



Dział Badań i Analiz NCK

**1
/
2021**

**NOWOŚCI
BADAWCZE
NCK**

styczeń – czerwiec 2021

	WPROWADZENIE	4
1.	ZASTOSOWANIA SZTUCZNEJ INTELIGENCJI W BADANIACH KULTURY	8
	oprac. Ewa Majdecka	
•	Modelowanie słuchaczy i kontekstowe rekomendacje muzyki oparte na archetypach krajów	9
•	Obliczanie dystansu identyfikacji kulturowej z wykorzystaniem sztucznej inteligencji jako narzędzia do analizy rdzennych społeczności Amazonii. Studium przypadku w społeczności Waorani	12
•	Uczenie maszynowe w segmentacji turystów – w oparciu o postrzeganie zrównoważonego rozwoju oraz wiarygodności celów podróży	15
2.	NOWE PODEJŚCIA DO ANALIZ MATERIAŁÓW AUDIOWIZUALNYCH	19
	oprac. Joanna Byszuk	
•	Stylometria dialogów filmowych: zalety i pułapki	20
	oprac. dr Radosław Bomba	
•	Połączenie analizy obliczeniowej i ludzkiego doświadczenia: sztuki performatywne i humanistyka cyfrowa	23
•	Eksploracja zdigitalizowanych kolekcji ruchomych obrazów: projekt SEMIA, analiza wizualna i zwrot w stronę abstrakcji	27

3.	DZIEDZICTWO KULTUROWE W ERZE CYFRYZACJI	30
	oprac. Joanna Byszuk	
•	Archiwa sieci kontaktów: historia ilościowa i zbiory grup dokumentów	31
•	Po co mapować literaturę? Prototypowanie geoprzestrzenne w badaniach literackich i humanistyce cyfrowej	34
•	Dywergencja i złożoność różnic w tekstach i kulturze	37
	oprac. Marta M. Świetlik	
•	<i>Close-up Cloud</i> : wizualizacja detali w dynamicznym przeglądzie kolekcji sztuki	40
	oprac. dr Radosław Bomba	
•	Cyfrowy kolonializm kulturowy: pomiar uprzedzeń w zagregowanych i zdigitalizowanych treściach gromadzonych w Google Arts and Culture	45
4.	BADANIA SPOŁECZNOŚCI INTERNETOWYCH, NOWE MEDIA	47
	oprac. dr Michał Wyrwa	
•	Kluczowi influencerzy dyplomacji publicznej 2.0: analiza sieci społecznościowych w wybranych krajach	48
	oprac. dr Aleksandra Janus	
•	Marzenia o podróży. Netnografia Pinterestu	52
•	Analiza interakcji na Facebooku dotyczących przemocy w relacjach zawodowych w służbie publicznej	55

WPROWADZENIE

W prezentowanych „Nowościach Badawczych NCK” 2021, nr 1 podjęliśmy tematykę badań społeczności internetowych (wirtualnych) oraz badań kultury z wykorzystaniem metod i narzędzi cyfrowych. Impulsem do zgłębienia powyższych zagadnień była obserwacja dwóch dynamicznie rozwijających się dyscyplin: humanistyki cyfrowej i socjologii cyfrowej.

Humanistyka cyfrowa (HC) to dziedzina szeroka, niejednoznaczna, a przede wszystkim – za sprawą postępującej cyfryzacji i rozwoju technologii – coraz bardziej rozpowszechniona. Stworzenie jednej uniwersalnej definicji przysparza trudności¹, a także niesie ryzyko zawężenia opisywanego obszaru. Pod pojęciem humanistyki cyfrowej rozumiemy tradycyjne dyscypliny humanistyczne operujące na dużych zbiorach danych z wykorzystaniem możliwości, jakie dają dostępne i stale rozwijające się narzędzia cyfrowe². Jak pokazują polskie i zagraniczne praktyki badawcze, HC stopniowo obejmuje swoim zasięgiem różne teksty kultury: słowo pisane, obraz, zapis audio itp. Ten szeroki zakres działań, interdyscyplinarność oraz mnogość analizowanych form nasuwa wniosek, że HC staje się coraz bardziej rozpowszechniona wśród humanistów. HC wkracza również w sferę biznesową i ma szerokie zastosowanie na rynku, a dzięki temu poszerza również możliwości zawodowe humanistów³. Konieczność pogłębienia i rozbudowania współpracy w zakresie HC zaowocowała utworzeniem przez kilkanaście polskich instytucji konsorcjum DARIAH-PL⁴, zajmującego się rozwijaniem i utrzymywaniem infrastruktury HC w Polsce.

Po metody i narzędzia powstałe na gruncie nauk ścisłych częściej sięga także socjologia, gdyż bez nich coraz trudniej jest wyjaśniać zmiany spo-

-
- ¹ J. Ucińska, *Humanistyka cyfrowa: problemy definiowania pojęcia*, „Zagadnienia Informacji Naukowej – Studia Informacyjne” 2017, z. 1 (109), s. 124–145.
 - ² M. Przystek-Szamokowa, *Czym jest humanistyka cyfrowa? Pole semantyczne pojęcia (zarys)*, „Zagadnienia Informacji Naukowej – Studia Informacyjne” 2016, z. 2 (108), s. 82–93.
 - ³ W szkołach wyższych tworzy się kierunki studiów skoncentrowane na humanistyce cyfrowej (np. Humanistyka Cyfrowa na Katolickim Uniwersytecie Lubelskim czy Humanistyka 2.0 na Uniwersytecie Kazimierza Wielkiego w Bydgoszczy), a przy instytutach badawczych powstają jednostki zajmujące się obecnością humanistyki w sieci, jak np. Centrum Humanistyki Cyfrowej Instytutu Badań Literackich PAN.
 - ⁴ Strona internetowa DARIAH-PL, <https://dariah.pl> [dostęp: 31.03.2021].

teczne⁵. Znaczna część dziedzin życia i relacji międzyludzkich przeniosła się do rzeczywistości wirtualnej, która stała się ważnym obszarem aktywności społecznej. Dariusz Jemielniak (2018) pisał, że socjologia stoi przed koniecznością włączania badań społeczności internetowych do większości projektów badawczych, w przeciwnym wypadku sporo utraci na szczególności w wyjaśnianiu zjawisk. Implikuje to konieczność czerpania z warsztatu analitycznego wywodzącego się z nauk ścisłych oraz włączanie *big data* do etnografii cyfrowej⁶.

Połączenie tradycji z nowoczesnością – humanistyki i socjologii z postępującą technologizacją i cyfryzacją – skutkuje zawiązywaniem przez naukowców interdyscyplinarnej współpracy. Wspólny cel łączy „tradycyjnych” badaczy humanistycznych i socjologów, jak również naukowców reprezentujących nauki ścisłe, w szczególności informatyków i analityków *big data*. Tworzenie i analiza dużych zbiorów danych związanych z wytworami kultury, np. cyfrowa archiwizacja i późniejsza analiza tekstów literatury polskiej, wymaga zaangażowania wielu grup specjalistów.

W Polsce od kilkunastu lat prowadzone są przedsięwzięcia, w których zaimplementowano nowe technologie do projektów badawczych z zakresu humanistyki: cyfrowe repozytoria i korpusy językowe, programy upowszechniające obecność kultury w przestrzeni internetowej. Jednym z przykładów jest realizowany w latach 2018–2021 projekt „KRONIK@ – Krajowe Repozytorium Obiektów Nauki i Kultury”. Ma on na celu: „stworzenie w ramach istniejącej infrastruktury państwa miejsca do długotrwałego i bezpiecznego przechowywania zasobów kultury i nauki w postaci repozytorium cyfrowych zasobów, a także budowę platformy integrującej wszystkie dotychczasowe portale w tym obszarze, wraz z wyszukiwarką umożliwiającą przeszukiwanie metadanych zintegrowanych zbiorów”⁷.

Inne znaczące polskie projekty to m.in.: największy korpus współczesnej polszczyzny – Narodowy Korpus Języka Polskiego⁸, biblioteka cyfrowa –

⁵ M. Szpunar, *Humanistyka Cyfrowa a Socjologia Cyfrowa. Nowy paradygmat badań naukowych*, „Zarządzanie w Kulturze” 2016, nr 4 (17), s. 355–369.

⁶ D. Jemielniak, *Socjologia 2.0: o potrzebie łączenia Big Data z etnografią cyfrową, wyzwaniach jakościowej socjologii i systematyzacji pojęć*, „Studia Socjologiczne” 2018, z. 2 (229).

⁷ Strona internetowa Serwis Rzeczypospolitej Polskiej, <https://www.gov.pl/web/cyfryzacja/kronik-krajowe-repozytorium-obiektow-nauki-i-kultury> [dostęp: 31.03.2021].

⁸ Strona internetowa Narodowego Korpusu Języka Polskiego: <http://nkjp.pl/> [dostęp: 15.03.2021].

Polona⁹, największy zbiór wywiadów ze świadkami historii XX w. – Archiwum Historii Mówionej¹⁰ czy też jedyne jak dotąd przedsięwzięcie z zakresu historii sztuki – Urus¹¹ – baza informacji o grafikach związanych z Polską, Litwą i Prusami.

Jak wynika z doświadczeń 2020 r., cyfryzacja w kulturze to nie tylko przejaw rozwoju, lecz także przystosowania się do zaistniałych pandemicznych warunków. Przykładem może być program dotacyjny „Kultura w sieci”. Odpowiadał on na problem czasowego zamknięcia instytucji kultury w czasie pandemii i stworzony został w celu wspierania projektów rozwijających kapitał kulturowy odbiorców za pomocą narzędzi internetowych¹².

Zainspirowani obserwacją dokonań w zakresie humanistyki cyfrowej oraz ekspansją tej dyscypliny, podjęliśmy się sprawdzenia, jakie obecnie prowadzi się na świecie oraz w naszym kraju projekty badawcze łączące tradycyjne humanistyczne dyscypliny z infrastrukturą technologiczną, a także analizy socjologiczne z użyciem narzędzi cyfrowych. Przedmiotem naszego zainteresowania są zarówno najnowsze cyfrowe narzędzia badawcze, jak i wirtualne obiekty badań – internetowe społeczności, nowe media i aplikacje. Do cyfrowego instrumentarium zaliczają się bowiem oba te elementy – nie tylko nowe narzędzia, lecz także badania zdarzeń toczących się w internecie¹³.

Na łamach „Nowości Badawczych NCK” 2021, nr 1 dokonaliśmy przeglądu cyfrowych metod i narzędzi służących do badania kultury. Sprawdziliśmy, jakie zastosowania znalazły one w najnowszych projektach badawczych z zakresu humanistyki cyfrowej. Wskazaliśmy również na aktualne badania społeczności funkcjonujących w przestrzeni internetu. W numerze zaprezentowaliśmy syntetyczne opisy projektów badawczych odnoszących się do literatury, sztuk audiowizualnych, dziedzictwa kulturowego, a także nowych mediów. Pokazaliśmy różnorodne przykłady

⁹ Strona internetowa Polony: <https://polona.pl/> [dostęp: 15.03.2021].

¹⁰ Strona internetowa Archiwum Historii Mówionej: <https://dsh.waw.pl/81-archiwum-historii-mowionej,ahm> [dostęp: 15.03.2021].

¹¹ Strona internetowa bazy Urus: <http://urus.uw.edu.pl/oprojekcie.html> [dostęp: 15.03.2021].

¹² O różnych formach wsparcia w ramach programu „Kultura w sieci” szerzej pisaliśmy w najnowszym wydaniu „Rocznika Kultury Polskiej” 2020, <https://sklep.nck.pl/pl/p/Rocznik-Kultury-Polskiej-2020-do-pobrania-PDF-/917> [dostęp: 15.03.2021].

¹³ Por. M. Szpunar, *Nowe media – nowe metody badawcze? Metodologia pracy cyfrowego humanisty* [w:] *Metody badań medioznawczych i ich zastosowanie*, red. A. Szymańska, M. Lisowska-Magdziarz, A. Hess, Kraków 2018, s. 325–339.

zastosowań sztucznej inteligencji w badaniach kultury. Wyboru i opisu zagranicznych artykułów naukowych z ostatnich miesięcy dokonali eksperci związani z obszarem HC. Mamy nadzieję, że projekty zaprezentowane w najnowszym numerze będą stanowiły inspirację dla badaczek i badaczy zainteresowanych humanistyką i socjologią cyfrową.

Serdecznie zapraszamy do lektury.

Redakcja „Nowości Badawczych NCK”

1.



**ZASTOSOWANIA
SZTUCZNEJ
INTELIGENCJI
W BADANIACH KULTURY**

MODELOWANIE SŁUCHACZY I KONTEKSTOWE REKOMENDACJE MUZYKI OPARTE NA ARCHETYPACH KRAJÓW

LISTENER MODELING AND CONTEXT-AWARE MUSIC RECOMMENDATION BASED ON COUNTRY ARCHETYPES

Markus Schedl, Christine Bauer, Wolfgang Reisinger, Dominik Kowald, Elisabeth Lex

2021

[Link do raportu źródłowego](#)

Systemy zwane silnikami rekomendacyjnymi to algorytmy, których celem jest wskazanie użytkownikowi najodpowiedniejszych czy najistotniejszych dla niego obiektów. Algorytmy są trenowane na bazie wyborów dokonanych wcześniej przez użytkowników i powszechnie wykorzystywane w technologiach cyfrowych w niemal każdym obszarze. Nie tylko są korzystne dla użytkowników, którzy szybciej i trafniej odnajdują poszukiwane przez siebie treści, lecz także mają niewątpliwą wartość ekonomiczną.

Systemy rekomendacyjne ukształtowały dzisiejszy cyfrowy rynek muzyczny, który tworzony jest przez wielu interesariuszy (np. artystów, wytwórnie muzyczne, internetowe platformy muzyczne). W omawianym artykule badacze – wywodzący się z Johannes Kepler University i Graz University of Technology w Austrii oraz Utrecht University w Holandii, a także z austriackiej firmy zajmującej się analizą biznesową opartą na danych i algorytmach – opisali, w jaki sposób przyjęli perspektywę słuchaczy.

Powszechnie wykorzystywanymi typami algorytmów do rekomendowania treści użytkownikom były do tej pory:

- filtrowanie z analizą zawartości (*content-based filtering*): zakłada, że użytkownikowi spodobać się utwory podobne do tych, które poznał poprzednio, tj. podobne zgodnie z wybraną miarą podobieństwa;
- filtrowanie kolaboratywne (*collaborative filtering* – CF): polecenie utworów polubionych przez innych użytkowników o podobnych preferencjach.

Autorzy zdecydowali się oprzeć swój system rekomendacyjny na kraju pochodzenia użytkownika. Nie włączali danych dodatkowych o samym kraju, np. o jego sytuacji ekonomicznej czy kulturze. Uznali bowiem, że

informacje ogólnokrajowe nie będą odzwierciedlać sytuacji poszczególnych użytkowników i zaburzają wynik.

Zbrano dane z serwisu muzycznego Last.fm o 1 088 161 692 zdarzeniach muzycznych, które wydarzyły się pomiędzy styczniem 2005 a sierpniem 2014 r. Na zdarzenie składały się: identyfikatory użytkownika, artyści, albumu i utworu oraz czas, w którym był on odsłuchany. W wykorzystywanej bazie danych każdy z użytkowników miał przypisanych średnio 8879 zdarzeń muzycznych.

Dane pozyskane z serwisu zostały w pewnym stopniu „wyczyszczone”, m.in.:

- wykorzystano dane 46% użytkowników (N=55 186), którzy zadeklarowali kraj pochodzenia. Spośród tej grupy 8% stanowiły osoby z Polski; ze zbioru danych usunięto te kraje, w których nie było dostępnych co najmniej 80 tys. zdarzeń muzycznych i co najmniej 25 użytkowników;
- połączono zbiór danych z serwisu Last.fm z danymi o podgatunkach muzycznych z serwisu Spotify.

Przyjęta metoda doprowadziła do wydzielenia dziewięciu typów klastrów, w których zawierały się różne kraje. Polska znalazła się w klastrze razem z Austrią, Belgią, Czechami, Finlandią, Holandią, Niemcami, Słowacją oraz Szwajcarią. Dzięki analizie najbardziej popularnych utworów we wszystkich klastrach autorzy porównali podgatunki muzyczne, które są najczęściej reprezentowane w poszczególnych klastrach, i nazwali to archetypicznymi preferencjami muzycznymi. Okazały się one wspólne w ramach bliskości geograficznej, na co w dużym stopniu wskazują klastry.

Kolejnymi etapami w omawianym badaniu były:

1. identyfikacja klastrów krajów (za pomocą analizy składowych głównych [PCA]¹ wybrano 100 cech, które wyjaśniały 99,8% wariacji danych, a następnie dokonano klasteryzacji za pomocą algorytmów t-SNE i OPTICS);
2. identyfikacja archetypów konsumpcji muzycznej odpowiadających klastrów (oparta na danych behawioralnych z serwisu muzycznego Last.fm);
3. wykorzystanie zdarzeń muzycznych pobranych od użytkowników, którzy podali swój kraj pochodzenia, jako zestawu danych do treningu algorytmicznego;
4. wydzielenie dziewięciu archetypicznych klastrów krajowych.

¹ PCA to algorytm redukcji wymiarowości, który transformuje wejściowe zmienne na tzw. składowe główne, zachowujące możliwie dużo informacji ze zbioru. Kolejne składowe są mniej ważne od poprzednich. Zwykle niewielka liczba początkowych składowych zawiera większość informacji, a dalsze odrzuca się jako szum.

Autorzy za szczególnie istotną kwestię w realizacji tego typu badania uznali czyszczenie bazy danych z obserwacji odstających, tj. takich, które mogą wpłynąć w nieuzasadniony sposób na wyniki całości (np. użytkownicy słuchający tego samego i niepopularnego wśród innych utworu; niewiarygodne dane o kraju pochodzenia lub wieku).

Modyfikacja i rozwój algorytmu planowane są w kolejnych badaniach, m.in. w zakresie wykorzystania modeli probabilistycznych, które byłyby w stanie określić prawdopodobieństwo przynależności użytkownika do klastra krajowego. Ponadto autorzy chcieliby porównać opisywane wyniki z wynikami badania opartego na analizie artystów, którzy są najczęściej słuchani, a nie na analizie samych utworów muzycznych.

Niewątpliwym osiągnięciem autorów było wydzielenie oraz zdefiniowanie krajowych klastrów i przypisanych im charakterystycznych upodobań muzycznych. Wyniki mają nie tylko istotną wartość ekonomiczną, o której mowa w artykule. Stanowią istotny wkład w badania nad muzyką jako częścią kultury i stylu życia. Tym samym są uzupełnieniem innych podejść do analizy preferencji muzycznych. Nie byłoby to możliwe, gdyby nie wykorzystanie algorytmów. Z tego też względu warto podkreślić interdyscyplinarny model pracy badaczy, z których część to przedstawiciele uniwersytetów, jeden zaś jest pracownikiem centrum badawczego wyspecjalizowanego i doświadczonego w wykorzystywaniu oraz modyfikowaniu algorytmów.

Badacze udostępnili zbiór danych² oraz kod³ wykorzystywany w powyższych badaniach do publicznego użytku.

oprac. Ewa Majdecka

² Strona internetowa Zenodo, <https://zenodo.org/record/3907362#.XvRq1CgzZPY> [dostęp: 3.03.2021].

³ Strona internetowa platformy GitLab, https://gitlab.cp.jku.at/markus/fiai2020_country_clusters [dostęp: 3.03.2021].

OBLICZANIE DYSTANSU IDENTYFIKACJI KULTUROWEJ Z WYKORZYSTANIEM SZTUCZNEJ INTELIGENCJI JAKO NARZĘDZIA DO ANALIZY RDZENNYCH SPOŁECZNOŚCI AMAZONII. STUDIUM PRZYPADKU W SPOŁECZNOŚCI WAORANI

CULTURAL IDENTITY DISTANCE COMPUTATION THROUGH ARTIFICIAL INTELLIGENCE AS AN ANALYSIS TOOL OF THE AMAZON INDIGENOUS PEOPLE. A CASE STUDY IN THE WAORANI COMMUNITY

Aldrin Marcel Espín-León, Antonio Jimeno-Morenilla, María Luisa Pertegal-Felices,
Jorge Azorín-López
2020

[Link do raportu źródłowego](#)

Autorzy badania, wywodzący się z uniwersytetów w Hiszpanii (University of Alicante) oraz w Ekwadorze (Central University of Ecuador), podjęli się stworzenia narzędzia (kwestionariusza oraz algorytmu analizującego zebrane dane), które pozwoliłoby na zmierzenie stanu akulturacji rdzennej społeczności amazońskiej i południowoamerykańskich społeczności miejskich. Ze względu na brak wcześniejszych badań ilościowych nie istniały dane o zmianach tożsamości kulturowej rdzennej społeczności Waorani w trwających procesach globalizacji, migracji oraz zniszczeń środowiskowych na terenie Puszczy Amazońskiej.

Spółeczność Waorani, która poddana została analizom, zamieszkuje amazońskie tereny Ekwadoru i Peru. Liczy ona ok. 13 tys. członków. Do porównania wybrano mieszkańców dwóch ekwadorskich miast: Tena i Quito. Quito jako stolica Ekwadoru reprezentuje w badaniu najbardziej zwesternioną kulturę. Tena to amazońskie miasto turystyczne, do którego migrują również przedstawiciele rdzennych społeczności amazońskich.

Próba w przedmiotowym badaniu wyniosła 299 osób, z czego 88 to Waorani, 100 to mieszkańcy Quito, a 111 – mieszkańcy miasta Tena.

Badanie zostało podzielone na cztery etapy:

1. Przygotowanie narzędzia badawczego umożliwiającego zebranie danych od rdzennych społeczności – oddającego kulturową perspektywę Waorani i nie-narzucającego kategorii ujmowania świata przyjętych w kulturze zachodniej.
2. Proces zbierania danych jakościowych od społeczności Waorani, mieszkańców miast Tena oraz Quito.
3. Wydzielenie 99 elementów składających się na tożsamość kulturową, pogrupowanie ich w szersze kategorie oraz stworzenie pięciu wymiarów tożsamości kulturowej społeczności Waorani przy założeniu, że:
 - tożsamość kulturowa Waorani traktowana jest jako wzorzec, do którego zostały porównane tożsamości kulturowe mieszkańców dwóch wybranych miast;
 - każdy z 99 elementów mógł uzyskać wartość od 0 do 4.
4. Przy użyciu sieci Kohonena (ang. *self-organising maps* – SOM)¹ mapowanie zebranych od poszczególnych jednostek danych ilościowych dotyczących ich tożsamości kulturowej. Uzyskane mapy przedstawiały dystans kulturowy pomiędzy poszczególnymi jednostkami wobec grupy badanych, do której należały, ale również wobec pozostałych dwóch grup badanych.

Mieszkańcy Quito, stolicy Ekwadoru, są najbardziej oddaloną grupą w zakresie tożsamości kulturowej od Waorani. Można wśród nich dostrzec jednak badanych reprezentujących kulturową tożsamość Waorani oraz osoby w pewnym stopniu do niej zbliżone. Mieszkańcy miasta Tena, turystycznego ośrodka amazońskiego i miejsca migracji części przedstawicieli Waorani, są kulturowo bliżej tej rdzennej społeczności. Choć wśród mieszkańców Tena można dostrzec przedstawicieli kultury Waorani, jak i kultury porównywanej do tej z Quito, to ta grupa badanych oscyluje w zasadzie między nimi na zaproponowanej skali tożsamości kulturowej.

Badani przedstawiciele społeczności Waorani nie są kulturowo jednolitą grupą. Najbardziej reprezentatywnymi przedstawicielami kultury Waorani są kobiety. Mężczyźni pełnią w tej społeczności funkcje ekonomiczne, które wiążą się z cyklicznymi podróżami do miast. Mają zatem bezpośredni kontakt z kulturą spoza własnej społeczności rdzennej. Kobiety zazwyczaj

¹ Sieć Kohonena to typ sztucznej sieci neuronowej. Ta samoucząca się sieć przy kolejnych wielokrotnych iteracjach wydziela węzły, inaczej zwane wektorami odniesienia. Pojawiają się one wokół obszarów o dużej gęstości podobnych punktów. W ten sposób w zbiorze danych powstają podstawowe klastry.

nie podróżują do miast w takim stopniu jak mężczyźni, więc nie mają takiej styczności z kulturą inną niż ich własna.

Powtarzane cyklicznie badanie może służyć do pomiaru zmian w dystansie tożsamości kulturowej pomiędzy społecznością rdzenną (w tym przypadku Waorani) a mieszkańcami miast (w tym przypadku w Ekwadorze) i dzięki temu przyczyniać się do ochrony rdzennej kultury. Wyniki badania wskazują na dystans tożsamości kulturowej jako całości, ale również na ewentualne zmiany poszczególnych czynników składających się na tę tożsamość.

oprac. Ewa Majdecka

UCZENIE MASZYNOWE W SEGMENTACJI TURYSTÓW – W OPARCIU O POSTRZEGANIE ZRÓWNOWAŻONEGO ROZWOJU ORAZ WIARYGODNOŚCI CELÓW PODRÓŻY

A MACHINE LEARNING APPROACH TO SEGMENTATION OF TOURISTS BASED ON PERCEIVED DESTINATION SUSTAINABILITY AND TRUSTWORTHINESS

Gabriel I. Penagos-Londoño, Carla Rodriguez-Sanchez, Felipe Ruiz-Moreno, Eduardo Torres
2021

[Link do raportu źródłowego](#)

Autorzy badania – reprezentujący uniwersytety w Kolumbii, Hiszpanii i Chile – przy tworzeniu jego koncepcji wyszli od potrzeby zrozumienia, w jakim stopniu przemysł turystyczny może przysłużyć się do zrównoważonego rozwoju¹ istniejących ośrodków turystycznych. Celem badania była segmentacja turystów pod kątem ich subiektywnej oceny zrównoważonego rozwoju oraz subiektywnej wiarygodności obranych celów podróży². Kolejnym celem badania było także sprawdzenie, co charakteryzuje takich turystów, dla których cele zrównoważonego rozwoju są istotne.

-
- ¹ Organizacja Narodów Zjednoczonych w 2015 r. przyjęła Agendę na rzecz zrównoważonego rozwoju 2030. Poszczególne zadania przypisane do różnych jej Celów odnoszą się do zrównoważonego rozwoju w turystyce oraz ochrony dziedzictwa i różnorodności kulturowej. Dla przykładu jednym z zadań Celu 8 („Promować stabilny, zrównoważony i inkluzywny wzrost gospodarczy, pełne i produktywnie zatrudnienia oraz godną pracę dla wszystkich ludzi”) jest „Do 2030 r. opracować i wdrożyć polityki promujące zrównoważoną turystykę, która tworzy miejsca pracy oraz promuje lokalną kulturę i produkty”. Więcej: Strona internetowa ONZ, http://www.un.org/files/164/Agenda%202030_pl_2016_ostateczna.pdf [dostęp: 5.03.2021].
- ² „Subiektywna wiarygodność celów podróży” jest to termin, którym posługują się autorzy artykułu w odniesieniu do tego, w jaki sposób badani subiektywnie postrzegają za wiarygodne dane miejsce, do którego zamierzają się udać.

W próbie badawczej znalazło się 438 turystów powyżej 17. roku życia przede wszystkim z Chile i Ekwadoru, w niemal równej liczbie kobiety i mężczyźni, o średniej wieku 31,05 lat, z których 76% deklaroowało wyższe wykształcenie. Osoby te w roku poprzedzającym badanie odbyły podróże krajowe. Autorzy pozyskali dane za pomocą kwestionariusza internetowego zrealizowanego pomiędzy czerwcem a wrześniem 2019 r. W efekcie przeprowadzonego badania stworzono model, który pozwolił na charakterystykę trzech segmentów turystów.

W artykule opisano trzy etapy tworzenia modelu segmentacji:

1. Wybrano podzbiór istotnych zmiennych ze zbioru wszystkich możliwych kombinacji zmiennych na podstawie wskaźników zrównoważonego rozwoju. Zmienne obejmowały wymiary zrównoważonego rozwoju (ekonomiczne, socjokulturowe, środowiskowe³) oraz wiarygodności (kompetencje, życzliwość, uczciwość) miejsca podróży⁴. Selekcja ta miała na celu odrzucenie zmiennych, które obniżyłyby jakość budowanego modelu.

Na wymiar socjokulturowy złożyło się docenianie wartości dziedzictwa materialnego (np. pomników), dziedzictwa niematerialnego, wydarzeń kulturalnych, lokalnej kultury, autentyczności i dbałość o nie.

2. Ze względów obliczeniowych do optymalizacji zastosowano algorytm genetyczny⁵.

³ Autorzy przyjmują to trójwymiarowe podejście do zrównoważonego rozwoju jako najbardziej rozpowszechnione w analizach. Opierają się m.in. na artykule P.V. Mathew i S. Sreejesh, *Impact of responsible tourism on destination sustainability and quality of life of community in tourism destinations*, „Journal of Hospitality and Tourism Management” 2017, z. 31, s. 83, <https://www.sciencedirect.com/science/article/pii/S1447677015300474> [dostęp: 5.03.2021].

⁴ Według autorów artykułu wymiar kompetencji (ew. umiejętności, zdolności) z punktu widzenia turystów oznacza, czy dostawca (miejsce, samorząd miejsca itp.) jest w stanie zrealizować to, czego turyści potrzebują. Wymiar życzliwości to okazywanie szacunku i zaangażowanie w kontakt z turystami, dostrzeganie ich potrzeb i perspektywy. Wymiar uczciwości to etyczne postępowanie i działanie zgodnie z obietnicami – oczywiście jako subiektywnie postrzegana przez turystów kategoria dotycząca miejsca ich podróży.

⁵ Algorytm genetyczny to inspirowana doбором naturalnym metoda optymalizacji, w której abstrakcyjnym osobnikom przypisuje się wartości pewnych cech (genów), a następnie na przestrzeni wielu pokoleń krzyżuje się lub mutuje osobniki (proces treningu) celem znalezienia nowych, doskonalszych. Geny najlepszej jednostki będą reprezentowały optymalne rozwiązanie danego problemu.

3. Do klasteryzacji i klasyfikacji zastosowano modele oparte na mieszaninie rozkładów Gaussowskich⁶.

Model pozwolił na wyodrębnienie trzech segmentów, którym nadano nazwy własne, a które posiadają następujące cechy:

- Segment 3 – „Szczególnie optymistyczni” wobec rozwoju turystyki (3,2% próby badawczej):
 - średnia wieku była nieco wyższa (33 lata), więcej mężczyzn (71,4%), osób z wykształceniem średnim i Chilijczyków;
 - najwyżej oceniany był cel podróży pod kątem zrównoważonego rozwoju (socjokulturowego, ekonomicznego oraz środowiskowego) oraz jego wiarygodności;
 - miejsce podróży według przedstawicieli segmentu 3 miało wymienną infrastrukturę i inwestycje przyciągające turystów, co ich zdaniem przynosiło temu miejscu pozytywne ekonomiczne skutki; usługi miały dobry stosunek jakości do ceny, a zanieczyszczenie środowiska, zapachy odczuwane w otoczeniu oraz liczba osób na ulicach były akceptowalne;
 - odwiedzanie obiektów historycznych, ale również inwestowanie i zachowywanie historycznego dziedzictwa były dla nich ważną częścią turystyki;
 - spośród wszystkich segmentów, przedstawiciele tej grupy najwyżej oceniali poziom bezpieczeństwa w miejscu docelowym swojej podróży;
 - charakterystyka tego segmentu potwierdza wcześniejsze wyniki badań, w których mężczyźni są bardziej chętni do działalności turystycznej i w mniejszym stopniu dostrzegają negatywne skutki rozwoju turystyki⁷.
- Segment 2 – „Optymistyczni” wobec rozwoju turystyki:
 - największy segment – 54,6% próby badawczej;
 - średnia wieku 31,3 lat, 52,7% mężczyzn, w porównaniu do innych segmentów stosunkowo najwięcej osób z wykształceniem wyższym niepełnym (60,2%), Ekwadorczycy i Chilijczycy równo reprezentowani;
 - wysoka (ale niższa niż w segmencie 3) ocena miejsca podróży pod kątem zrównoważonego rozwoju oraz wiarygodności w odniesieniu do realizacji założeń zrównoważonego rozwoju;
 - wysoka ocena dziedzictwa historycznego, lokalnej kultury i autentyczności docelowego miejsca podróży.

⁶ Model probabilistyczny, który zakłada, że dane pochodzą z jednego z wielu (tytu, ile jest klas/klastrów) rozkładów normalnych (o tylu wymiarach, ile jest zmiennych w danych).

⁷ Zob. N.G. McGehee, K.L. Andereck, *Factors predicting rural residents' support of tourism*, „Journal of Travel Research” 2004, z. 43 (2), s. 131-140.

- Segment 1 – „Umiarkowanie optymistyczni” wobec rozwoju turystyki:
 - 42,2% próby badawczej;
 - średnia wieku 30,6 lat, 53,5% mężczyzn;
 - największe zróżnicowanie pod kątem wykształcenia – stosunkowo największa grupa z wykształceniem średnim (18,9%) oraz wyższym magisterskim i w stopniu doktora (19,5%);
 - stosunkowo największe grupy studentów/osób bezrobotnych oraz osób z zawodem niespecjalistycznym;
 - największa grupa Ekwadorczyków (45,4%);
 - najniższa ocena celu podróży pod kątem zrównoważonego rozwoju i wiarygodności oraz poziomu bezpieczeństwa w miejscu podróży;
 - charakterystyka segmentu 1 potwierdza wcześniejsze wyniki badań⁸ – jeżeli ocena zrównoważonego rozwoju ośrodka turystycznego jest niska, to analogicznie niska jest również ocena jego wiarygodności.

Prezentowany projekt badawczy jest jednym z pierwszych, w którym wykorzystano trzy zmienne zrównoważonego rozwoju jako podstawy do segmentacji. Badania w takim kształcie mogą i powinny być kontynuowane, aby lepiej rozumieć, co może mieć wpływ na skuteczne (z różnych perspektyw) wdrażanie zasad zrównoważonego rozwoju. Jak wynika z przedmiotowego badania, istnieje grupa podróżnych, dla której ważne są aspekty takiego rozwoju w miejscach, do których docierają jako turyści. Autorzy opracowania uważają, że wiedza o cechach poszczególnych segmentów może pomóc władzom ośrodków turystycznych tworzyć lub modyfikować polityki kształtowania strategii turystycznych. Ponadto autorzy są zdania, że wnioski płynące z badań można wykorzystać do tworzenia ofert zachęcających do przyjazdu tych turystów, dla których zrównoważony rozwój stanowi ważny aspekt nieodłącznie związany z podróżowaniem.

oprac. Ewa Majdecka

⁸ Zob. D.W. Greening, D.B. Turban, *Corporate social performance as a competitive advantage in attracting a quality workforce*, „Business & Society” 2000, z. 39 (3), s. 254-280; I.A. Hengst, P. Jarzabkowski, M. Hoegl, M. Muethel, *Toward a process theory of making sustainability strategies legitimate in action*, „Academy of Management Journal” 2000, z. 63 (1), s. 246-271.

2.



**NOWE PODEJŚCIA DO
ANALIZ MATERIAŁÓW
AUDIOWIZUALNYCH**

STYLOMETRIA FILMOWA

**STYLOMETRIA DIALOGÓW FILMOWYCH:
ZALETY I PUŁAPKI****THE STYLOMETRY OF FILM DIALOGUE: PROS AND PITFALLS**

Agata Hołobut, Jan Rybicki

2020

[Link do raportu źródłowego](#)

Chociaż w badaniach filmów i seriali więcej uwagi poświęca się aspektom wizualnym, w ostatnich latach także język filmu staje się obiektem różnych analiz. W niniejszym artykule badacze z Uniwersytetu Jagiellońskiego podjęli się próby ilościowego zbadania 278 filmów, wyprodukowanych na przestrzeni 84 lat. Celem badań było przede wszystkim wykrycie podobieństw i różnic w dialogach filmowych, odkrycie zależności gatunkowych i tematycznych, powiązań z kontekstem historycznym, a także zweryfikowanie przydatności proponowanych przez autorów metod – stylometrii i analizy wydziwisku (ang. *sentiment analysis*, po polsku czasem nazywanej też analizą sentymentu) w zastosowaniu do języka filmu.

Autorzy uzasadnili użycie ilościowych metod w badaniach języka utworów filmowych wcześniejszymi wynikami badań, które wykazały, że język filmowy jest wyrazisty stylistycznie i różny od języka pisanego (nawet języka pisanego przemówień czy wystąpień teatralnych) oraz spełnia odrębne funkcje językowe, związane zwłaszcza z kreowaniem świata przedstawionego. Wcześniejsze badania autorów pozwoliły im także postawić tezę, że możliwe jest wykrycie pewnych śladów autorskich scenarzysty i/lub reżysera, choć w przypadku filmów, inaczej niż w literaturze pięknej, dominujący wydaje się ślad stylistyczny gatunku.

Zgromadzony materiał badawczy opierał się na transkrypcjach; każdy z 278 filmów został dodatkowo opatrzony etykietą z nazwiskiem scenarzysty oraz jednym z sześciu gatunków przydzielonych na podstawie opisów w bazie IMDb¹. Ponadto w części badania wykorzystującej analizę wydziwisku

¹ Baza informacji o filmach i twórcach filmowych, <https://www.imdb.com> [dostęp: 29.03.2021].

ku teksty zostały poddane lematyzacji, a więc sprowadzeniu wyrazów do ich podstawowych form gramatycznych.

Pierwsza część badań była oparta na stylometrii i analizie sieciowej. Stylometria to metoda polegająca na porównaniu tekstów poprzez obliczenie ich „stylistycznej odległości” od siebie – mierzonej poprzez zestawienie rozkładu statystyk użycia określonych cech, tutaj najczęstszych słów w danych utworach. Metoda opiera się na obserwacji wskazującej na to, że najczęściej używane słowa, a więc raczej jednostki funkcyjne (czyli np. spójniki ‘i’, ‘a’, ‘lub’, ‘oraz’ itp.), a nie dłuższe wyrazy posiadające własne znaczenie, stanowią o stylu danego twórcy i pozwalają na jego dość precyzyjne zidentyfikowanie. Tak zdobyte informacje o podobieństwie poszczególnych tekstów autorzy wykorzystali następnie do zbudowania siatki stylistycznych powiązań między autorami oraz zwizualizowania tych informacji w programie Gephi.

Analiza wydźwięku jest kwestią bardziej skomplikowaną i niejednoznaczną. Autorzy przedstawili w swoim artykule trudności w stosowaniu tej metody w badaniach literatury – zauważając, że cechy, które uznajemy za stanowiące o literackości czy jakości stylistycznej tekstu, np. metafory czy złożone środki stylistyczne, niekoniecznie pasują do schematu działania algorytmów w tej metodzie (w większości rozwiązań opartych na klasyfikowaniu słów według leksykonów słów pozytywnych i negatywnych lub nacechowanych emocjonalnie). Badacze zwrócili także uwagę na trudność komputerowego odczytania niejasności, ironii, metafor czy po prostu negacji.

Mimo tych zastrzeżeń autorzy podjęli się analizy wydźwięku z wykorzystaniem pakietu *syuzhet* w języku R oraz słownika powiązań między słowami i emocjami – NRC Word-Emotion Association Lexicon. Zauważając nieprecyzyjność leksykonu opisali oni przykładowe problemy, takie jak przypisywanie słowom stereotypowego wymiaru emocjonalnego (np. dentysta = strach) oraz ryzyko powielania uprzedzeń i nieprawidłowej interpretacji niejasności. Ich decyzja o podjęciu badań na podstawie analizy wydźwięku była oparta na opinii, że problemy cechuje duża regularność, a więc ewentualne skrzywienie perspektywy będzie dotyczyć całego korpusu, a przy dużej uważności i ostrożności interpretacyjnej można uniknąć wyciągania pochopnych i błędnych wniosków.

Wyniki analizy stylometrycznej i sieciowej wykazały dobrą rozpoznawalność wyrazistych podgatunków – zwłaszcza thrillerów i kryminałów – co autorzy powiązali z użyciem dużej liczby słów niższego rejestru (np. wulgaryzmów) i dynamiką języka „dedukcji i groźby”. Wyrazisty jest także styl

chick flick, czyli filmów „kobiecych”, głównie romantycznych albo rodzinnych, do których algorytm zakwalifikował też niekiedy stereotypowe pozycje, takie jak filmy Damiena Chazelle, Richarda Linklatera czy niektóre prace Woody’ego Allena.

Styl autorski okazał się nie mieć wielkiej siły przebicia ponad gatunek, z wyjątkiem Alfreda Hitchcocka (zarówno jako reżysera, jak i scenarzysty), którego wyrazistość w kontekście rozważanego korpusu jest dużo bardziej zauważalna niż np. indywidualny styl takich twórców jak Quentin Tarantino czy bracia Coen.

Powyższe obserwacje pozwoliły autorom potwierdzić, że statystyki użycia słów pojawiających się najczęściej w filmach umożliwiają wykrycie pewnych zmian w funkcjach języka filmowego. Pozwalają również zauważyć to, że cechy charakterystyczne dla różnych gatunków filmowych, takie jak ekspresyjne wyrażanie uczuć czy budowanie suspense, da się zbadać za pomocą stosowanych tu metod. Analiza sentymentu wskazała z kolei na dość neutralny wydźwięk języka większości utworów filmowych. Ponadto badanie ukazało, że możliwy jest inny podział gatunków: *chick flicks* i *romanse*; filmy o wampirach; filmy akcji, thrillery i filmy o superbohaterach.

Autorzy zwrócili uwagę na pewną ironię – metoda służąca do badań semantycznych okazała się mniej czuła w wykrywaniu różnic semantycznych niż stylometria, która z założenia jest bardziej zorientowana na analizę słów. Artykuł stanowi istotny wkład w potwierdzenie możliwości szerszego niż dotychczas stosowania stylometrii czy analizy wydźwięku oraz wskazuje nowe ścieżki badań filmoznawczych.

oprac. Joanna Byszuk

POŁĄCZENIE ANALIZY OBLICZENIOWEJ I LUDZKIEGO DOŚWIADCZENIA: SZTUKI PERFORMATYWNE I HUMANISTYKA CYFROWA

MATCHING COMPUTATIONAL ANALYSIS AND HUMAN EXPERIENCE: PERFORMATIVE ARTS AND THE DIGITAL HUMANITIES

Jan-Hendrik Bakels, Matthias Grotkopp, Thomas Scherer, Jasper Stratil

2020

[Link do raportu źródłowego](#)

Celem badań zaprezentowanych w tekście jest próba połączenia analiz obliczeniowych z zakresu humanistyki cyfrowej z badaniami jakościowymi. Autorzy zauważają, że od wielu lat w humanistyce dominowały metody *close reading*¹ oparte na studium przypadku. Zmieniło się to wraz z masową digitalizacją zasobów kultury i łatwością dostępu do metadanych. Pojawiły się metody oparte na cyfrowym przetwarzaniu i wizualizacji dużych korpusów danych. Za Frankiem Morettim nazwać je można jako *distant reading*². Metody tego typu opierają się na zastosowaniu analiz statystycznych w badaniach zarówno pojedynczych utworów, jak i grup tekstów. Metoda *distant reading* kładzie nacisk na wyszukiwanie wzorców i podobieństw w dużych zbiorach danych. Według autorów problem stanowi jej adaptacja do badań nad materiałami performatywnymi, w szczególności audiowizualnymi. Metoda ta zastosowana w badaniach filmu dostarcza niekompletnych danych. Dzięki niej można policzyć szereg formalnych aspektów związanych z filmem: liczbę scen, długość poszczególnych ujęć, dominującą kolorystykę,

¹ *Close reading* jest jakościową techniką interpretacji, która skupia się na uważnym czytaniu tekstu i wyszukiwaniu w nim elementów charakterystycznych, szczególnie i specyficznych tylko dla niego.

² Więcej na temat podejścia *distant reading* zob. F. Moretti, *Wykresy, mapy, drzewa. Abstrakcyjne modele na potrzeby historii literatury*, Kraków 2016.

natężenie ruchu w poszczególnych scenach itp., ale w takim podejściu często umyka konkretne znaczenie danych scen. Według badaczy skutkiem takich analiz są wysoce abstrakcyjne wizualizacje.

Problemem badawczym artykułu jest pytanie, czy można połączyć metody jakościowe charakterystyczne dla badań mediów audiowizualnych z metodami ilościowymi stosowanymi w humanistyce cyfrowej? Autorzy przybliżają to zagadnienie, opierając się na zrealizowanym przez nich projekcie badawczym zatytułowanym *Audio-visual rhetorics of affect*, którego istotę stanowi kombinacja jakościowych metod interpretacji dzieła i analiz wykorzystujących duże ilości danych, umożliwiających porównanie kompozycji materiałów audiowizualnych.

Z perspektywy filmoznawstwa wszelkie pytania metodologiczne dotyczące analizy w kontekście humanistyki cyfrowej poprzedzone są wyzwaniem opisowego zaangażowania się w „ulotną” tematykę, która istnieje wyłącznie w czasie jej postrzegania. Dodatkowo sprawę komplikuje fakt, że film jest oglądany przez określonego widza, który może nadawać różne znaczenia i sensy oglądanym scenom. Przywołując jednak psychologię ewolucyjną i koncepcje ucieleśnionego poznania, które wyrosły na kanwie zwrotu performatywnego i cielesnego, autorzy tekstu zauważyli, że możliwa jest intersubiektywna analiza treści audiowizualnych. W celu poparcia swojej tezy badacze odwołali się do założeń neofenomenologicznej teorii filmu, według której obraz audiowizualny jest doświadczany intersubiektywnie jako akt percepcji kinestetycznej, ekspresji ruchu, rytmu i kompozycji audiowizualnej. Badacze zaangażowani w projekt stwierdzili, że wszyscy tworzymy własne doświadczenia, oglądając dane materiały audiowizualne, wzięwszy jednak pod uwagę podstawowe zasady rytmu, ruchu i ucieleśnionej percepcji, te indywidualne doświadczenia odnoszą się do wspólnej dynamiki doświadczeń. Takie ujęcie stało się podstawą badań estetycznego doświadczania ruchomych obrazów w opisywanym projekcie.

W celu zdefiniowania tej wspólnej płaszczyzny doświadczenia filmowego autorzy tekstu wykorzystali metodę eMAEX (ang. *electronically based media analysis of expressive movement images*), rozwijaną przez grupę filmoznawców pod kierunkiem Hermanna Kappelhoffa z Freie Universität w Berlinie. Metoda ta składa się z trzech etapów:

1. Przedmiot badań jest najpierw dzielony na pojedyncze sceny zdefiniowane jako jednostki.
2. Wyodrębnione sceny poddawane są jakościowemu opisowi, który skupia się na pięciu cechach formalnych określających kompozycję badanego ma-

teriału audiowizualnego: choreografii, kamerze, dźwięku, gestach i mimice oraz kompozycji obrazu.

3. Na podstawie tego wstępnego opisu scena jako całość zostaje podzielona na kolejne segmenty ruchu ekspresyjnego (ang. *expressive movement units*, EMU), są to wzajemne oddziaływania różnych poziomów inscenizacji, które są opisywane jako jedna figura. Dwa z pięciu formalnych poziomów zostają następnie wybrane jako poziomy dominujące w ramach każdej EMU.

W projekcie *Audio-visual rhetorics of affect* (AdA)³ wykorzystano tę koncepcję do empirycznych badań nad dużymi korpusami materiałów audiowizualnych (filmów, w tym dokumentów, reportaży telewizyjnych) związanych z globalnym kryzysem finansowym w 2007 r. Zaprezentowana metoda stała się podstawą rozbudowanego systemu anotacji i wizualizacji materiału audiowizualnego. Do tego celu wykorzystano otwarty program do anotacji wideo – Advene⁴. Jak wskazują autorzy projektu, opracowany przez nich system przedstawia utwór audiowizualny, podobnie jak partytura przedstawia utwór orkiestrowy, wyświetlając audiowizualny rytm jako specyficzny wzór graficzny.

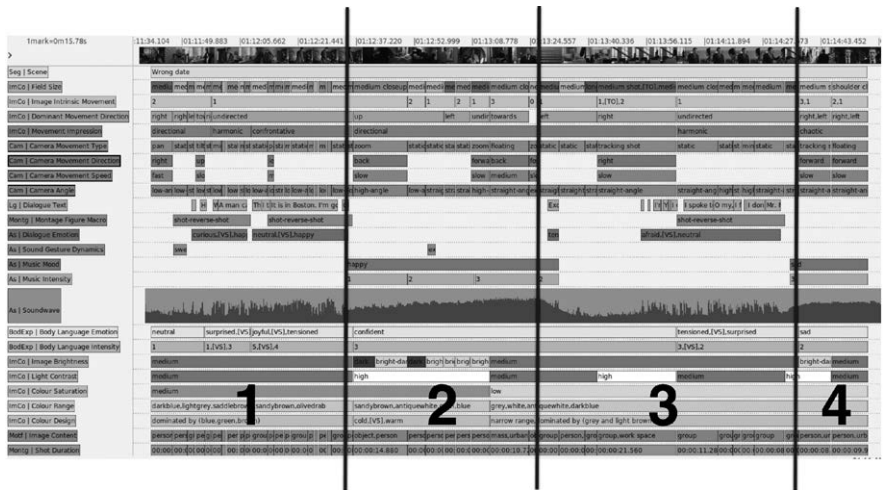
W ramach przedsięwzięcia utworzono także rozbudowaną bazę anotacji, która umożliwia nie tylko analizę oryginalnych sygnatur graficznych, czyli specyficznych wizualizacji obrazujących różne kategorie badawcze charakterystyczne dla danego utworu audiowizualnego (zob. ilustrację na s. 26), ale także pozwala porównywać je między sobą. W ten sposób powstaje możliwość porównywania dużych korpusów danych audiowizualnych. Narzędzie do eksploracji materiałów opracowanych w ramach projektu jest dostępne pod tym adresem: <http://ada.filmontology.org/explorer/>.

Metody i narzędzia zastosowane w projekcie AdA otwierają nowe perspektywy dla analiz audiowizualnych, pokazując strategię łączenia badań jakościowych i ilościowych oraz oferując nowe formy wizualizacji i anotacji.

³ Strona internetowa projektu AdA: <https://projectada.github.io/> [dostęp: 16.04.2021].

⁴ Strona internetowa projektu Advene: <https://www.advene.org/> [dostęp: 16.04.2021].

PRZYKŁAD SEGMENTACJI SCEN I ICH ANOTACJI Z FILMU *THE COMPANY MEN*. WIZUALIZACJA WYKONANA W OTWARTYM PROGRAMIE DO ANOTACJI ADVENE



Źródło: J.-H. Bakels, M. Grotkopp, T. Scherer, J. Stratil, *Matching computational analysis and human experience: performative arts and the digital humanities*, „DHQ: Digital Humanities Quarterly” 2020, z. 4 (14), <http://www.digitalhumanities.org/dhq/vol/14/4/000496/000496.html> [dostęp: 16.04.2021]. Artykuł na licencji CC BY-ND 4.0.

oprac. dr Radosław Bomba

ANALIZA WIZUALNA ZDIGITALIZOWANYCH KOLEKCJI

**EKSPLORACJA ZDIGITALIZOWANYCH
KOLEKCJI RUCHOMYCH OBRAZÓW:
PROJEKT SEMIA, ANALIZA WIZUALNA
I ZWROT W STRONĘ ABSTRAKЦИИ****EXPLORING DIGITISED MOVING IMAGE COLLECTIONS: THE SEMIA
PROJECT, VISUAL ANALYSIS AND THE TURN TO ABSTRACTION**

Eef Masson, Christian Gosvig Olesen, Nanne van Noord, Giovanna Fossati

2020

[Link do raportu źródłowego](#)

Tradycyjne podejście do kolekcji audiowizualnych ma szereg ograniczeń, które uniemożliwiają analizę kolekcji jako całości. Utrudnia to z kolei dostrzeganie nowych powiązań i reinterpretację zgromadzonych materiałów. Biorąc to pod uwagę, zespół badawczy pracujący przy projekcie SEMIA (ang. *Sensory Moving Image Archive*) postanowił wykorzystać narzędzia używane do analizy wizualnej, aby zbudować interfejs umożliwiający otwarte, eksploracyjne podejście do analizy zarchiwizowanych materiałów audiowizualnych. Celem projektu było również stworzenie interfejsu bazującego nie na opisach, a na zmysłowym odbiorze materiałów audiowizualnych, adekwatnych do natury zgromadzonych obiektów. Projekt SEMIA to wynik 2,5-letniej współpracy pomiędzy naukowcami z Uniwersytetu Amsterdamskiego, Uniwersytetu Nauk Stosowanych w Amsterdamie, agencją Studio Louter, Eye Filmmuseum oraz Holenderskim Instytutem Dźwięku i Wizji.

Problemem archiwów przechowujących materiały audiowizualne jest nagromadzenie zasobów, których cechy zmysłowe są prezentowane za pomocą opisów, co niejednokrotnie upraszcza lub zafałszowuje ich specyfikę. Nawet jeżeli w opisach archiwalnych pojawiają się materiały sensoryczne – są one najczęściej fragmentaryczne. Badacze zaangażowani w projekt SEMIA zauważyli, że selektywny charakter opisów materiałów audiowizualnych naraża użytkowników dwóch problemów:

- wymusza przeszukiwanie zasobów na bazie uprzednich interpretacji i założeń katalogujących – co stoi w opozycji do swobodnej eksploracji;
- uniemożliwia opieranie się na indywidualnych doświadczeniach użytkownika, które powstają w procesie obcowania z materiałami audiowizualnymi (dźwięk, kolor, kształt, ruch) i które nie zawsze są odpowiednio oddane przez opis.

Zasadniczym wyzwaniem, przed którym stanął zespół badawczy, było przetworzenie oprogramowania do analizy i wizualizacji: koloru, kształtu, wizualnej złożoności i ruchu na narzędzie umożliwiające alternatywny wgląd w kolekcję materiałów audiowizualnych. Interfejs projektu, bazujący na cechach formalnych materiałów audiowizualnych (tj. dominujących w scenach kolorach, kształtach itp.), miał jednocześnie zachęcać użytkownika do swobodnej eksploracji zasobów, a nie prostego wyszukiwania na podstawie opisów.

Projekt składał się z dwóch faz. Pierwszy etap opierał się na ekstrakcji cech obrazów do analizy, a drugi zakładał skonstruowanie bogatego interfejsu¹ wizualizującego relacje pomiędzy analizowanymi fragmentami. Inspiracją dla projektu SEMIA były badania stylometryczne i współczesne badania mediów z wykorzystaniem narzędzi do analiz wizualnych. Istotne były również działania artystyczne, szczególnie te oparte na zawłaszczaniu ruchomych obrazów. Zespół badawczy wskazał tutaj takich artystów jak Michael Pigott, który swoją działalność artystyczną określał jako „wywoływanie nieczytelności”. Według artysty uczynienie obrazu nieczytelnym pozwalało odczytać go w nowy nieoczekiwany sposób. Z drugiej strony przywołano projekt Geerta Mula *Match of the Day* (2004–2008) – jednego z wczesnych przykładów wykorzystania algorytmów w sztuce. Praca opierała się na obrazach, których dane formalne zostały przez artystę pozyskane, a następnie posłużyły mu do ukazania podobieństw w niezwykle heterogenicznej kolekcji. Takie strategie nie mają na celu znalezienia wzorców, ale nastawione są na przypadkowe lub nieoczekiwane znaleziska ujawniające zaskakujące podobieństwa (lub kontrasty) w materiałach audiowizualnych. Inspiracja wyżej wymienionymi artystami pozwoliła wypracować badaczom

¹ Więcej na temat kategorii „bogaty interfejsów” w tekście R. Bomba, *Bazodanowe interfejsy. Projektowanie interakcji z dużymi zasobami danych kulturowych*, „Kultura Popularna” 2016, z. 4 (50), s. 64–73. Tekst w otwartym dostępie można pobrać ze strony <https://kulturapopularna-online.pl/resources/html/article/details?id=147329> [dostęp: 16.04.2021].

oryginalne podejście, które umożliwia swobodne przeglądanie i otwartą eksplorację zasobów, dając jednocześnie możliwość odkrywania podobieństw wyykających się dotychczasowym interpretacjom.

W projekcie skupiono się na formalnych cechach analizowanych obiektów obejmujących takie charakterystyki jak: kształt, kolor, złożoność wizualna i ruch. Do celów projektu SEMIA zebrano ok. 7 tys. filmów, które następnie podzielono na ponad 100 tys. ujęć za pomocą specjalnego algorytmu. Korpus użyty do testów składał się z fragmentów pobranych ze zbiorów Eye Filmmuseum i Holenderskiego Instytutu Dźwięku i Wizji, dostępnych na platformie Open Images².

Projekt SEMIA miał pokazywać cechy nieoczekiwane oraz połączenia i podobieństwa między zarchiwizowanymi obiektami. Dlatego też zespół badawczy postawił na zróżnicowane podejście łączące elementy nowoczesnych algorytmów uczących się (ang. *deep learning*) z klasycznym podejściem inżynierii cech (ang. *feature engineering*). Podejście takie wynikało z głównego założenia projektu, który miał oferować odbiorcy dużą swobodę eksploracji kolekcji, niezdeterminowaną przez uprzednią interpretację. Celem projektu SEMIA było zainspirowanie użytkowników poprzez ujawnienie potencjalnie istotnych relacji między elementami bazy danych. Podejście to uwzględniało wykorzystanie algorytmów uczących się, które zostały przeszkolone w taki sposób, aby wyszukiwać w materiałach audiowizualnych podobieństwa, operując jednak na dużym poziomie ogólności.

Projekt ukazał, jak w innowacyjny sposób można wykorzystać metody wizualizacji mediów do kreatywnej eksploracji archiwów audiowizualnych. Podejście to oferuje nowe możliwości i pozwala zobaczyć całą kolekcję zasobów audiowizualnych oraz nieoczekiwane powiązanie pomiędzy nimi. Więcej na temat projektu można znaleźć na stronie <https://sensorymovin-gimagearchive.humanities.uva.nl/index.php/about/>.

oprac. dr Radosław Bomba

² Strona internetowa Open Images: <https://www.openbeelden.nl/> [dostęp: 16.04.2021].

3.



**DZIEDZICTWO
KULTUROWE W ERZE
CYFRYZACJI**

ARCHIWA SIECI KONTAKTÓW: HISTORIA ILOŚCIOWA I ZBIORY GRUP DOKUMENTÓW

NETWORKING ARCHIVES: QUANTITATIVE HISTORY AND THE CONTINGENT ARCHIVE

Yann Ryan, Sebastian Ahnert, Ruth Ahnert

2020

[Link do raportu źródłowego](#)

Badacze z Uniwersytetów w Cambridge i Londynie zauważyli, że rzadko podejmowane są próby ustalenia, na ile „wycinek rzeczywistości”, który znajduje się w archiwach, jest reprezentatywny dla danego kontekstu historycznego. Postawili sobie zatem pytanie, czy metody cyfrowe mogą pomóc w ustaleniu, jak duży fragment znanej nam bogatej siedemnastowiecznej korespondencji czołowych angielskich intelektualistów stanowią dostępne im dane. Podjęli próbę identyfikacji, gdzie w siatce powiązań między tymi dokumentami znajdują się potencjalne brakujące ogniwa.

W artykule poruszono trzy kwestie. Po pierwsze, autorzy przedstawili strategię radzenia sobie z brakującymi danymi, wychodząc od próby zrozumienia, jak duże są deficyty w badanych przez nich zbiorach oraz jakiego typu dokumentów brakuje. Po drugie, zaprezentowali opartą na teorii sieci metodę pozwalającą zrozumieć wpływ braków danych na historyczne archiwa listów. Po trzecie, przekonywali, że za pomocą połączonych sieci korespondencji można lepiej zrozumieć strukturę siedemnastowiecznej wymiany myśli intelektualnej.

Za obiekt badań autorzy obrali zbiór angielskiej korespondencji z XVII w., który utworzyli na podstawie metadanych z różnych archiwów i katalogów, przede wszystkim Early Modern Letters Online (EMLO) i State Papers Online (SPO). Dzięki temu uzyskali metaarchiwum liczące ok. 450 tys. listów. EMLO to kolekcja zgromadzona przez Bibliotekę Bodlejańską i Wydział Humanistyczny Uniwersytetu Oxfordzkiego, zawierająca wymianę myśli uznanych w tym czasie intelektualistów. SPO to archiwum dokumentów państwowych, zdigitalizowane przez współpracującą z bibliotekami firmę Gale, powstające od XVII w. w celu zachowania państwowej korespondencji.

Chcąc odkryć, których listów brakuje w metaarchiwum, autorzy zastosowali analizę sieci społecznych, czyli wywodzącą się z teorii sieci metodę badania powiązań między jednostkami oraz analizy ich wzajemnych wpływów. Swój wybór uzasadnili oni tym, że zbiór korespondencji jako zapis interakcji społecznych jest z definicji bardzo dobrym obiektem do badań z zastosowaniem takiej metody. Zwrócili uwagę na to, że pomimo popularności metody sieciowej jej zastosowania w humanistyce rzadko podejmują temat deficytów danych oraz ich wpływu na interpretację zawartości zasobów archiwalnych. Sformułowali również wnioski dotyczące możliwych przyczyn deficytów występujących na etapie realizacji badań, które wykorzystują analizę sieci społecznych: m.in. brakujące lub źle skonstruowane dane czy też źle określone granice. Autorzy postawili hipotezę, że w przypadku rozważanego przez nich korpusu metaarchiwum listów główny problem stanowią utracone lub obecnie niedostępne archiwa osobiste.

Następnie przeanalizowano strukturę zasobów, tzn. sposób opisu listów. Już na tym etapie analiza materiału pokazała, że braki w danych archiwalnych wynikają z przyczyn systemowych i mają charakter regularny – w badanych zbiorach często brakowało paralelnych informacji, np. o autorze i nadawcy. Analiza dat powstania poszczególnych listów pozwoliła autorom zauważyć, że wraz z upływem czasu zwiększała się precyzyjność zapisu – wczesne teksty często datowane były przedziałem kilku lat lub roku, a znaczna część nieco późniejszych listów – przedziałem czasowym 10 dni (wynikającym z rozbieżności między kalendarzem gregoriańskim a juliańskim). Jak podkreślili autorzy, zrozumienie i uwzględnienie wpływu takich braków i rozbieżności było kluczowe do przeprowadzenia rzetelnej analizy.

Próbując odkryć brakujące elementy sieci, autorzy wykorzystali zaadaptowaną wersję metody Peeplesa¹, która pozwala zmierzyć wrażliwość sieci poprzez losowe usuwanie z niej poszczególnych węzłów oraz coraz większych fragmentów, a następnie porównywanie jej parametrów z parametrami pełnej sieci. Pozwala to na wnioskowanie, jak różnego typu nieciągłości w danych wpływają na pozycję pojedynczych węzłów. Stosując tę metodę, autorzy chcieli odkryć, w których miejscach sieci powiązań między listami prawdopodobnie brakowało istotnych elementów.

¹ M. Peeples, *Network science and statistical techniques for dealing with uncertainties in archaeological datasets*, <http://www.mattpeeples.net/netstats.html> [dostęp: 16.04.2021].

Wyniki badania korpusu tą metodą wskazały, że nawet przy usunięciu dużej liczby węzłów pozostałe zajmowały podobną pozycję zarówno w oryginalnej, jak i „podziurawionej” sieci. Pokazało to, że stopień węzła (czyli liczba węzłów, z którymi jest połączony) stanowi najważniejszy aspekt jego centralności w sieci. Badacze odkryli także, że centralni autorzy w zbiorze listów zwykle są ze sobą silnie połączeni, nawet jeśli nie poprzez bezpośrednią komunikację ze sobą nawzajem, to przez wspólnych korespondentów. Ta cecha posiadania dużej liczby wspólnych korespondentów okazała się kluczowa dla określenia, których listów bądź autorów brakuje w bazie. Odkrycie, że w danej sieci relacji utrzymywanych poprzez korespondencję powinny się znaleźć określone osoby byłoby trudne bez wsparcia takich metod, gdyż nawet dogłębna lektura tekstów nie może pomóc w przypadku, gdy np. jedynie domyślamy się, że skoro dwie osoby intensywnie korespondowały z trzecią, to mogły również pozostawać w korespondencji ze sobą nawzajem.

Analiza pozwoliła badaczom odkryć, których autorów brakuje w archiwum. Okazało się, że w przypadku części z nich, np. Johna Dury’ego (jego zaginionej korespondencji autorzy poświęcają więcej uwagi), można uzupełnić zbiory o ich twórczość lub wskazać, w których kolekcjach (również z innych krajów) będą znajdowały się ślady takiej twórczości.

Osoby zainteresowane tą tematyką mogą sięgnąć po obszerniejszą publikację autorów – książkę *Network Turn Changing Perspectives in the Humanities* lub śledzić ich prace na stronie <https://networkingarchives.org/>. Na uwagę zasługuje też miejsce publikacji artykułu, tj. materiały konferencyjne warsztatu *Computational Humanities Research*, który stanowi jedną z ciekawszych propozycji ostatnich lat z zakresu innowacyjnych zastosowań metod komputerowych w badaniu tekstów.

oprac. Joanna Byszuk

MAPOWANIE LITERATURY

PO CO MAPOWAĆ LITERATURĘ? PROTOTYPOWANIE GEOPRZESTRZENNE W BADANIACH LITERACKICH I HUMANISTYCE CYFROWEJ

WHY MAP LITERATURE? GEOSPATIAL PROTOTYPING FOR LITERARY STUDIES AND DIGITAL HUMANITIES

Randa El Khatib, Marcel Schaeben

2020

[Link do raportu źródłowego](#)

Mapowanie literatury to młoda gałąź humanistyki cyfrowej umiejscowiona na granicy literaturoznawstwa i komputerowych metod analizy informacji przestrzennej. Pozwala na nowe odczytanie i interpretację tekstu. Umożliwia również lepsze zrozumienie relacji między tekstem a miejscami w nim opisanymi, a także tego, jak postrzega je autor. Artykuł podejmuje problem prototypowania, czyli ewolucyjnego tworzenia kolejnych, coraz bardziej kompletnych wersji projektu w literackiej analizie geoprzestrzennej. Autorzy prezentują go na przykładzie procesu projektowania i tworzenia aplikacji wizualizującej nawiązania do rzeczywistych miejsc w *Raju utraconym* Johna Milтона.

Badacze z Uniwersytetu w Victorii (Randa El Khatib) i Kolonii (Marcel Schaeben) wyszli od tezy postawionej przez Morgana Ng, głoszącej, że chcąc odczytać przestrzenne odniesienia w twórczości Milтона, konieczne jest zapoznanie się z paratekstami w formie map i obrazów, które towarzyszyły mu przy opracowywaniu koncepcji. Milton reprezentował popularny w siedemnastowiecznej Anglii pogląd historycznej i geograficznej ciągłości wydarzeń biblijnych, czyli przekonanie, że tereny zajmowane przez Imperium Osmańskie były wcześniej sceną wydarzeń biblijnych. Zdaniem autorów często przypisywał on przedstawianym miejscom kategorie moralne, co można ukazać poprzez zwizualizowanie na mapie odniesień do nich, a także podanie odpowiadających im fragmentów tekstu i dodatkowych informacji. Zdaniem autorów pozwala to lepiej zrozumieć moralność w ujęciu pisarza.

Badacze rozpoczęli przygotowania od analizy kompletnej edycji krytycznej *Raju utraconego* oraz cytowanych w niej źródeł, zauważając, że już na tym etapie odbiorca ma trudności z nadążaniem za mnogością nawiązań. Tworzony przez nich projekt „Mapy *Raju utraconego*” miał odpowiedzieć na ten problem, prezentując wizualnie „geograficzną ciągłość” w koncepcji Milтона. Metodologicznie artykuł jest poświęcony dwóm kwestiom: zbieraniu danych i nadawaniu im struktury oraz projektowaniu i budowaniu aplikacji jako procesowi produkcji wiedzy, który przynosi nowe sposoby odczytania tekstu i opracowywania rozwiązań technicznych. Badacze starali się podkreślić wagę prototypowania w tworzeniu rozwiązań komputerowych, zwłaszcza w kontekście tworzenia (cyfrowych) edycji krytycznych, czyli takich edycji tekstu, które powstają na podstawie różnych wersji oryginalnego tekstu (pozwalając w ten sposób ukazać taką jego wersję, która byłaby możliwie najbliższa autorskiej) oraz są wyposażone w komentarz badawczy i obudowę edytorską, np. indeksy.

„MAPA RAJU UTRACONEGO”: GRUPOWANIE PRZESTRZENNE POWIĄZANYCH ZE SOBĄ MIEJSC (W TYM PRZYPADKU RABBY I MIEJSC ZWIĄZANYCH Z KULTEM BOŻKÓW)



Źródło: R. El Khatib, M. Schaeben, *Why map literature? Geospatial prototyping for literary studies and digital humanities*, „Digital Studies/le Champ Numérique” 2020, z. 1 (10), Publikacja na licencji CC BY 3.0.

Stworzona w ramach projektu aplikacja jest oparta na systemie informacji geograficznej GIS (ang. *Geographic Information System*) i ma służyć jako narzędzie eksploracyjne dla badaczy, studentów oraz innych czytelników. Dostępna w niej mapa zawiera odnośniki do konkretnych miejsc wraz z opowiadającymi o nich fragmentami tekstu (w przypadku miejsc wspomnianych wielokrotnie – każda z tych wzmianek dostała osobny rekord), które są opatrzone komentarzami redaktorskimi wyjaśniającymi ich kontekst. W przeciwieństwie do wielu podobnych projektów aplikacja nie poprzestaje na naniesieniu miejsc na współczesne mapy, ale stara się zwizualizować temporalności i zmiany poprzez georektyfikację (czyli naniesienie na główną mapę informacji z innych map) z dwoma historycznymi mapami: mapą krain biblijnych z Biblii króla Jakuba (1611) oraz mapą Imperium Osmańskiego Johna Speeda (1626). Jak podkreślili autorzy, ta funkcja jest szczególnie istotna w przypadku dzieła nawiązującego do tych samych miejsc w różnych okresach, ponieważ umożliwia zbudowanie kontekstu historycznego. Paralelne, zsynchronizowane mapy pozwalają w bardziej intuicyjny i wydajny sposób porównywać szersze kategorie czasowe i warstwy danych.

Jednym z przykładów wpływu prototypowania i modyfikowania decyzji podjętych na wcześniejszych etapach projektu było dodanie opcji pozwalającej użytkownikom aplikacji na wyłączanie poszczególnych funkcji. Jedną z nich była możliwość wyboru wersji z określonym typem i liczbą informacji. Jak wyjaśnili autorzy, stwarza to nowe możliwości korzystania z aplikacji, w tym takie, które wychodzą poza zakres zastosowań przewidziany przez jej twórców. Ponadto, chcąc włączyć w proces tworzenia mapy badaczy niezwiązanych z projektem, autorzy zdecydowali się na opublikowanie jego całości w wolnym dostępie, a w szczególności na użycie arkusza kalkulacyjnego zamiast bazy danych do przechowywania informacji.

Projekt stanowi istotny wkład w dyskurs na temat prototypowania i tworzenia aplikacji z zakresu humanistyki cyfrowej oraz przede wszystkim klarownie prezentuje sposób przygotowywania map literackich.

oprac. Joanna Byszuk

DYWERGENCJA I ZŁOŻONOŚĆ RÓŻNIC W TEKSTACH I KULTURZE

DIVERGENCE AND THE COMPLEXITY OF DIFFERENCE IN TEXT AND CULTURE

Kent K. Chang, Simon DeDeo

2020

[Link do raportu źródłowego](#)

Jednym z największych wyzwań w badaniach literatury na dystans (ang. *distant reading*) jest znalezienie odpowiednich metod porównywania poszczególnych utworów. Kwestia ta dotyczy również badania innych tekstów kultury. W swoim artykule Kent Chang i Simon DeDeo z Carnegie Mellon University podjęli polemikę z postrzeganiem relacji między tekstami jako pewnej przestrzeni, w której różnice da się „zmierzyć” za pomocą matematycznych miar odległości, np. metryki euklidesowej czy manhattańskiej. Kwestie te są już utrwalone w badaniach z zakresu stylometrii czy szerzej rozumianego *distant reading*.

Badacze podważyli niektóre z aksjomatów geometrycznego postrzegania różnic między utworami. Zauważyli, że nie zawsze sprawdzają się one w przypadku badania relacji między tekstami kultury, np. literatury. Zwrócili oni uwagę na rolę „efektu zaskoczenia” przy lekturze. Zauważyli również to, jak niesymetrycznie w rzeczywistości postrzegamy różnice między tekstami – czy to w obrębie jednego utworu (i wpływu kolejności rozdziałów na to, jak odbieramy daną historię i rozumiemy problem), czy to w skali np. twórczości danego autora. Swoją tezę uzasadnili tym, że np. w przypadku Lwa Tołstoja, czytelnik przechodzący od *Anny Kareniny* do *Wojny i pokoju* odczuje większe zaskoczenie mnogością poruszanych tematów i wojennym rozmachem niż ten, który będzie się poruszał w kierunku przeciwnym, czyli od *Wojny i pokoju* do *Anny Kareniny*. Ten drugi odczuje wprawdzie pewne zaskoczenie głębią psychologiczną przedstawienia miłości w *Annie Kareninie*, ale już nie tyloma nowymi motywami, co w *Wojnie i Pokoju*.

Zdaniem autorów można podważyć również kwestię braku skrótów w rozumieniu relacji między utworami. Podkreślili oni rolę tekstów-kluczy,

używanych zwłaszcza w szkolnictwie, które pomagają zapoznać się z najważniejszymi punktami rozwoju literatury, bez konieczności czytania wszystkiego bądź wybierania lektur „po omacku”. Według autorów czytelnik, który zapoznawszy się z dramatem greckim, sięgnie po Szekspira, a dopiero potem po Ibsena, będzie mniej zaskoczony kształtem czy treścią jego dzieł niż ten, który pominąłby w tej drodze angielskiego dramaturga. Autorzy posłużyli się zagadnieniami z zakresu teorii informacji. Przy porównywaniu tekstów zaproponowali wykorzystanie dwóch metod: metody dywergencji Kullbacka-Leibera (czyli miary rozbieżności między dwoma rozkładami prawdopodobieństwa) oraz entropii. Obu tych metod użyto do badania prawdopodobieństwa wystąpienia określonych tematów w literaturze, a także do pomiaru tzw. efektu zaskoczenia. Szczegółowy opis zastosowanej metodologii można znaleźć w załącznikach do artykułu źródłowego.

Sposób badania omawianych w artykule przykładów zawierał, typowy dla komputerowych badań tekstu, element analizy proporcji słów, w tym przypadku charakterystycznych dla poszczególnych haseł tematycznych, które następnie posłużyły do wyliczenia udziału poszczególnych tematów i prawdopodobieństwa wystąpienia „efektu zaskoczenia”.

Podjęcie to badacze zastosowali do skontrastowania ze sobą wspomnianej już twórczości Lwa Tołstoja oraz do zbadania powiązań między poszczególnymi rozdziałami innego dzieła – *O powstawaniu gatunków* Karola Darwina. Ten drugi przypadek jest szczególnie ciekawy ze względu na to, że na jego przykładzie można pokazać różnice między wzorami analitycznego i syntetycznego argumentowania. Jak wykazał eksperyment autorów Changa i DeDeo, Darwin przedstawiał swoje teorie w sposób syntetyczny, prezentując poszczególne elementy w kolejnych rozdziałach i związując je w kompletne, dopracowane tezy w rozdziałach 10. i 11. Jak się okazało, zbadane za pomocą metody dywergencji tezy są punktem kulminacyjnym wyводу Darwina.

Autorzy wspomnieli również o brakach takich metod. Na przykład klasyczne metody porównywania tekstów oparte na symetrycznym mierzeniu odległości – rozumianej jako różnica stylometryczna wyliczana na podstawie proporcji najczęstszych słów czy innych cech *stricte* językowych – nie dają, według badaczy, możliwości zidentyfikowania kierunku rozwoju argumentacji. Symetryczne metody stylometryczne wskazują na różnice i podobieństwa między tekstami, jednak nie wyjaśniają przyczyn tych różnic oraz nie mogą wytłumaczyć, w jaki sposób teksty wpływają na siebie; możliwe jest więc odkrycie za ich pomocą, że dwa rozdziały są do siebie podobne,

ale bez lektury tekstów nie da się stwierdzić, że przyczyną podobieństwa jest to, że np. jeden z nich podsumowuje treść drugiego lub powtarza zawarte w nim argumenty. Autorzy uznali zatem, że proponowana przez nich metoda – dzięki temu, że pozwala na automatyczne wykrycie kierunków wzajemnego wpływania na siebie przez teksty – może być sposobem na ich pogłębioną analizę oraz łatwiejszą i bardziej precyzyjną interpretację.

Artykuł DeDeo i Changa stanowi pierwszą próbę wprowadzenia metody dywergencji do kanonu metod badania tekstów „na dystans”. I może być inspiracją do kolejnych badań z jej wykorzystaniem. Innowacyjność rozwiązania – które, jak się wydaje, pozwala na analizę wzajemnego wpływu tekstów – może zainteresować sympatyków humanistyki cyfrowej i przynieść nowe kierunki rozwoju tej dyscypliny.

oprac. Joanna Byszuk

INTERFEJSY WIZUALNE

***CLOSE-UP CLOUD: WIZUALIZACJA
DETALI W DYNAMICZNYM PRZEGLĄDZIE
KOLEKCJI SZTUKI*****THE CLOSE-UP CLOUD: VISUALIZING DETAILS OF IMAGE COLLECTIONS
IN DYNAMIC OVERVIEWS**

Pauline Junginger, Dennis Ostendorf, Barbara Avila Vissirini, Anastasia Voloshina,
Timo Hausmann, Sarah Kreiseler, Marian Dörk
2020

[Link do raportu źródłowego](#)

W artykule przedstawiono nową, eksperymentalną technikę wizualizacji danych, zaprojektowaną na potrzeby odnajdywania wzorców ikonograficznych występujących w obrębie kolekcji dzieł sztuki. Technika ta znalazła odzwierciedlenie w projekcie *Close-up Cloud*¹ – jest to interfejs, który umożliwia połączenie dwóch przeciwstawnych sposobów oglądu dzieł sztuki: całościowego przeglądu kolekcji z jednoczesną możliwością widzenia detali. Taka metoda oglądania ukazuje użytkownikowi całe bogactwo ikonograficzne zbioru dzieł.

Badacze związani z ośrodkami naukowymi w Poczdamie podjęli się opracowania wizualnego interfejsu dla historycznych kolekcji szklanych negatywów należących do Museum für Kunst und Gewerbe w Hamburgu (Muzeum Sztuki Użytkowej i Dekoracyjnej). Zainspirowani widokiem negatywów na podświetlanym stole, postanowili przenieść tradycyjną dla historii sztuki metodę bliskiego oglądania artefaktów na medium cyfrowe, łącząc ją z dynamicznym widokiem całości kolekcji.

¹ Link do interfejsu (w j. niem.): <https://uclab.fh-potsdam.de/closeupcloud/#/viz> [dostęp: 31.03.2021].

Badacze ustalili, że wcześniejsze rozwiązania bazowały na interfejsach zorientowanych na wyszukiwanie informacji zamiast na wspieranie swobodnej eksploracji zbioru. Zespół postawił sprecyzowane cele projektowe:

1. prowadzenie odbiorcy do szczegółów,
2. wyeksponowanie niepowtarzalnego charakteru kolekcji,
3. wspieranie otwartej eksploracji (ang. *open-ended exploration*).

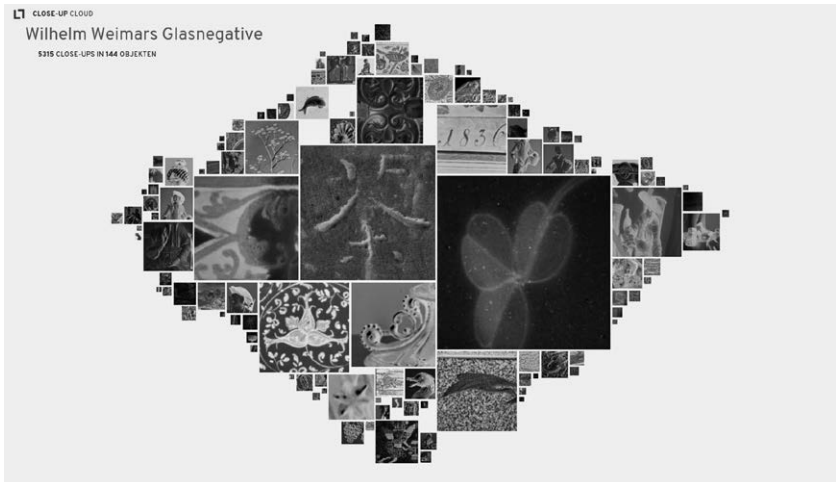
Podczas spotkania warsztatowego (ang. *co-design workshop*) z muzealnymi ekspertami pojawiły się pierwsze pomysły na interfejs, które wyrastały z różnorodnych doświadczeń uczestników wywodzących się z interdyscyplinarnego grona specjalistów w zakresie muzealnictwa, medioznawstwa, komunikacji i projektowania graficznego. Poznawanie historii kolekcji szklanych negatywów, a przede wszystkim możliwość obejrzenia ich na wyostrzającym szczegóły podświetlanym stole, zafascynowało badaczy i ustrukturyzowało późniejszy projekt.

Kolekcja, która stała się obiektem eksperymentu, powstała w wyniku działalności pierwszego kustosa muzeum w Hamburgu. W okresie od 1897 do 1915 r. Wilhelm Weimar stworzył dokumentację w postaci ok. 1700 szklanych negatywów przedstawiających obiekty muzealne. Ich jakość i precyzja są wyjątkowe, także ze względu na skupienie na plastyczności eksponatów i ich detalach ornamentalnych, nierozzerwalnie związanych z funkcją i kształtem obiektu.

Negatywy zostały zdigitalizowane przez muzealnego eksperta, który oznaczył je terminami ikonograficznymi odnoszącymi się do konkretnych szczegółów. Do wizualizacji wybrano 144 szklane negatywy, proporcjonalnie reprezentujące zbiory różnych działów muzeum. Były wśród nich obrazy przedstawiające przedmioty z działów: sztuki starożytnej, europejskiej sztuki dekoracyjnej i rzeźby, Azji Wschodniej, sztuki islamu, secesji, mody i instrumentów muzycznych.

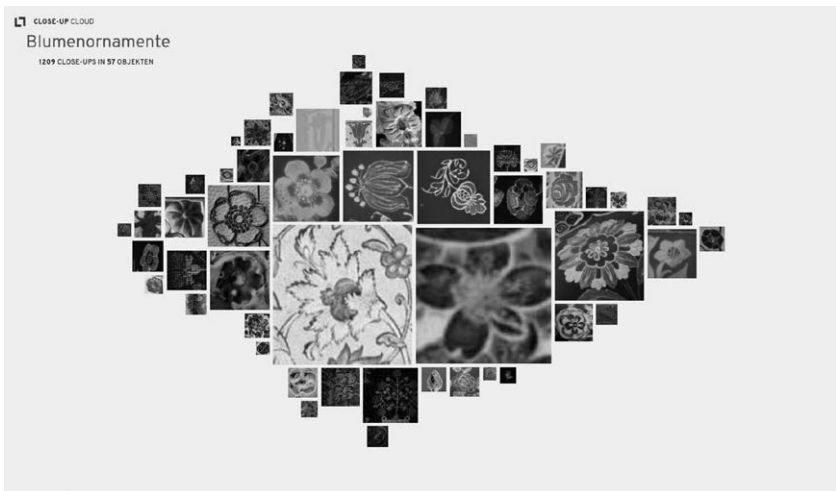
Dzięki *Close-up Cloud* możliwe stało się oglądanie obrazów na trzy sposoby, poprzez: widok kolekcji, znaczników (tagów) lub obiektu. Wszystkie wspierają zarówno horyzontalną eksplorację, jak i doświadczenie immersyjne (zanurzenia w kolekcji). Kluczowe dla interfejsu są znaczniki ikonograficzne, które zostały ręcznie powiązane z określonymi detalami i w ten sposób stworzono cyfrowe wycinki służące jako zbliżenia obiektów.

CLOSE-UP CLOUD. GŁÓWNY WIDOK INTERFEJSU



Źródło: Kolekcja szklanych negatywów stworzona przez Wilhelma Weimara, zdigitalizowana i opublikowana przez autorów artykułu na <https://uclab.fh-potsdam.de/closeupcloud/#/viz> [dostęp: 17.04.2021].

CLOSE-UP CLOUD. WIDOK ZBIORU POD TAGIEM: ORNAMENT ROŚLINNY (BLUMENORNAMENTE)



Źródło: Kolekcja szklanych negatywów stworzona przez Wilhelma Weimara, zdigitalizowana i opublikowana przez autorów artykułu na <https://uclab.fh-potsdam.de/closeupcloud/#/viz> [dostęp: 17.04.2021].

W początkowym widoku interfejsu – trybie kolekcji (fot. 1, s. 42) – prezentowany jest przegląd wybranych 144 negatywów w formie wycinków zawierających szczegóły obiektów, które połączone są ze znacznikami ikonograficznymi. Podobnie jak w chmurach tekstowych rozmiar obrazka reprezentuje częstotliwość występowania danego znacznika w kolekcji. Nie wymagając interakcji użytkownika, *Close-up Cloud* stopniowo odsłania całość zbioru. Zastępując jeden widok detalu innym, oznaczonym tym samym tagiem, tworzy nieustannie nowe kolaże motywów ikonograficznych, które pojawiają się przed oczami użytkownika. Najeżdżenie kursorem na jeden z obrazków umożliwi zobaczenie podpowiedzi tekstowej z tagiem i przejście do widoku tagu (fot. 2, s. 42) prezentującego wszystkie odpowiadające temu znacznikowi detale ikonograficzne. W widoku obiektu (fot. 3, s. 43) wszystkie obrazki z detalami przenoszą się na swoje miejsca w oryginalnym obrazie, osadzając szczegóły we właściwym kontekście. W każdym z widoków możliwe jest powiększanie i przesuwanie interfejsu, odwrócenie negatywu w pozytyw i zobaczenie obiektu w bardziej naturalnej kolorystyce oraz uzyskanie dostępu do metadanych negatywów, takich jak tytuł, nazwisko autora, miejsce i data wykonania.

CLOSE-UP CLOUD. WIDOK OBIEKTU: TALERZ Z ORNAMENTEM KWIATOWYM. FRAGMENTY PRZESUWAJĄ SIĘ NA SWOJE MIEJSCA W OBIEKCIE



Źródło: Kolekcja szklanych negatywów stworzona przez Wilhelma Weimara, zdigitalizowana i opublikowana przez autorów artykułu na <https://uclab.fh-potsdam.de/closeupcloud/#/viz> [dostęp: 17.04.2021].

Aby zweryfikować wartość interfejsu, ocenić jego potencjał w zakresie odkrywania kolekcji i zasugerować możliwe ulepszenia, przeprowadzono testy z udziałem 10 uczestników: pięciu profesjonalistów z doświadczeniem w muzealnictwie oraz pięciu zainteresowanych amatorów. Podczas indywidualnych sesji trwających ok. 30 minut, zawierających badanie prototypu z „protokołem głośnego myślenia” oraz częściowo ustrukturyzowany wywiad, badani potwierdzili zrozumienie widoku kolekcji jako wyodrębnionych fragmentów całego obrazu w celu podkreślenia szczegółów. Wielu uczestników określiło główny obraz jako „atrakcyjny”, a widok zbioru jako „dobrze zorganizowany” i „ustrukturyzowany”. Badania z użytkownikami potwierdziły także, że nowa metoda wizualizacji może być przydatna zarówno dla wyspecjalizowanych badaczy ornamentyki i motywów ikonograficznych, jak i przeciętnych odbiorców zainteresowanych zgłębianiem muzealnych kolekcji.

oprac. Marta M. Świetlik

CYFROWY KOLONIALIZM KULTUROWY: POMIAR UPREDZEŃ W ZAGREGOWANYCH I ZDIGITALIZOWANYCH TREŚCIACH GROMADZONYCH W GOOGLE ARTS AND CULTURE

DIGITAL CULTURAL COLONIALISM: MEASURING BIAS IN AGGREGATED DIGITIZED CONTENT HELD IN GOOGLE ARTS AND CULTURE

Inna Kizhner, Melissa Terras, Maxim Rumyantsev, Valentina Khokhlova,
Elisaveta Demeshkova, Ivan Rudov, Julia Afanasieva
2021

[Link do raportu źródłowego](#)

Google Art and Culture jest jednym z pierwszych przedsięwzięć, w którym firma Google skupiła się na dobrach kulturowych. Ideą projektu jest współpraca z instytucjami zajmującymi się gromadzeniem zasobów kulturowych (muzea i galerie), digitalizacja użyczanych przez nie obiektów w ultrawysokich rozdzielczościach i udostępnianie ich na stronach platformy. Dodatkowo Google Art and Culture wykorzystuje znaną z Map Google'a funkcję *street view*, umożliwiającą zwiedzanie wybranych instytucji, głównie muzeów. Projekt wystartował w lutym 2011 r. i, jak informują sami twórcy przedsięwzięcia, platforma udostępnia obecnie blisko 6 milionów zdigitalizowanych obrazów i stale współpracuje z 2500 muzeami i galeriami na całym świecie.

Dyrektor projektu, Amit Sood, stwierdził, że Google Art and Culture rozpoczął nową erę „łatwego” dostępu do dóbr kultury. Niemniej jednak bliższe przeanalizowanie zasobów udostępnianych przez firmę Google pokazuje istotne niedociągnięcia i wady projektu związane z marginalizowaniem i niedoreprezentowaniem wielu kultur. Zagadnieniem tym zajęł się zespół badawczy z Uniwersytetu Edynburskiego oraz Syberyjskiego Uniwersytetu Federalnego. Badacze po bliższym zapoznaniu się z danymi udostępnianymi przez platformę Google Art and Culture odnotowali, że w zbiorach cyfrowego giganta wiele kultur jest nieobecnych bądź zmarginalizowanych. Zauważyli m.in., że zdecydowana większość prezentowanych zasobów pochodzi z instytucji funkcjonujących w USA. Znacznie mniej zdigitalizowanych

materiałów wywodzi się z krajów europejskich (głównie z Włoch, Niderlandów, Wielkiej Brytanii), a te z Afryki praktycznie nie występują. Wiele krajów jest pominiętych lub pokazywanych jedynie w kontekście współpracy z USA. Dodatkowym problemem, na który wskazał zespół badawczy, jest skupienie się na wielkich centrach, najczęściej stolicach danych krajów i niedoreprezentowanie prowincji i mniejszych miast. W Google Art and Culture dominuje także sztuka XX w.

Według badaczy, przeprowadzone przez nich w 2019 r. analizy pokazały, że projekt Google'a jest daleki od zrównoważonej reprezentacji wszystkich kultur świata i utrwała niebezpieczne, kolonialne stereotypy. Autorzy zwrócili uwagę na fakt, że projekt Google Art and Culture nie jest tworzony przez profesjonalnych naukowców i promuje kolonialną perspektywę uprzywilejowującą kraje szeroko pojętego Zachodu (z nadreprezentacją instytucji z USA). Jest to także niebezpieczne z tego względu, że firma Google tworzy pewien standard, cyfrowy kanon udostępnianych dzieł, który już na wstępie jest stronniczy i obciążony uprzedzeniami. Ta stronniczość może być powielana i rozpowszechniana przez algorytmy sztucznej inteligencji i uczenia maszynowego, które będą bazować na zasobach Google'a. Ten stan rzeczy badacze określili cyfrową kolonizacją kulturową.

Wnioski z badań zakończył apel o dekolonizację danych i dbanie o zrównoważoną reprezentację innych kultur świata. Zespół badawczy zaproponował także szereg rekomendacji, które mogłyby się do tego przyczynić. Między innymi pomocne mogłoby być wdrożenie polityki otwartego dostępu, co pozwoliłoby pominiętym instytucjom partycypować w projekcie. Autorzy spostrzegli także, że warto w podobny sposób przebadać inne duże portale zajmujące się rozpowszechnianiem zdigitalizowanych zasobów kulturowych, takie jak Europeana lub Useum.

oprac. dr Radosław Bomba

4.



**BADANIA
SPOŁECZNOŚCI
INTERNETOWYCH,
NOWE MEDIA**

ANALIZA SIECI SPOŁECZNOŚCIOWYCH

KLUCZOWI INFLUENCERZY DYPLOMACJI PUBLICZNEJ 2.0: ANALIZA SIECI SPOŁECZNOŚCIOWYCH W WYBRANYCH KRAJACH

KEY INFLUENCERS IN PUBLIC DIPLOMACY 2.0: A COUNTRY-BASED SOCIAL NETWORK ANALYSIS

Diana Ingenhoff, Giada Calamai, Efe Sevin

2021

[Link do raportu źródłowego](#)

Działania organów państwowych skoncentrowane na kształtowaniu opinii publicznej za granicą, sprzyjającej celom polityki zagranicznej, tradycyjnie były rozumiane dość jednostronnie. To państwa decydowały o tym, do kogo konkretnie był skierowany przekaz oraz jaka była jego treść. Rozgrywająca się w przestrzeni mediów społecznościowych dyplomacja publiczna 2.0 ma natomiast charakter bardziej otwarty i równy. Polega w znacznej mierze na treściach generowanych przez samych użytkowników mogących wejść w bezpośrednią interakcję z przedstawicielami organów państwowych. W takim kontekście upowszechnienie się metod komunikacji cyfrowej, w tym mediów społecznościowych, diametralnie zmieniło charakter dyplomacji publicznej.

Autorzy badania, pracujący na Uniwersytecie we Fryburgu (Szwajcaria) i Uniwersytecie Towson (USA), podjęli – słabo dotąd przebadaną – kwestię poziomu wpływu, jaki użytkownicy portali społecznościowych mają na kształtowanie wizerunku kraju (ang. *country image*). Opierając się na danych z Twittera dotyczących wpisów na temat trzech państw europejskich (Austrii, Holandii i Szwajcarii), poszukiwali odpowiedzi na trzy pytania:

- Który rodzaj użytkowników ma największy wpływ na budowanie wizerunku państwa na Twitterze?
- Do którego aspektu wizerunku państwa najczęściej się odnoszą wpływowi użytkownicy?

- Jakie różnice w kształtowaniu wizerunku państwa przez użytkowników można dostrzec pomiędzy państwami?

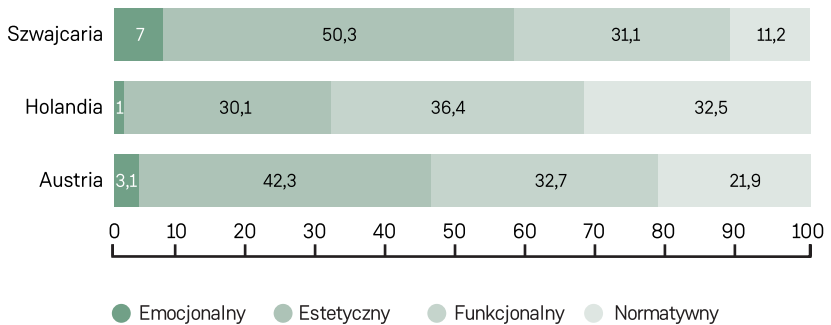
Autorzy skorzystali z narzędzi znanych z badań sieci społecznościowych. Do pobrania danych z Twittera wykorzystano pakiet NodeXL Pro¹, a do przebadania sieci powiązań między użytkownikami – program Gephi². Pobrane zostały wpisy z hashtagami #Austria, #Netherlands i #Switzerland, pochodzące z okresu między kwietniem a czerwcem 2018 r. Z takiego, głównie anglojęzycznego, zbioru liczącego ponad 12 tys. wpisów do dalszej analizy wybrano tylko 848. Były to wpisy najbardziej wpływowych użytkowników wyłonionych przez usunięcie z analizy tych kont, które miały mniejsze niż przeciętne miary centralności³. Każde wpływowe konto było oznaczane jako należące do: (1) organizacji państwowej lub międzynarodowej; (2) organizacji pozarządowej non-profit lub korporacji; (3) użytkownika politycznego (dyplomaci, politycy i inni oficjele); (4) prominentnego użytkownika (celebryci i dziennikarze); (5) zwykłego użytkownika (osoby prywatne). Ostatnim krokiem analizy danych była identyfikacja, który wymiar wizerunku państwa opisuje dany wpis: funkcjonalny, normatywny, estetyczny czy emocjonalny. Do pierwszego wymiaru zaliczane były wpisy o ekonomii, polityce, oświacie, a także bezpieczeństwie, wojskowości, innowacji technologicznej i efektywności działań publicznych. Wpisy w kategorii normatywnej odnosiły się do wartości i norm panujących w danym państwie oraz jego działań w obszarze praw człowieka, walki z terroryzmem, kwestii socjalnych i środowiska naturalnego. Wpisy w kategorii estetycznej dotyczyły turystyki, rozrywki oraz szeroko pojętej kultury, m.in.: zwyczajów, muzyki, gier, sztuki i architektury danego

-
- ¹ M. Smith, A. Ceni, N. Milic-Frayling, B. Shneiderman, E. Mendes Rodrigues, J. Leskovec, C. Dunne, *NodeXL: A free and open network overview, discovery and exploration add-in for Excel 2007/2010/2013/2016*, Social Media Research Foundation, 2010, <https://www.smrfoundation.org/research/publications/> [dostęp: 31.03.2021].
 - ² M. Bastian, S. Heymann, M. Jacomy, *Gephi: an open source software for exploring and manipulating networks*, International AAAI Conference on Weblogs and Social Media, 2009, <https://gephi.org/publications/gephi-bastian-feb09.pdf> [dostęp: 31.03.2021].
 - ³ Parametry często wykorzystywane w analizach sieciowych. W tym badaniu były to: centralność stopnia relacji (związana z liczbą relacji danego użytkownika), centralność przechodniości (związana z tym, na ile dany użytkownik pośredniczy w wymianie informacji), modularność (związana z określaniem grup użytkowników) oraz centralność bliskości (związana z częstotścią bezpośrednich interakcji między użytkownikami).

państwa. W kategorii emocjonalnej natomiast uwzględniane były wpisy zawierające terminy odnoszące się do uczuć.

Najliczniejszą grupą pośród wpływowych użytkowników w każdym z przebadanych krajów były osoby prywatne. Co więcej, stanowiły one najbardziej aktywny rodzaj użytkowników. Obie te zależności potwierdziły rolę odbiorców działań dyplomacji publicznej w przestrzeni cyfrowej w kształtowaniu wizerunku państwa.

WPISY NA TWITTERZE POŚWIĘCONE CZTEREM WYMIAROM WIZERUNKU KRAJÓW: AUSTRII, HOLANDII I SZWAJCARII



Źródło: Opracowanie własne na podstawie D. Ingenhoff, G. Calamai, E. Sevin, *Key influencers in public diplomacy 2.0: a country-based social network analysis*, „Social Media + Society” 2021, z.1 (7), <https://doi.org/10.1177/2056305120981053> [dostęp: 17.04.2021]. Artykuł na licencji CC BY-NC.

Działania użytkowników różniły się pod względem opisywanych wymiarów wizerunku kraju: osoby prywatne skupiały się na wymiarze estetycznym, podczas gdy konta organizacji i osób publicznych opisywały przede wszystkim wymiar funkcjonalny. Różnice były widoczne również między wpisami o poszczególnych krajach. W przypadku Szwajcarii i Austrii wpływowi użytkownicy odnosili się przede wszystkim do estetycznego wymiaru wizerunku państwa, natomiast w Holandii najpowszechniej poruszany był funkcjonalny wymiar wizerunku. Wymiar estetyczny wpisów dotyczył przede wszystkim piękna krajobrazu, zabytków i walorów turystycznych trzech państw. W wymiarze funkcjonalnym skupiano się na innowacyjności i konkurencyjności Szwajcarii i Holandii, szczególnie w zakresie finansów i ekonomii, a także na austriackich wydarzeniach politycznych. Co zaskakujące, we wpisach wymiaru normatywnego raczej nie krytykowano działań państwowych ani nie poruszano tematów kontrowersyjnych.

Dzięki zastosowaniu narzędzi do analizy sieci społecznościowych okazało się, że wpisy wpływowych użytkowników Twittera są cennym źródłem danych na temat wizerunku państw w przestrzeni cyfrowej. Wyniki badania potwierdziły istotną rolę zwykłych użytkowników sieci społecznościowych w kształtowaniu wizerunku kraju, dla których często najważniejszy jest wymiar estetyczny tego wizerunku, związany m.in. z obszarem kultury.

oprac. dr Michał Wyrwa

NETNOGRAFIA

MARZENIA O PODRÓŻY. NETNOGRAFIA PINTERESTU**DREAMING ABOUT TRAVEL: A PINTEREST NETNOGRAPHY**

Ulrike Gretzel

2021

[Link do raportu źródłowego](#)

Media społecznościowe i platformy internetowe stanowią istotne narzędzia wyszukiwania informacji o możliwych celach podróży, a także instrumentarium do ich planowania i dokumentacji. Celem badania była próba zrozumienia, w jaki sposób platformy społecznościowe, w których dominują treści wizualne, inspirują marzenia o podróżach. Badanie przeprowadzono na przykładzie platformy Pinterest. Autorka artykułu, Ulrike Gretzel z University of Southern California, zaproponowała podejście netnograficzne do badania danych zgromadzonych w serwisie Pinterest, który był jak dotąd rzadko analizowany jako źródło danych wizualnych.

Jak stwierdziła autorka artykułu, już samo marzenie o podróży jest istotnym etapem procesu podróżowania. Treści wizualne odgrywają w nim szczególną rolę, kształtującą konsumenckie pragnienia dotyczące przyszłych wyjazdów. Jej zdaniem, taki sposób konstrukcji wizualnej marzeń może wywierać realny wpływ na sektor turystyczny.

Jak zauważyła autorka, podkreślając kontekst swoich badań, w realiach globalnej pandemii COVID-19 wielu użytkowników sieci aktywnie zajmuje się planowaniem, wyobrażaniem sobie i wyszukiwaniem informacji na temat możliwych wycieczek. W serwisie Pinterest od początku jego istnienia jedną z podstawowych tablic sugerowanych użytkownikom jest ta zatytułowana „Miejsca, które chciał(a)bym odwiedzić” (ang. *Places I'd like to go*). Sam serwis, działający od 2010 r., ma ok. 400 milionów użytkowników miesięcznie. Opiera się na treściach wizualnych organizowanych tematycznie przez użytkowników przy pomocy „tablic”, co go odróżnia od innych platform (takich jak Facebook czy Instagram), które pozwalają na udostępnianie treści, ale nie ich porządkowanie i grupowanie.

Do analizy treści wizualnych dostępnych w serwisie Pinterest wykorzystano metodę netnograficzną¹. Autorka odniosła się do sześciu etapów badania, opisując jego przebieg od znalezienia i zdefiniowania pytań badawczych, przez zbieranie, analizę i interpretację danych, po komunikowanie wyników. W omawianym badaniu połączono dwie strategie gromadzenia danych netnograficznych (gromadzenie danych i immersja), co w praktyce oznaczało systematyczne gromadzenie istniejących danych cyfrowych przy jednoczesnym „zanurzeniu” badaczki w badaną przestrzeń cyfrową². Autorka, sama będąc użytkowniczką platformy od 10 lat, zintensyfikowała swoje zaangażowanie w jej ramach na czas badania we wrześniu 2020 r., koncentrując się na badaniu tablic podróżniczych, które publikują użytkownicy Pinterestu, i odnotowywaniu interesujących wzorców zarówno pod względem struktury, jak i treści. W omawianym badaniu zebrano tylko ogólnodostępne dane i posługiwano się pseudonimami (*nickami*) w odniesieniu do konkretnych użytkowników Pinterestu. Analizie poddano łącznie 609 „pinów”, czyli „przypiętych” na wirtualnych tablicach treści wizualnych. Kodowanie ich zawartości zostało przeprowadzone przez autorkę i polegało na ustaleniu cech danego obrazu. Struktura i zawartość 100 poddanych analizie tablic została zbadana za pomocą statystyk opisowych, aby ujawnić także szersze wzorce w danych. Analiza ta przeplatana była spostrzeżeniami z danych immersyjnych. Na poziomie mikro analiza obejmowała dogłębną eksplorację poszczególnych tablic i identyfikację interesujących wzorców w obrębie tych tablic i pomiędzy nimi.

Część poświęcona omówieniu wyników podzielona została na dane natury ogólnej oraz pogłębiony opis czterech wybranych tablic. Badanie dowiodło, że Pinterest jest szeroko wykorzystywany jako narzędzie wyszukiwania informacji o podróżach i że jego rola obejmuje zarówno inspirację wizualną, motywację, jak i funkcję archiwizacyjną dla przyszłych planów podróżniczych. „Tablice” tworzą w przeważającej większości kobiety. Marzenia o podróży często wiążą się z odległymi miejscami (podróż na inny kontynent).

-
- ¹ Stanowi ona jakościową metodę badawczą, która pozwala na pogłębione kontekstowe zrozumienie danych pochodzących z sieci, w tym np. stron internetowych, mediów społecznościowych lub aplikacji mobilnych.
- ² R. Kozinets, *Netnography – The Essential Guide to Qualitative Social Media Research*, Los Angeles 2020.

Wyniki wskazały również na to, że obserwacja wzorców zachowań i trendów na Pinterescie może posłużyć do przewidywania zmian preferencji dotyczących podróży. Treści wizualne są kluczowe w fazie konstruowania marzeń o podróży. Niektóre „pinezki” dodawane są przez użytkowników wyłącznie ze względu na ich walory estetyczne, bez wskazania lokalizacji. Autorka zasugerowała, aby w przyszłych badaniach rozważyć także pozyskanie danych od użytkowników Pinterestu, którzy korzystają z platformy (Pinterest posiada funkcję wymiany wiadomości). Ponadto zaplanowana na dłuższy czas analiza konkretnych tablic pozwoliłaby uzyskać dodatkowe informacje na temat częstotliwości i wzorców „przypinania”. Co istotne, w omawianym badaniu analiza treści dotycząca typów atrakcji była szczątkowa i stanowi inspirację do dalszych badań, w tym np. dla sugerowanego przez autorkę wykorzystania analizy sieci.

oprac. dr Aleksandra Janus

NETNOGRAFIA

ANALIZA INTERAKCJI NA FACEBOOKU DOTYCZĄCYCH PRZEMOCY W RELACJACH ZAWODOWYCH W SŁUŻBIE PUBLICZNEJ

ANÁLISE DE INTERAÇÕES NO FACEBOOK SOBRE ASSÉDIO MORAL NO SERVIÇO PÚBLICO

Kauane Gonçalves da Mata, Paula Fernandes Furbino Bretas

2020

[Link do artykułu źródłowego](#)

Autorki artykułu, współpracujące z Instituto Federal de Brasília, postawiły sobie za cel wstępną analizę interakcji użytkowników sieci społecznościowych, którzy obserwowali tę samą stronę na Facebooku. Użytkownicy mediów społecznościowych byli jednocześnie członkami dwóch grup facebookowych.

W omawianym badaniu wykorzystano metodę netnograficzną zakładającą obserwację najczęściej publikowanych zagadnień, typów treści oraz reakcji na nie (komentarze, polubienia i udostępnienia). Chociaż netnografia charakteryzuje się podejściem jakościowym i zmierza do pogłębionego opisu badanych zjawisk, oferuje także strategię wstępnej analizy pola i to właśnie w takim wydaniu zastosowały podejście netnograficzne autorki omawianego artykułu, zawierającego wyniki wstępnej eksploracji zagadnienia. W badaniu nacisk położony został przede wszystkim na interakcje społeczne zachodzące w środowisku wirtualnym platformy (przy czym warto zaznaczyć, że komentarze, polubienia i udostępnienia rozumiane są właśnie jako interakcje) i na dążenie do zidentyfikowania niektórych wzorców zachowań użytkowników.

Zamierzeniem badaczek było wstępne rozpoznanie, w jaki sposób ludzie korzystają z sieci społecznościowej (Facebook), aby rozwiązać problem nękania/mobbingu w służbie publicznej. W tym celu w badaniu netnograficznym uwzględnione zostały dwie publiczne grupy Facebookowe poświęcone temu zagadnieniu (pod nazwami: ASSÉDIO MORAL NO SERVIÇO PÚBLICO i Assédio Moral No Serviço Público) oraz strona internetowa pod taką samą nazwą. Jedna z badaczek, współautorek artykułu, poprosiła o zgodę

na członkostwo w obu grupach w celu prowadzenia obserwacji. Jej udział w nich nie był aktywny, nie komentowała udostępnianych treści, nie publikowała i nie reagowała na zamieszczane treści i komunikaty. Dane były gromadzone przez badaczki w okresie od sierpnia 2018 do sierpnia 2019 r. Wszystkie posty zostały zebrane w arkuszu kalkulacyjnym i skategoryzowane na podstawie podobieństw tematycznych. Zidentyfikowano także reakcje na każdy post, natomiast analiza miała na celu wskazanie prawdziwości i została przeprowadzona w oparciu o częstotliwość pojawiania się tematów, formatów i reakcji, będąc adaptacją ilościowej analizy treści Bardin¹.

Wyniki obserwacji i analizy interakcji w obu grupach wskazały na to, że najczęściej publikowane i spotykające się z największą liczbą reakcji były informacje na temat typów praktyk przemocowych w środowisku pracy (mobbing, nękanie) oraz samych relacji w miejscu pracy. Jeśli chodzi o formę, dominowały informacje (mające charakter „newsów”) i plakaty. Stwierdzono, że użytkownicy nie pisali wielu osobistych postów czy komentarzy dotyczących własnych doświadczeń, a raczej śledzili informacje na dany temat i dzielili się nimi.

Autorki zauważyły, że uczestnictwo w grupach wydaje się przede wszystkim służyć zdobywaniu i poszerzaniu wiedzy na temat praktyk przemocowych w środowisku pracy oraz w relacjach zawodowych, a także przekazywaniu informacji. Jak przyznały badaczki, może się wydawać, że grupy tego rodzaju służą szukaniu wsparcia i pomocy wobec przemocy osobiście doświadczanej w miejscu pracy. Wyniki omawianego badania dowiodły jednak, że w obu analizowanych grupach dominowało raczej dzielenie się informacjami i poszukiwanie ich, nie zaś wyrażanie uczuć, opinii, wątpliwości czy dzielenie się własnymi doświadczeniami. Ta wstępna analiza, której efekty podsumowuje artykuł, zaledwie sonduje terytorium, a poczynione przez autorki rozpoznanie sugeruje potrzebę pogłębionego badania, potencjalnie również z wykorzystaniem netnografii, które obejmować mogłoby także wywiady z osobami obserwującymi stronę i uczestnikami/uczestniczkami obu grup oraz poszerzoną analizę interakcji.

oprac. dr Aleksandra Janus

¹ Por. L. Bardin, *Análise de Conteúdo*, tłum. L. A. Reto, A. Pinheiro, São Paulo 2016.

Nowości Badawcze NCK 1/2021
© Narodowe Centrum Kultury, Warszawa 2021

REDAKCJA

Kamila Węglarska, redaktor naczelny (NCK)
Kamila Sondej, sekretarz redakcji (NCK)

ZESPÓŁ AUTORSKI

dr Radosław Bomba (Instytut Nauk o Kulturze UMCS)
Joanna Byszuk (Instytut Języka Polskiego PAN)
dr Aleksandra Janus (Instytut Etnologii i Antropologii Kulturowej UJ)
Ewa Majdecka (Instytut Stosowanych Nauk Społecznych UW)
Marta M. Świetlik (Instytut Badań Literackich PAN, Polsko-Japońska Akademia
Technik Komputerowych)
dr Michał Wyrwa (Wydział Psychologii i Kognitywistyki UAM)

KONTAKT

Dział Badań i Analiz Narodowego Centrum Kultury
tel. 22 209 80 97, e-mail: kweglarska@nck.pl

Opracowanie graficzne: Poważne Studio
Skład: Inter Line SC. P. Hrehorowicz, M. Punzet
Redaktor językowy: Zofia Lisiewicz
Korekta: Ida Radziejowska
Redaktor techniczny: Maryla Broda
Redaktor prowadzący: Aleksandra Pawlińska

ISSN: 2719-5481

WYDAWCA



ul. Płocka 13
01-231 Warszawa
<http://nck.pl>
<http://sklep.nck.pl>

Redaktor Naczelny Wydawnictw NCK: dr hab. Rafał Wiśniewski, prof. ucz.
Narodowe Centrum Kultury jest instytucją państwową
działającą na rzecz rozwoju kultury w Polsce.



Publikacja finansowana ze środków
Ministerstwa Kultury, Dziedzictwa Narodowego i Sportu



NARODOWE
CENTRUM
KULTURY