

BIG DATA&OPEN DATA У СВІТОВІЙ ТА ВІТЧИЗНЯНІЙ ЖУРНАЛІСТИЦІ: РОЛЬ, ЗНАЧЕННЯ, ТЕНДЕНЦІЇ

Василь Полюга,

канд. соц. наук, доцент кафедри

нових медіа та медіадизайну

Одеського національного університету імені І. І. Мечникова

e-mail: v.poljuga@onu.edu.ua

Нові формати цифрових даних — Big&Open Data — трансформують сучасну модель професійної соціальної комунікації та змінюють параметри комунікації. Поява нових форматів цифрової інформації, зокрема великих даних, передбачає нові способи її пошуку, обробки, аналізу, представлення й візуалізації в журналістиці. Ці зміни повинні знаходити відображення в текстах журналістських матеріалів. Однак є підстави припускати, що вплив великих та відкритих даних на структуру журналістських матеріалів у провідних вітчизняних засобах інформації залишається мінімальним.

Ключові слова: великі дані, відкриті дані, датафікація, дата-журналістика/журналістика метаданих, дигіталізація, конвергенція, мас-медіа, модель комунікації.

Актуальність дослідження. Постановка проблеми та її зв'язок із науковими й практичними завданнями.

Ключовою тенденцією сучасності стає датафікація соціального життя — процес подання явищ у кількісному цифровому форматі, перетворення соціальних фактів у цифрову інформацію, доступну для подальших аналітичних досліджень [8]. Теоретики й практики мас-медіа у пошуку оптимальних економічних і творчих моделей мас-

медіа позначили ряд тенденцій, які впродовж найближчих декількох років стануть визначальними у повсякденних практиках засобів масової комунікації: розширення використання новітніх технологій, поява нових екранів і порталів для доставки інформації, посилення участі аудиторії у медіакомунікаціях на основі реалізації функцій творчості й самовираження, запровадження ігрових технологій дискурсу. У зв'язку з використанням нових платформ, основними стають трансмедійні стратегії розвитку засобів масової комунікації, зокрема в журналістиці. Трансмедійність і демократизація (на основі підвищення технічної доступності) системи розподілу і, відповідно, споживання контенту означають зміну характеристик інформації як базової категорії журналістики, зокрема відкритість даних, вільний доступ до джерел інформації (Open Data).

Аналіз останніх досліджень і публікацій.

Основними джерелом відкритих та великих даних є інтернет, де масиви даних відкриті та накопичуються, де функціонують масштабні пошукові системи Google, (google.com), Bing (bing.com), Baidu (baidu.com), Yahoo! (yahoo.com), мають місце глобальні громадські проекти Вікіпедія, OpenStreetMap тощо. Інформаційні ресурси та потоки інтегруються у бази даних, що значно збільшує їхні обсяги. Останнім часом трендом стає інтеграція сервісів, створення відкритих порталів та каталогів відкритих даних, як результат подібних трансформацій, джерела структурованих даних великих обсягів стають загальнодоступними [3].

Рух відкритих даних подібний до інших «відкритих рухів», таких як відкриті вихідні коди (open source), відкритий контент (open content) і відкритий доступ (open access). Система поглядів, що стоїть за цими явищами, сформована давно, однак сам термін «відкриті дані» можна сміливо назвати молодим: своєї популярності він набув із широким розвитком Всесвітньої павутини і, особливо із

запуском урядових ініціатив з організації вільного доступу громадян до даних, пов'язаних із прийняттям управлінських рішень. Для ближчого ознайомлення з явищем відкритих даних і розуміння їхнього призначення та практичної цінності, варто звернутися до його визначень і нормативно-правової бази. Так, у вільній енциклопедії «Вікіпедія» міститься наступне визначення: «Відкриті дані — це концепція, за якою певні дані мають бути вільними для використання та розповсюдження будь-якою особою з будь-якою метою» [2]. В Україні єдиний державний вебпортал відкритих даних створено на вимогу Закону України «Про доступ до публічної інформації» та постанови Кабінету Міністрів України від 21 жовтня 2015 р. № 835 «Про затвердження Положення про набори даних, які підлягають оприлюдненню у формі відкритих даних». Відкриті дані наділені іманентними ознаками великих даних: вони цифрові, машинозчитувані, але при цьому не обов'язково мають значний обсяг. Великі дані те ж саме джерело визначає як «Великі дані (англ. Big Data) в інформаційних технологіях — набори інформації (як структурованої, так і неструктурованої) настільки великих розмірів, що традиційні способи та підходи (здебільшого засновані на рішеннях класу бізнесової аналітики та системах управління базами даних) не можуть бути застосовані до них» [1].

Виклад основного матеріалу дослідження.

Завдяки поширенню відкритих даних та датафікації, у світовій практиці журналістики останнього десятиріччя формується низка тенденцій щодо використання даних, зокрема вплив нових технологій на процес створення й тематику журналістського матеріалу, виникнення нових можливостей доставки та розповсюдження інформації, творчості та самовираження, оформляється новий напрям — журналістика даних (data journalism). Сучасна журналістика даних спирається на традиції

«комп'ютерної», або «обчислювальної» журналістики (computer-assisted reporting — CAR), і «прецизійної журналістики» (precision-journalism), заснованої на застосуванні наукових методів збору й аналізу даних, яка останнім часом набула поширення у провідних зарубіжних мас-медіа. З'являється технологічний тренд і концепція, що представляють собою серію підходів, інструментів і методів обробки структурованих і неструктурованих машинозчитуваних даних. Система включає технології аналізу та використання, методи пошуку необхідної інформації у великих масивах і наразі є однією з найбільш динамічних. Варто також відзначити, що ініціатива щодо розкриття великої частини даних є фундаментом відкритого державного управління у багатьох країнах світу. За допомогою такого підходу забезпечується прозорість роботи органів державної влади, формується база для контролю, створюються нові послуги для громадян і бізнесу. За висновками незалежних міжнародних досліджень, відкриті дані — один із найбільш динамічних і перспективних стратегічних напрямків для урядів різних країн як з розвинутою економікою, так і з економікою, що ще розвивається [6]. Це положення підтверджується й досвідом багатьох країн і міжнародних організацій, які в своїй роботі активно використовують відкриті дані. Сьогодні лідерами в зазначеному напрямку виступають:

- Австралія з масштабним проєктом із відкриття картографічних даних і систем позиціонування;
- Великобританія, де активно публікуються відкриті дані державних витрат, ринку праці та нерухомості;
- США, країна відома безліччю «відкритих» проєктів та активно стимулює некомерційні ініціативи в області відкритих даних.

Варто наголосити, що список країн, де відкриті дані впроваджуються й використовуються, досить великий. Адже сьогодні

значна кількість урядів світу запустила офіційні портали відкритих даних. Цей процес почався в 2010 році та досяг своєї пікової точки в 2012. Підкреслимо, що це явище супроводжувалося громадським підйомом, пов'язаним із риторикою відкритості та підзвітності, і відбувалося в контексті перспективних прогнозів, які віщували помітні поліпшення в області економіки за порівняно малих витрат: нові продукти й послуги, нові робочі місця, оптимізація роботи державних органів за рахунок якісного поліпшення їхньої взаємодії. Тим часом, втілення амбітного проєкту в життя, в більшості випадків, миттєвих результатів не давало, і на зміну життєствердній абстракції прийшли цілком закономірні питання про те, як саме домогтися заявленого ефекту на практиці. Адже, по суті, очікувалося, що варто тільки урядам країн розкрити дані, як розробники, журналісти й громадські активісти неодмінно почнуть використовувати їх у своїй роботі. Насправді ж все це сталося, але далеко не в тих масштабах, що планувалися, перед зацікавленими урядами закономірно постало питання про популяризацію відкритих даних серед населення [6].

З проблемою слабого попиту на відкриті дані стикаються багато урядів різних країн. Супутні обставини цієї ситуації багато в чому схожі: недолік інформованості, нерозвинена культура публікації даних (від країни до країни це проявляється по-різному), технічні перешкоди у вигляді цифрового розриву й недолік необхідних технічних компетенцій як у громадян, так і у влади. Ще один загальний і суттєвий фактор — це відсутність координації між попитом і пропозицією: потенційно затребувані й важливі дані значною мірою залишаються нерозкритими [6].

Проблема відкритих даних особливо актуальна для журналістики. Важливо співпрацювати, обмінюватися досвідом, джерелами й враженнями, щоб бути в курсі того, що відбувається, й орієнтуватися в обстановці. Використання відкритих великих обсягів даних стає

все більш значущим напрямком роботи для журналістів всього світу. Поява повністю відкритих джерел даних, а також безкоштовних інструментів для їхнього очищення, обробки й візуалізації дозволяє робити прогноз про збільшення кількості журналістських матеріалів на основі аналітики великих обсягів даних у найближчі роки.

Варто зазначити, що актуальні дослідження журналістики даних, незважаючи на істотний обсяг і велику географію, поки що мають описовий, систематизуючий характер, і лише на початку 2018 року в науковій літературі починають говорити про те, що накопичений емпіричний матеріал дозволяє аналізувати журналістику даних як явище й процес [1, 4, 5]. Більшість авторів підкреслюють, що журналістика даних і досі перебуває на етапі становлення, вивчають її, ґрунтуючись на методах спостереження й інтерв'ювання. Комплексні дослідження джерел інформації публікацій журналістики даних і досі відсутні. У провідних виданнях зростає кількість матеріалів, які базуються на відкритих даних та публікуються на регулярній основі в якості окремих проєктів або в спеціалізованих розділах. Однак, незважаючи на поширеність, це явище викликає жваві дискусії як в середовищі журналістів-практиків, так і серед представників академічної спільноти. Практики журналістики даних достатньою мірою не систематизовані та комплексно не відображені в науковій літературі. Наявні дослідження в зазначеній області знаходяться на стику різних наукових напрямків і концепцій та мають міждисциплінарний характер. Однак уже зараз багато ЗМІ починають працювати з цифрами й базами даних, буквально знаходячи в них історії для публікацій. Можна констатувати, що сьогодні зароджується співтовариство людей, які називають себе журналістами даних, які значно відрізняються від спільноти експертів в області комп'ютерної журналістики. Це нове співтовариство володіє більш інноваційними способами надання даних, створення

сюжетів та подання підсумкових репортажів, що поєднують у собі глибоко аналітичний і критичний підходи спільноти комп'ютерної та візуальної журналістики.

Особливістю нового формату є грамотний аналіз даних як ще один спосіб критичного дослідження світу і, по суті, інструмент, що допомагає отримати й представити громадськості важливі відомості, які могли залишитися непоміченими у величезних масивах інформації про сучасну реальність. У журналістиці даних цифри — це головне, вони можуть говорити самі за себе, розповідати історію. Текст в таких матеріалах практично відсутній, натомість використовується наочний інтерактивний інструмент, який легко вбудовується в усі статті, пов'язані з темою, для довідки. Головний принцип — матеріал будується не навколо новинного приводу або явища, а навколо статистики, цифр, зведень, звітів та іншої довідкової інформації. Про журналістику даних як про сформований жанр, на жаль, поки що можна говорити тільки в Європі та США, де вона вже отримала достатнє поширення й визнання. Стимулом до розвитку цього явища та його оформлення в самостійний жанр послужив розвиток інтернету й інформаційних технологій, що вмщують в собі великі обсяги статистичних та інших даних. Для дата-журналістів основним джерелом інформації виступають бази даних, накопичені в різних державних інформаційних системах. Наприклад, державні контракти, бюджет, статистика, індикатори, звіти. Найбільш процвітають у роботі з подібною відкритою державною інформацією американські газети, зокрема всім відомі «The New York Times», «Chicago Tribune», «Los Angeles Times» тощо. Ці видання не просто аналізують отримані дані, але й ведуть постійні успішні спецпроекти на їхній основі.

За результатами аналізу українських журналістських видань, сьогодні, на жаль, ми не зустріли жодного, яке могло б заявити

журналістику даних в якості одного зі своїх основних жанрів. Хоча варто зазначити, що досить активно пробують себе в цьому напрямі «Bird in Flight», «Liga.net» та «KyivPost». Є сильні проєкти журналістики даних у низці недержавних громадських організацій: «Чесно», «Опора», «CifraGroup», «infolight.org.ua» тощо. Приклади такої журналістської роботи можна нерідко зустріти на сторінках друкованих видань і в Інтернеті. Ось тільки самі журналісти не звикли уналежнювати подібні матеріали до якогось конкретного жанру і вже, тим більше, до дата-журналістики.

Висновки та перспективи дослідження.

Таким чином, незважаючи на актуальність і зростання популярності в аудиторії, в Україні журналістика даних залишається у фокусі небагатьох ЗМІ. На нашу думку, основна проблема, пов'язана з реалізацією нових віянь у медіасфері, полягає в нестачі кадрів. Наразі курси з журналістики даних викладають в основному в зарубіжних університетах. У більшості університетів нашої країни подібних курсів немає (окремі курси з журналістики даних пропонує магістерська програма Могилянської школи журналістики та ще декілька факультетів журналістики у провідних вишах), журналістів не вчать працювати з великими масивами інформації, і вони не володіють необхідним набором технічних навичок. А між тим, журналіст, який працює з даними, повинен поєднувати в собі чимало технічних і гуманітарних знань.

Разом з тим, основа для формування журналістики даних як жанру давно готова: створено законодавчу базу, є дані, є розробники, які цілком вміють їх візуалізувати, але не вистачає системного підходу з боку видань. Можливо, криза в галузі журналістики, а також зростаюча конкуренція з боку інтернет-компаній, соціальних мереж і розважальних брендів, змусить видавців звернути увагу на цей новий перспективний жанр і почати активно його впроваджувати.

Поява нових форматів цифрової інформації, зокрема великих даних, передбачає нові способи її пошуку, обробки, аналізу, уявлення та візуалізації в журналістиці. Ці зміни повинні знаходити відображення в текстах журналістських матеріалів, підготовлених у форматі журналістики даних. Разом з тим, є підстави припускати, що сьогодні у вітчизняній журналістиці вплив використання даних на джерела інформації матеріалів провідних журналістських організацій світу залишився мінімальним. Але, незважаючи на цілу низку невирішених питань, відкриті дані вже відіграли важливу роль у житті суспільства, а подальший розвиток цього напрямку полегшить його застосування.

Література:

1. Великі дані. Матеріал з Вікіпедії — вільної енциклопедії. 2020. URL: https://uk.wikipedia.org/wiki/Великі_дані (дата звернення: 29.11.2020).
2. Відкриті дані. Матеріал з Вікіпедії — вільної енциклопедії. 2020. URL: https://uk.wikipedia.org/wiki/Відкриті_дані (дата звернення: 29.11.2020).
3. Кірей К.О. Розвиток і трансформація поняття «Big Data». *Вісник Черкаського державного технологічного університету*. Серія: Технічні науки. 2019. №1. С. 33–40. URL: http://nbuv.gov.ua/UJRN/Vchdtu_2019_1_7 (дата звернення 26.11.2020).
4. Шилина М. Г. Big Data, Open Data, Linked Data, метаданные в PR: актуальные модели трансформации теории и практики. *Медиаскоп*. Вып. №1. Москва, 2014. URL: <http://www.mediascope.ru/node/1486> (дата звернення: 26.11.2020).
5. Anderson B. & Borges-Rey E. Encoding the UX: User Interface as a Site of Encounter between Data Journalists and Their Constructed Audiences. *Digital Journalism*, 2019, 7:9, pp. 1253–1269.

6. Borges-Rey E. and Stalph F. A random walk through data + journalism: assessing the data journalism ecosystem. *Paper presented on NODA-2018 (The Nordic Data Journalism Conference 15–17 Mar. 2018)*. Sodertorn University : Sweden, 2018.
7. Ester Appelgren, Carl-Gustav Linden & Arjen van Dalen. Data Journalism Research: Studying a Maturing Field across Journalistic Cultures, Media Markets and Political Environments. *Digital Journalism*. 2019, 7:9, pp. 1191–1199.
8. Mayer-Schönberger V., Cukier K., Big Data: A Revolution That Will Transform How We Live, Work, and Think. *American Journal of Epidemiology*, Volume 179, Issue 9, 1 May 2014, pp. 1143–1144. URL: <https://doi.org/10.1093/aje/kwu085>.

УДК 007:004.912:311.213

DOI: 10.18524/2308-3255.2020.26.225317

Big Data&Open Data в журналистике: роль, значение, тенденции

Василий Полюга,

канд. соц. наук, доцент кафедры

новых медиа и медиадизайна

Одесского национального университета имени И. И. Мечникова

e-mail: v.poljuga@onu.edu.ua

Новые форматы цифровых данных — Big & Open Data — трансформируют современную модель профессионального социального общения и меняют параметры общения. Появление новых форматов цифровой информации, включая большие данные, открывает новые способы поиска, обработки, анализа, представления и визуализации в журналистике. Эти изменения должны быть отражены в текстах публицистических материалов. Однако есть основания полагать, что влияние больших и открытых данных на

структуру журналистских материалов в ведущих отечественных средствах массовой информации остается минимальным.

Ключевые слова: большие данные, открытые данные, датафикация, дата-журналистика/журналистика метаданных, дигитализация, конвергенция, масс-медиа, модель коммун

UDC 007:004.912:311.213

DOI: 10.18524/2308-3255.2020.26.225317

Big Data&Open Data in journalism: role, meaning, trends

Wasil Poljuga,

PhD (Sociology), Associate Professor
of the New Media and Media Design Chair,
Odessa I. I. Mechnikov National University
e-mail: v.poljuga@onu.edu.ua

New digital data formats — Big & Open Data — are transforming the modern model of professional social communication and changing the parameters of communication. The emergence of new formats of digital information, including big data, provides new ways to search, process, analyze, present and visualize it in journalism. These changes should be reflected in the texts of journalistic materials. However, there is reason to believe that the impact of large and open data on the structure of journalistic materials in the leading domestic media remains minimal.

Data journalism is still in its infancy, being studied based on observation and interviewing techniques. Comprehensive studies of data sources of data journalism publications are still lacking. In leading publications, the number of materials based on open data and published on a regular basis as separate projects or in specialized sections is growing. However, despite its prevalence, this phenomenon is still controversial both among practicing journalists and academics. The practices of data journalism, which are not sufficiently systematized, are not comprehensively reflected in the scientific literature.

Existing research in this area is at the intersection of different scientific fields and concepts and is interdisciplinary. However, many media outlets are already starting to work with numbers and databases, literally finding stories in them for publication. It can be stated that today a community of people is emerging who call themselves data journalists, which is significantly different from the community of experts in the field of computer journalism. This new community has more innovative ways of providing data, storytelling, and final reporting, combining the deeply analytical and critical approaches of the computer and visual journalism community.

Keywords: big data, open data, dating, data journalism/metadata journalism, digitization, convergence, mass media, communication model.

References:

1. Big data. (2020). Velyki dani. Material z Vikipedii — vilnoi entsyklopedii. [From Wikipedia, the free encyclopedia]. URL: <https://uk.wikipedia.org/wiki/> (accessed: 29.11.2020).
2. Open data. (2020). Vidkryti dani. Material z Vikipedii — vilnoi entsyklopedii. [From Wikipedia, the free encyclopedia]. URL: <https://uk.wikipedia.org/wiki/> (accessed: 29.11.2020).
3. Kirey, E. A. (2019). Rozvytok i transformatsiia poniattia «Big Data». [Development and transformation of big data concept]. *Visnyk Cherkaskoho derzhavnoho tekhnolohichnoho universytetu. Seriya: Tekhnichni nauky*. №1. pp. 33–40. URL: http://nbuv.gov.ua/UJRN/Vchdtu_2019_1_7 (accessed: 26.11.2020).
4. Shilina, M. G. (2014). Big Data, Open Data, Linked Data, metadannye v PR: aktual'nye modeli transformacii teorii i praktiki. [Big Data, Open Data, Linked Data, metadata in PR: current models of transformation of theory and practice]. *Mediakop*. Issue №1. Moscow. URL: <http://www.mediascope.ru/node/1486> (accessed: 26.11.2020).
5. Anderson, B. & Borges-Rey, E. (2019). Encoding the UX: User Interface as a Site of Encounter between Data Journalists and Their Constructed Audiences, *Digital Journalism*, 7:9, pp. 1253–1269.
6. Borges-Rey, E. and Stalph, F. (2018) A random walk through data +

- journalism: assessing the data journalism ecosystem. *Paper presented on NODA-2018 (The Nordic Data Journalism Conference)*. Sodertorn University: Sweden.
7. Ester Appelgren, Carl-Gustav Linden & Arjen van Dalen (2019). Data Journalism Research: Studying a Maturing Field across Journalistic Cultures, Media Markets and Political Environments, *Digital Journalism*, 7:9, pp. 1191–1199.
 8. Mayer-Schönberger, V., Cukier, K. (2014). Big Data: A Revolution That Will Transform How We Live, Work, and Think. *American Journal of Epidemiology*, Vol. 179, Issue 9, pp. 1143–1144, URL: <https://doi.org/10.1093/aje/kwu085> (accessed: 26.11.2020).