content/uploads/2014/02/Inferring-International-and-Internal-Migration-Patterns-from-Twitter-Data.pdf> (дата звернення: 18.01.2019).
4. State B., Rodriguez M., Helbing D., Zagheni E. Migration of Professionals to the U.S. Evidence from LinkedIn data [Електронний ресурс] // SocInfo: Social Informatics. – 2014. – Режим доступу: http://zagheni.net/uploads/3/1/7/9/3179747/migration_professionals_linkedin.pdf (дата звернення: 18.01.2019).
5. Blumenstock J. Inferring patterns of internal migration from mobile phone call records: evidence from Rwanda [Електронний ресурс] // Information Technology and Development. – 2012. – Vol. 18, № 2. – Режим доступу: https://www.researchgate.net/publication/233346836_Inferring_patterns_of_internal_migration_from_mobile_phone_call_records_evidence_from_Rwanda (дата звернення: 18.01.2019).
6. Novak J., Ahas R., Aasa A., Silm S. Application of mobile phone location data in mapping of commuting patterns and functional regionalization: a pilot study of Estonia [Електронний ресурс] // Journal of Maps. – 2013. – P. 10–15. – Режим доступу: <https://doi.org/10.1080/17445647.2012.762331> (дата звернення: 18.01.2019).
7. Wladyka D.K. Queries to Google Search as Predictors of Migration Flows from Latin America to Spain [Електронний ресурс] // Journal of Population and Social Studies. – 2017. – Vol. 25, No. 4. – P. 312–327. – Режим доступу: http://www.jpss.mahidol.ac.th/PDF/v25n4_paper2.pdf (дата звернення: 18.01.2019).
8. Böhme M., Gröger A., Stöhr T. Searching for a Better Life: Predicting International Migration with Online Search Keywords [Електронний ресурс]. – 2017. – Режим доступу: https://www.researchgate.net/publication/321358705_Searching_for_a_Better_Life_Predicting_International_Migration_with_Online_Search_Keywords (дата звернення: 18.01.2019).
9. Кислова О.Н. Социология в контексте вызова «больших данных»: роль интеллектуального анализа данных в становлении новых подходов к социологическим исследованиям [Електронний ресурс] // Вісник Харківського національного університету імені В.Н. Каразіна. – 2015. – № 1148. – С. 57–62. – Режим доступу: <https://periodicals.karazin.ua/ssms/article/view/4470/4041> (дата звернення: 18.01.2019).
10. Сариогло В.Г. «Великі дані» як джерело інформації та інструментарій для офіційної статистики: потенціал, проблеми, перспективи // Статистика України. – 2016. – № 4. – С. 12–19.

11. Хмелевська О.М. Репетиторство як складова тіньової освіти та можливості його оцінювання в Україні // Демографія та соціальна економіка. — 2017. — № 1 — С. 37–53. — doi: <https://doi.org/10.15407/dse2017.02.037>
12. Кучерук В.Ю., Головащенко І.О., Глушко М.В. Як Big Data (великі дані) впливають на буття людини [Електронний ресурс]. — Режим доступу: <https://ir.lib.vntu.edu.ua/bitstream/handle/123456789/17811/2555.pdf?sequence=3&isAllowed=y> (дата звернення: 18.01.2019).
13. Individuals using the Internet (% of population) / World Bank [Електронний ресурс]. — Режим доступу: <https://data.worldbank.org/indicator/IT.NET.USER.ZS> (дата звернення: 18.01.2019).
14. Letouzé E. Big Data for Development: Challenges & Opportunities [Електронний ресурс]. — 2012. — Режим доступу: <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf> (дата звернення: 18.01.2019).
15. Sangen M van der, Sandijk J. van. Mapping migrant flows with satellites [Електронний ресурс]. // Statistics Netherlands (CBS). — 2018. — Режим доступу: <https://www.cbs.nl/en-gb/corporate/2018/09/mapping-migrant-flows-with-satellites> (дата звернення: 18.01.2019).
16. Чудиновских О.С. Большие данные и статистика миграции [Електронний ресурс] // Вопросы статистики. — 2018. — № 25. — С. 48–56. — Режим доступу: https://istina.msu.ru/download/106684550/1fCOP:n_b2GlsCo1Gj_Pk4Wlseug7N_SI/ (дата звернення: 18.01.2019).
17. Google Trends [Електронний ресурс]. — Режим доступу: <https://trends.google.com/trends> (дата звернення: 18.01.2019).
18. Звіт щодо методології, організації проведення та результатів модульного вибіркового обстеження з питань трудової міграції в Україні / за ред. Е.М. Лібанової; Міжнародна організація праці; Група технічної підтримки з питань гідної праці та Бюро МОП для країн Центральної та Східної Європи. — Будапешт, 2013. — 96 с.
19. First permits issued for remunerated activities by reason, length of validity and citizenship [Електронний ресурс]. — Режим доступу: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=migr_resocc&lang=en (дата звернення: 18.01.2019).
20. Opinion Software Media. Данные работы Панели [Електронний ресурс] // Factum Group. — 2018. — Режим доступу: https://inau.ua/sites/default/files/file/1810/top-25_saytiv_u_veresni_2018.pdf (дата звернення: 18.01.2019).
21. SimilarWeb [Електронний ресурс]. — Режим доступу: <https://www.similarweb.com> (дата звернення: 18.01.2019).
22. Dexi.io [Електронний ресурс]. — Режим доступу: <https://dexi.io> (дата звернення: 18.01.2019).
23. Государственная программа по содействию добровольному переселению в Российскую Федерацию соотечественников. Официальный информационный пакет. — Режим доступу: https://xn--b1aew.xn--p1ai/mvd/structure1/Glavnie_upravlenija/guvvm/compatriots/information-pack (дата звернення: 18.01.2019).

REFERENCES

1. Tyshchuk, T. (2018). Velyka mihratsiia: chomu nikhto v Ukraini ne znaie, skilky nashykh hromadian vyikhalo za kordon [Large migration: why nobody knows in Ukraine how many of our citizens moved abroad]. *voxukraine.org*. Retrieved from <https://voxukraine.org/uk/velika-migratsiya-chomu-nihtov-ukrayini-ne-znaie-skilki-nashih-gromadyan-viyihalo-za-kordon/> [in Ukrainian].
2. Nowcasting (economics). (n.d.). *en.wikipedia.org*. Retrieved from [https://en.wikipedia.org/wiki/Nowcasting_\(economics\)](https://en.wikipedia.org/wiki/Nowcasting_(economics))
3. Zagheni, E., Garimella, V., Weber, I., & State, B. (2014). Inferring International and Internal Migration Patterns from Twitter Data. Proceedings of the 23rd International Conference on World Wide Web. *ingmarweber.de*. Retrieved from <https://ingmarweber.de/wp-content/uploads/2014/02/Inferring-International-and-Internal-Migration-Patterns-from-Twitter-Data.pdf>
4. State, B., Rodriguez, M., Helbing, D., & Zagheni, E. (2014). Migration of Professionals to the U.S. Evidence from LinkedIn data. SocInfo: Social Informatics. *zagheni.net*. Retrieved from http://zagheni.net/uploads/3/1/7/9/3179747/migration_professionals_linkedin.pdf.
5. Blumenstock, J. (2012). Inferring patterns of internal migration from mobile phone call records: evidence from Rwanda. *Information Technology and Development*, Vol. 18, 2. Retrieved from https://www.researchgate.net/publication/233346836_Inferring_patterns_of_internal_migration_from_mobile_phone_call_records_evidence_from_Rwanda

6. Novak, J., Ahas, R., Aasa, A., & Silm, S. (2013). Application of mobile phone location data in mapping of commuting patterns and functional regionalization: a pilot study of Estonia. *Journal of Maps*, 10-15. Retrieved from <https://doi.org/10.1080/17445647.2012.762331>
7. Wladyka, D.K. (2017). Queries to Google Search as Predictors of Migration Flows from Latin America to Spain. *Journal of Population and Social Studies*, 25, 312-327. Retrieved from http://www.jpss.mahidol.ac.th/PDF/v25n4_paper2.pdf
8. Böhme, M., Gröger, A., & Stöhr, T. (2017). Searching for a Better Life: Predicting International Migration with Online Search Keywords. *researchgate.net*. Retrieved from https://www.researchgate.net/publication/321358705_Searching_for_a_Better_Life_Predicting_International_Migration_with_Online_Search_Keywords
9. Kyslova, O.N. (2015). Sotsyologiya v kontekste vuzova «bolshykh dannukh»: rol intellektualnogo analiza dannukh v stanovlenii novukh podkhodov k sotsiologicheskim issledovaniyam [Sociology in the context of the challenge of “big data”: the role of data mining in the development of new approaches to sociological research]. *Visnyk Kharkivskoho natsionalnoho universytetu imeni V.N. Karazina. - The journal of V.N. Karazin Kharkiv National University*, 1148, 57-62. Retrieved from <https://periodicals.karazin.ua/ssms/article/view/4470/4041> [in Ukrainian].
10. Sarioglo, V.G. (2016). «Velyki dani» yak dzherelo informatsii ta instrumentarii dlia ofitsiinoi statystyky: potentsial, problemy, perspektyvy. [«Big data» as a source of information and tool for official statistics: potential, problems, perspectives]. *Statystyka Ukrainy - Statistics of Ukraine*, 4, 12-19 [in Ukrainian].
11. Khmelevska, O.M. (2017). Repetytorstvo yak skladova tinovoi osvity ta mozhlyvosti yoho otsiniuvannia v Ukraini [Tutoring as a component of the shadow education and approaches to its assessment in Ukraine]. *Demohrafiia ta sotsialna ekonomika - Demography and Social Economy*, 1, 37-53. doi: <https://doi.org/10.15407/dse2017.02.037> [in Ukrainian].
12. Kucheruk, V.Yu., Holovashenko, I.O., & Hlushko, M.V. (n.d.). Yak Big Data (Velyki dani) vplyvaiut na buttia liudyny [How Big data influences human life]. *ir.lib.vntu.edu.ua*. Retrieved from <https://ir.lib.vntu.edu.ua/bitstream/handle/123456789/17811/2555.pdf?sequence=3&isAllowed=y> [in Ukrainian].
13. Individuals using the Internet (% of population) (n.d.). World Bank. *data.worldbank.org*. Retrieved from <https://data.worldbank.org/indicator/IT.NET.USER.ZS>
14. Letouzé, E. (2012). Big Data for Development: Challenges & Opportunities. *unglobalpulse.org*. Retrieved from <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf>
15. Sangen, M van der., & Sandijk, J van. (2018). Mapping migrant flows with satellites. Statistics Netherlands (CBS). *cbs.nl*. Retrieved from <https://www.cbs.nl/en-gb/corporate/2018/09/mapping-migrant-flows-with-satellites>
16. Chudinovskih, O.S. (2018). Bolshie dannye i statistika migratsii [Big data and migration statistics]. *Voprosy statistiki - Questions of statistics*, 25, 48-56. Retrieved from https://istina.msu.ru/download/106684550/1f1COP:n_b2GlsCo1Gj_Pk4Wlseug7N_SI/ [in Russian]
17. Site «Google Trends» (n.d.). *trends.google.com*. Retrieved from <https://trends.google.com/trends/>
18. Libanova, E.M. (Eds.). (2013). *Zvit shchodo metodologii, orhanizatsii provedennia ta rezultativ modul-noho vybirkovoho obstezhennia z pytan trudovoi mihratsii v Ukraini [Report on the Methodology, Organization and Results of a Modular Sample Survey on Labour Migration in Ukraine]*. Decent Work Technical Support Team and Country Office for Central and Eastern Europe.
19. First permits issued for remunerated activities by reason, length of validity and citizenship. (n.d.). *appsso.eurostat.ec.europa.eu*. Retrieved from http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=migr_resocc&lang=en
20. Opinion Software Media. Dannye raboty Paneli [Opinion Software Media. Panel data] (2018). Retrieved from https://inau.ua/sites/default/files/file/1810/top-25_saytiv_u_veresni_2018.pdf.
21. Site «SimilarWeb» (n.d.). Retrieved from <https://www.similarweb.com>
22. Site «Dexi.io» (n.d.). dexi.io. Retrieved from <https://dexi.io>
23. Gosudarstvennaja programma po sodejstviju dobrovolnomu pereseleniju v Rossijskuju Federaciju sootchestvennikov. Oficialnyj informacionnyj paket [State program to promote the voluntary resettlement of compatriots to the Russian Federation. Official Information Pack] (n.d.). *xn--b1aew.xn--p1ai*. Retrieved from https://xn--b1aew.xn--p1ai/mvd/structure1/Glavnie_upravljenija/guvm/compatriots/informati-on-pack

Стаття надійшла до редакції 06.02.2019.