

ANDRZEJ RADOMSKI

ANALITYKA KULTUROWA

NOWY PARADYGMAT BADAŃ W HUMANISTYCE

ANDRZEJ RADOMSKI

Profesor w Instytucie Nauk o Kulturze Uniwersytetu Marii Curie-Skłodowskiej w Lublinie, zajmuje się humanistyką cyfrową, teorią i metodologią badania kultury cyfrowej, wizualizacją wiedzy i podstawami *data science*. Ostatnio opublikował *Humanistyka w świecie informacjonalizmu* (2015). ORCID: 0000-0002-1735-605X.

Rewolucja cyfrowa, będąca konsekwencją współczesnych technologii ICT¹, dokonała olbrzymich zmian we wszystkich praktykach kulturowych. Powstały nowe zawody, nowe sposoby produkcji, nowe formy aktywności, rozrywki, nauczania, kanały komunikacji czy uprawiania polityki. Zmieniła się również praktyka naukowa.

Codzienna działalność ludzi w ramach różnych praktyk kulturowych coraz częściej polega na „produkowaniu” danych czy informacji. Zarówno w środowisku fizycznym, jak i przede wszystkim wirtualnym umieszczają oni ogromne ilości postów, komentarzy, zdjęć, filmów czy projektów, a naukowcy coraz częściej publikują w internecie wyniki swoich badań. Co więcej, każde nasze kliknięcie w oknie przeglądarki czy wyszukiwarki też jest określoną informacją czy zestawem danych, które potem mogą być wykorzystane do różnych celów.

Na oznaczenie tych wszystkich cyfrowych śladów, które zostawiamy w rezultacie naszej profesjonalnej, jak i prywatnej działalności w poszczególnych praktykach współczesnego społeczeństwa, stworzono specjalny termin: *big data*. Humanistyka i większość dyscyplin społecznych (za wyjątkiem ekonomii) nie zwracały na początku szczególnej uwagi na zjawisko *big data* ani na konieczność zmierzenia się z problemami teoretyczno-metodologicznymi badania dużych

1 Skróć od *information and communication technology*.

kolekcji danych. Wyjątkiem na tym polu była działalność Lva Manovicha², który już w pierwszych latach naszego stulecia dostrzegł potrzebę zmiany przedmiotu badań humanistyki, dyscyplin społecznych i nauk o sztuce, a także wypracowania nowych narzędzi badawczych do analizy kulturowych *big data*.

W roku 2007 powołał on do życia Software Studies Initiative³ – rodzaj humanistycznego laboratorium, w ramach którego rozpoczął wraz z zespołem szeroko zakrojone badania nad cyfrowymi *big data* (głównie nad kolekcjami obrazów i filmami). Posiadając dzięki wykształceniu kompetencje z jednej strony artystyczne, a z drugiej informatyczne, zapoczątkował nowatorskie analizowanie współczesnej produkcji medialnej, wykorzystując do tego celu narzędzia ICT.

Najnowsza książka Lva Manovicha *Cultural Analytics*⁴ stanowi pewnego rodzaju podsumowanie dotychczasowej działalności naukowej tego badacza. Jest ona równocześnie rodzajem monografii o charakterze teoretyczno-metodologicznym, zarysowującej koncepcję nowego, jak już wszystko na to wskazuje, paradygmatu w humanistyce i naukach o sztuce, który ten rosyjski badacz nazwał analityką kulturową. Punktem wyjścia w rozważaniach autora jest stwierdzenie, że musimy na nowo odkryć, co to znaczy studiować kulturę⁵.

Analityka kulturowa jest autorskim pomysłem Lva Manovicha. W omawianej książce pisze:

czytałem istniejące prace o analizie i wizualizacji wielkich zestawów danych kulturowych i widziałem, że dzieje się to w wielu dziedzinach i że tworzy się nowy paradygmat badawczy, ale nie miał on własnej nazwy. Czułem potrzebę ustanowienia nowej kategorii, która odnosiłaby się do obliczeniowych analiz wzorów i trendów we współczesnej kulturze cyfrowej (jako opozycji do tylko historycznej kultury) i która obejmowałaby analizę różnych gatunków mediów (jako opozycji do tylko tekstu). Gdy wiosną 2007 roku powołaliśmy do życia nasze własne laboratorium do tego typu badań, wymyśliłem termin „analityka kulturowa”⁶.

Manovich w omawianej monografii zaplanował realizację dwóch głównych celów: praktycznego – czyli zademonstrowanie użycia metod nauk komputerowych, wizualizacji informacji i sztuki mediów do badania i analizy różnych rodzajów współczesnej produkcji medialnej, oraz teoretycznego – czyli przedyskutowanie zarówno możliwości, jak i ograniczeń stosowania wspomnianych metod do badania dużych zbiorów danych.

Projekt analityki kulturowej autorstwa Manovicha wyłożony w *Cultural Analytics* sytuuje się na przecięciu *data science*, *media studies* i *digital culture studies*⁷,

2 Od 2012 roku pracuje w City University of New York.

3 Zob. więcej informacji na temat tego Labu na stronie: <http://lab.softwarestudies.com/> (17 kwietnia 2021).

4 L. Manovich, *Cultural Analytics*, The MIT Press, Cambridge MA 2020. Wcześniej napisał: *The Language of New Media* (2001), *Software Culture* (2010), *Software Takes Command* (2013).

5 Tamże, s. 18.

6 Tamże, s. 7. Jeśli nie zaznaczono inaczej, cytaty w tłumaczeniu autora artykułu.

7 Jak dotychczas nie ma polskich odpowiedników tych anglojęzycznych nazw.

a więc stosunkowo młodych dyscyplin, które zajmują się badaniem współczesnej kultury cyfrowej z użyciem metod komputerowych. Jak pisze we wstępie autor, studiowanie jego propozycji nie wymaga znajomości podstaw *data science*, programowania, statystyki ani matematyki. Adresatami książki są badacze i studenci sztuki, humanistyki, nauk społecznych, mediów, analitycy danych i nauk komputerowych. Jego książka, jak zapewnia, jest przyjaznym, nieużywającym technicznego języka, wprowadzeniem do problematyki analizy danych kulturowych za pomocą algorytmów⁸.

W swym projekcie analityki kulturowej Manovich dystansuje się nieco od innego prężnego nurtu we współczesnej nauce, jakim jest humanistyka cyfrowa. Termin ten pojawił się w nauce na początku XXI stulecia, a głównym celem humanistów cyfrowych, jak zauważył Manovich, była analiza tekstów historycznych (na przykład z literatury). Tymczasem dla niego równie interesujące było to, jak globalizacja i powstanie nowych technologii komunikacyjnych zmieniły geografie kultury, czy możemy znaleźć określone kulturowe trendy w małych miastach leżących z dala od wielkich metropolii, jak współczesna kultura różnicuje rozwój – czy ulega on przyspieszeniu czy spowolnieniu⁹. Tego typu pytania stały się swoistym bodźcem do poszukiwań innych metod badania kultury niż te, które oferowała, jego zdaniem, humanistyka cyfrowa.

W omawianej książce Manovich wskazuje na pięć problemów, wokół których projektował swój paradygmat. Pierwszy można sformułować następująco: zobaczyć na podstawie codziennej twórczości, jak setki milionów ludzi na całym świecie kreują, postrzegają i wartościują różne wytwory. Wiąże się on z zaproponowaniem nowej/nowych koncepcji, za pomocą której/których można uchwycić tempo, zróżnicowanie i łączność między zjawiskami współczesnej globalnej kultury cyfrowej. Te nowe koncepcje, twierdzi Manovich, powinny nie tylko mieć teoretyczny charakter, ale również umożliwiać nam określanie wymiarów cyfrowej kultury i porównywać style, gusta, wyobrażenia i kulturowe zachowania w wielu miejscach i sieciach. Drugim celem było znalezienie nowego języka do opisu kulturowych artefaktów, doświadczeń i ich dynamiki. Człowiek bowiem nie jest w stanie uchwycić tysięcy, milionów czy bilionów artefaktów za pomocą dotychczasowych narzędzi metodologicznych. Niezbędne okazuje się użycie metod numerycznych i wizualizacji. Trzeci problem wiąże się ze zwróceniem szczególnej uwagi na media wizualne. Manovich uważa bowiem, że wiele interesujących zjawisk we współczesnej kulturze można dostrzec, badając media za pomocą metod komputerowych. Czwarty problem wiąże się z pytaniem: czy możemy badać wielkie zbiory danych medialnych (jak i innych) bez użycia metod ilościowych? Na przykład jak możemy zobaczyć bilion zdjęć? Ostatnim zagadnieniem jest dyskusja nad ograniczeniami metod *data science* do badań *big data*¹⁰.

8 L. Manovich, *Cultural Analytics*, dz. cyt., s. 2.

9 Tamże, s. 3.

10 Tamże, s. 9–12.

Wymienione problemy Manovich rozbija na 12 szczegółowych pytań:

- jak praca z dużymi kolekcjami danych może pomóc nam w przezwyciężeniu pewnych stereotypów na temat kultury/kultur?
- jakie są fundamentalne nowe sposoby rozumienia oraz badania wizualnych i medialnych kultur?
- w jaki sposób możemy badać olbrzymie zbiory wizualnych kolekcji obejmujących biliony zdjęć i wideo?
- jak można połączyć komputerową analizę mediów z jakościowymi metodami i teoriami *media studies*?
- jakich teoretycznych koncepcji i modeli potrzebujemy, aby zajmować się olbrzymią skalą tempa i zmienności treści generowanych przez użytkowników online?
- jak możemy analizować i wizualizować różnorodność współczesnej globalnej kultury cyfrowej, biorąc pod uwagę wielką liczbę twórców i obiektów, które oni tworzą?
- jaka będzie nowa nauka o kulturze zajmująca się *big data* i wykorzystująca metody obliczeniowe, jakie będą jej ograniczenia?
- czy da się zdefiniować ogólne ilościowe miary kulturowej zmienności, różnorodności, zmian w czasie, które będą wspólne dla różnych typów mediów i różnych typów okresów i kultur, w szczególności naszych czasów?
- jak możemy za pomocą metod statystycznych i tych charakterystycznych dla *data science* analizować szczególne i niepowtarzalne artefakty bądź doświadczenia?
- czy można opisać kultury jako rozkłady i kombinacje statystycznych elementów, tematów i strategii?
- czy powinniśmy całość wielkich kulturowych danych zredukować do mniejszej liczby struktur, czyli najczęściej pojawiających się idei, stylów, wzorów i zachowań?

Książka *Cultural Analytics* składa się z 10 krótkich rozdziałów podzielonych na trzy części. W części pierwszej (badanie kultury w dużej skali) Manovich omawia nową ideę badań nad kulturą, która zakłada szerokie użycie metod komputerowych i kwantyfikacji danych. Jest rzeczą pożądaną, twierdzi, aby studia nad mediami, kulturą cyfrową i w ogóle badania humanistyczne zaczęły używać kwantyfikacji, modeli matematycznych i wizualizacji danych. Kluczowym powodem jest nieadekwatność dotychczasowych metod wobec skali dzisiejszej produkcji kulturalnej, która zbliżyła się do skali charakterystycznej dla zjawisk fizycznych lub biologicznych (jeśli chodzi o wielkość). Wcześniej, zauważa badacz, teoretycy kultury i historycy tworzyli teorie i historie oparte na małych zbiorach danych (tak zwanych próbkach reprezentacyjnych). Posługiwanie się jednak próbkami o małej skali w badaniu dzisiejszej kultury może przynieść zafałszowane obrazy świata XXI wieku.

Już w początkach tego stulecia zaczęły się kształtować, jak zauważa Manovich, trzy nurty, które używają metod komputerowych do badania danych o zjawiskach kulturowych. Są to: humanistyka komputerowa i jakościowe nauki społeczne, ilościowe nauki społeczne i nauki komputerowe. Mają jednak one odmienne cele

i stawiają różne pytania badawcze¹¹. Analityka kulturowa próbuje łączyć owe trzy podejścia w jeden paradygmat¹².

Przedmiotem zainteresowania kulturowych analityków (w rozumieniu Manovicha) są przede wszystkim następujące zbiory danych: zapisy ludzkiej aktywności w sieci (cyfrowe ślady), typu odwiedzane strony, posty czy udostępnione materiały; zapisy fizycznych zachowań, ich lokalizacja geograficzna, data i czas zamieszczania postów w sieci; zawartość mediów tworzona przez kompanie medialne: piosenki, wideo, książki i filmy; zawartość mediów tworzona przez użytkowników: posty, dyskusje, zdjęcia i wideo¹³.

Do analizy wymienionych danych niezbędne są odpowiednie narzędzia i metody. Manovich zaleca te z obszaru *data mining*, uczenia maszynowego i sztucznej inteligencji (głębokiego uczenia maszynowego). Często są one oparte na rozwiązaniach *open source* i stąd łatwo dostępne dla akademików (napisane w Python i R). Szczególnie interesujące wydają się takie rozwiązania jak: przetwarzanie obrazu (*computer vision*), przetwarzanie języka naturalnego i lingwistyka obliczeniowa¹⁴.

W części drugiej (reprezentowanie kultury jako danych) Manovich rozważa nowe wyzwania związane z badaniem współczesnej cyfrowej kultury. Na czoło, według niego, wysuwają się dwa problemy: jakie typy danych powinny być uwzględnione przy próbach poznania kultury i jak wprowadzić nowy język opisu.

Klasyccy humaniści pracowali głównie na kanonie dzieł literackich czy źródeł pisanych, ale zrozumienie współczesnych zjawisk kulturowych wymaga poszerzenia zbioru danych. Manovich wskazuje na cztery obszary: media tworzone przez profesjonalistów i użytkowników sieci społecznościowych; nasze zachowania – zarówno w środowisku cyfrowym, jak i fizycznym, po których zostają tak zwane cyfrowe ślady; interakcje, w jakie wchodzimy (na przykład w sieciach społecznościowych) oraz wydarzenia kulturalne, miejsca i organizacje, które możemy analizować za pomocą algorytmów, i to w dużej skali¹⁵. Ten ostatni aspekt, czyli automatyczna analiza zjawisk w dużej skali, jest tutaj kluczowa. Píše on:

w drugim artykule o analityce kulturowej napisanym w marcu 2009 roku opisywałem moje doświadczenia z próbą użycia istniejących kolekcji cyfrowych obrazów dostępnych w tym czasie. Zainteresowałem się następującym zagadnieniem: co ludzie malowali na świecie w roku 1930 – oprócz małej liczby przedstawicieli ruchu modernistycznego obejmującego co najwyżej 200 artystów (pracujących w Paryżu, Amsterdamie, Berlinie i kilku innych miastach), którzy są teraz uwzględnieni w kanonie historycznym sztuki zachodniej? Chciałbym mieć obrazy z dziesiątek tysięcy małych muzeów sztuki w małych miastach na całym świecie, ale takich danych wtedy nie było i nie ma ich teraz. Pomyślałem jednak, że mogę z łatwością znaleźć kilka tysięcy obrazów online artystów ważnych dla danej naro-

11 Na przykład ilościowe nauki społeczne stawiają na opisy i analizy statystyczne, *computer science* szukają ogólnych praw, a humaniści interpretują wybrane teksty – najczęściej z tak zwanego kanonu.

12 L. Manovich, *Cultural Analytics*, dz. cyt., s. 45.

13 Tamże, s. 57.

14 Tamże, s. 58.

15 Tamże, s. 89.

dowości, którzy weszli do kanonu historii w ich krajach, chociaż nie byli częścią modernistycznego kanonu nakreślonego przez narrację historyczną sztuki zachodniej¹⁶.

Ten nieco długi cytat pokazuje być może kluczowy problem współczesnej humanistyki (i nie tylko zresztą jej), a mianowicie ogrom danych niezbędnych i wystarczających zarazem do ukazania całego zróżnicowania rzeczywistości. Manovich konsekwentnie utrzymuje, że ani kanoniczne źródła, ani statystyczne próbki nie wystarczą do ukazania złożoności świata kultury/kultur. Nie oznacza to jednak konieczności wykorzystania wszystkich dostępnych informacji na dany temat. Pokazuje to na przykładzie impresjonizmu. Zwykle o tym nurcie w malarstwie pisze się, uwzględniając 140 prac, które – jak twierdzi – stanowią ledwie 1 proc. obrazów zaliczanych do twórczości impresjonistycznej. A jeśli weźmiemy pod uwagę 5000 prac (38 proc.), to jak zmieni się nasze postrzeganie historii sztuki? Zatem, możemy zapytać, która wersja impresjonizmu jest poprawna: ta kanoniczna czy ta uwzględniająca wspomniane 5000 obrazów? Na tak postawione pytanie nie ma jednoznacznej odpowiedzi. Ważne jest, że mamy alternatywę¹⁷.

Dane rozumie Manovich jako obiekty wyposażone w odpowiednie cechy (właściwości, atrybuty czy metadane), które mogą być poddane zautomatyzowanej obróbce¹⁸. W związku z tym kolejnym problemem, jaki powstaje, jest kwestia języka używanego do analizy tego typu danych. Mają one bowiem charakter cyfrowy (dotyczy to również danych historycznych poddanych digitalizacji). Wcześniejsze badania w naukach humanistycznych czy społecznych odnosiły się do wytworów kulturowych mających najczęściej charakter językowy i były interpretowane także za pomocą kategorii języka naturalnego. Tymczasem wytwory (a właściwie dane) produkowane przez społeczeństwa XXI wieku mają już w przeważającej mierze charakter cyfrowy. Są one ustrukturyzowane i nieustrukturyzowane, trójwymiarowe, dyskretne i ciągłe. Wymagają zatem innego języka analizy. Dzisiaj, zauważa Manovich, to algorytmy automatycznie mierzą kompozycję, kolory i strukturę obrazów, a także wykrywają ich gatunki¹⁹. Zmieniają się zatem kategorie teoretyczne i ich punkt odniesienia (na przykład piksele). Tak więc to numeryczne pomiary kulturowych artefaktów, doświadczeń i procesów dają nam, według badacza, nowy język do opisu i dyskusji o kulturze²⁰.

W ostatniej części omawianej tu pozycji (badanie danych kulturowych) Manovich koncentruje się na badaniu mediów oraz wizualizacji. Autor *Cultural Analytics* zauważa, że na początku naszego wieku tylko w kilku dziedzinach istniało zainteresowanie wizualizacją wiedzy. W następnej dekadzie sytuacja uległa radykalnej zmianie i ukształtowały się nowe paradygmaty: wizualizacja informacji, wizualizacja danych oraz wizualizacja mediów. Wizualizację informacji można, jego zdaniem, uważać za mapowanie między danymi a ich wizualną

16 Tamże, s. 95–96.

17 Tamże, s. 119.

18 Tamże, s. 122–123.

19 Tamże, s. 125.

20 Tamże, s. 154.

reprezentacją²¹. Z kolei wizualizacja mediów jest dla niego tworzeniem nowej reprezentacji wizualnej z obiektu wizualnego, z kolekcji lub z części tego obiektu²². Podstawą tak rozumianej wizualizacji są redukcja danych i uprzywilejowanie zmiennych przestrzennych.

Posługiwanie się metodami opartymi na wizualizacji pozwala szybko wykryć wszystkie interesujące nas trendy i wzory, co ważne – w sposób zautomatyzowany. W przypadku badania mediów (filmów, zdjęć i innej grafiki) wizualizacja może nam ukazać określone cechy takiej czy innej produkcji medialnej. Manovich, odwołując się do własnych badań, wskazuje na takie cechy, jak: dominujące kolory (średnia i mediana), ich jasność, odcienie i stopnie nasycenia, kontrast czy kształty na obrazie²³.

Obecnie, jak wskazuje autor omawianej książki, mamy do dyspozycji szereg narzędzi i aplikacji do wizualizacji badań i ich wyników. Do najpopularniejszych zalicza się: Excel, Power BI, Tableau oraz biblioteki programistyczne napisane w Python i R, takie jak Plotly czy ggplot2.

W przypadku bardziej zaawansowanych badań stosuje się algorytmy sztucznej inteligencji. Pisze o tym następująco:

techniki komputerowego przetwarzania obrazów mogą być także użyteczne do badania dużych zestawów wizualnych mediów. Na przykład w 2015 roku porównaliśmy częstotliwość popularnych tematów w obrazach zamieszczonych na Twitterze w 20 amerykańskich miastach, używając przykładów z miliona zdjęć. W innym projekcie, także z 2015 roku, zastosowaliśmy sieci neuronowe do wykrycia obecności tysiąca rodzajów przedmiotów i zachowań na 100 tysiącach zdjęć z Instagrama umieszczonych w pięciu światowych metropoliach²⁴.

Współcześnie, zauważa Manovich, mamy do dyspozycji setki algorytmów i dziesiątki metod pracy z danymi, które możemy zastosować na wszystkich etapach badania, to jest przygotowywania danych, analizy, modelowania czy walidacji, a ponadto opisowe statystyki, nienadzorowane uczenie maszynowe, analizy szeregów czasowych czy analizy sieci społecznych²⁵.

Na koniec, aby przekonać czytelnika do analityki kulturowej i podobnych badań, Manovich znowu odwołuje się do własnych studiów, podając dający do myślenia przykład. Pisze:

Jeśli nie nauczymy się na podstawie obserwacji zdobywać wystarczającej pewności, co ludzie tworzą i jak zachowują się kulturowo, wszelkie teorie i interpretacje, które możemy zaproponować na podstawie naszej intuicji i otrzymanej wiedzy, będą prawdopodobnie błędne. Tak było, gdy przeanalizowaliśmy dane dotyczące 16 milionów zdjęć z Instagrama

21 Tamże, s. 189.

22 Tamże, s. 215.

23 Tamże, s. 228.

24 Tamże, s. 218.

25 Tamże, s. 246.

udostępnionych w 17 globalnych metropoliach w latach 2012–2015, miliona stron mangi, pięciu tysięcy obrazów impresjonistycznych artystów i innych zbiorów danych kulturowych. W każdym z tych przypadków moje przypuszczenia na temat tego, co zobaczą, oparte na intuicji i przyjętej wiedzy, zostały obalone²⁶.

Z powyższego cytatu wypływają co najmniej trzy wnioski: nie da się zrozumieć współczesnej kultury/kultur bez uwzględnienia cyfrowych danych masowo produkowanych w ramach różnych praktyk – zwłaszcza w internecie; dane te nie mogą być traktowane selektywnie, inaczej mówiąc: powinniśmy analizować ich duże zespoły (*big data*); dotychczasowa wiedza i dotychczasowe metody zwykle nie nadają się do interpretacji *big data*. Nikt nie jest przecież w stanie obejrzeć miliona stron mangi ani kilkunastu milionów zdjęć z portalu społecznościowego. Zastosowanie *computer vision* bądź podobnych narzędzi z obszaru *data science* lub AI jest w tym wypadku niezbędne! I najnowsza książka Manovicha w sposób przystępny przybliży nam ten świat nowej cyfrowej nauki, która nazywa się analityką kulturową²⁷.

Manovich, Lev. *Cultural Analytics*. Cambridge MA: The MIT Press, 2020.

Data wpłyńcia: 30 kwietnia 2021 r. Data zatwierdzenia do druku: 30 czerwca 2021 r.



SŁOWA KLUCZOWE: analityka kulturowa, kultura cyfrowa, media, wizualizacja
KEY WORDS: cultural analytics, digital culture, media, visualization

26 Tamże, s. 254.

27 Należy zaznaczyć, że Manovich prezentuje i inne swoje wielkie projekty, które też mogą pomóc nam zrozumieć jego ideę analityki kulturowej. Są to: The Exceptional and Everyday: 144 hours in Kiev, On Broadway, Van Gogh's paintings, Time magazine, 62.5 hours of Kingdom Hearts czy Elsewhere.