

Wprowadzenie. Metody humanistyki cyfrowej w socjologii jakościowej

Grzegorz Bryda 

Laboratorium CAQDAS TM; Uniwersytet Jagielloński

<https://doi.org/10.18778/1733-8069.20.4.01>

Słowa kluczowe:

socjologia jakościowa, analiza danych jakościowych, humanistyka cyfrowa, sztuczna inteligencja, przetwarzanie języka naturalnego, cyfrowa socjologia jakościowa

Abstrakt: Socjologia jakościowa stoi przed wieloma wyzwaniami i możliwościami dzięki postępom w takich technologiach, jak sztuczna inteligencja, uczenie maszynowe i przetwarzanie języka naturalnego. Technologie te zmieniają sposoby analizy danych jakościowych, oferując nowe perspektywy badawcze i metodologiczne. Sztuczna inteligencja i uczenie maszynowe umożliwiają automatyzację analizy danych, efektywne zarządzanie dużymi zbiorami danych oraz wyciąganie bardziej precyzyjnych i rygorystycznych wniosków przy zachowaniu głębi badań jakościowych. Integracja metod humanistyki cyfrowej z socjologią jakościową ma ogromny potencjał, umożliwiający lepsze zrozumienie złożonych zjawisk i procesów społecznych. Specjalny numer „Przeglądu Socjologii Jakościowej” pokazuje, jak metody humanistyki cyfrowej mogą być z powodzeniem stosowane w badaniach i analizie danych jakościowych, oferując teoretyczne i praktyczne wskazówki dla badaczy. Artykuły zawarte w tym numerze dostarczają analitycznego know-how, inspirując badaczy do dalszego eksplorowania oraz wdrażania innowacyjnych metod humanistyki cyfrowej i technologii cyfrowych w polu socjologii jakościowej.

Grzegorz Bryda

Doktor nauk społecznych, adiunkt w Instytucie Socjologii Uniwersytetu Jagiellońskiego. Specjalizuje się w socjologii kognitywnej, socjologii cyfrowej, socjologii edukacji, socjologii komunikowania oraz metodologii badań jakościowych i mieszanych, ze szczególnym uwzględnieniem zastosowań AI, LLM i NLP/NLU w analizie treści, narracji i analizie performatywnej. Jego zainteresowania obejmują interakcje człowiek – komputer, CAQDAS, komputerowe metody w naukach społecznych i lingwistyce, eksplorację tekstu oraz humanistykę cyfrową. Pracował jako starszy konsultant ds. analizy statystycznej i eksploracji danych w SPSS Polska (obecnie IBM Predictive Solutions), pełnił funkcję pełnomocnika rektora ds. oceny jakości na Uniwersytecie Jagiellońskim (2008–2017). Obecnie kieruje Laboratorium CAQDAS TM oraz Letnią Szkołą Analizy Danych, Metod Badań Jakościowych i Mieszanych.

e-mail: grzegorz.bryda@uj.edu.pl



© by the author, licensee University of Lodz, Poland
This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license CC-BY-NC-ND 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Współczesna socjologia jakościowa stoi przed wieloma wyzwaniami i możliwościami wynikającymi z dynamicznego rozwoju metod humanistyki cyfrowej, sztucznej inteligencji, metod uczenia maszynowego, przetwarzania i rozumienia języka naturalnego w dziedzinie komputerowej analizy danych jakościowych. Rozwój tych technologii nie tylko zmienia sposób, w jaki analizujemy dane jakościowe, ale również otwiera nowe perspektywy badawcze i metodologiczne, które mogą znacząco wzbogacić nasze rozumienie złożonych procesów społecznych. Problematyka ta stanowi od wielu lat obszar moich zainteresowań naukowych. W pracy badawczej koncentruję się na integracji nowoczesnych technologii z tradycyjnymi metodami socjologicznymi, co pozwala na holistyczne podejście do analizy zjawisk społecznych. Skupiam się na wykorzystywaniu technik przetwarzania języka naturalnego i uczenia maszynowego, które umożliwiają automatyzację i usprawnienie procesu analizy dużych zbiorów danych tekstowych i wizualnych, co wcześniej było nieosiągalne. Dlatego tym bardziej się cieszę, że udało się zebrać i po raz pierwszy w Polsce przedstawić w tym numerze specjalnym „Przeglądu Socjologii Jakościowej” wybrane zastosowania metod z obszaru humanistyki cyfrowej we współczesnej socjologii jakościowej. W niniejszym wprowadzeniu chcę przybliżyć zagadnienia rozwijane w kolejnych artykułach oraz elementy nowego podejścia analityczno-badawczego – cyfrowej socjologii jakościowej (Bryda, 2024). Łączy ona metody humanistyki cyfrowej z przetwarzaniem języka naturalnego i analizą danych jakościowych. Głównym celem niniejszego numeru jest pokazanie, jak metody te mogą być z powodzeniem zastosowane w badaniach socjologicznych oraz jakie korzyści płyną z ich integracji. Mam nadzieję, że publikacja ta będzie stanowić inspirację dla badaczy i praktyków, zachęcając do dalszego eksplorowania oraz wdrażania innowacyjnych metod i technologii cyfrowych w badaniach socjologicznych. Jestem przekonany, że integracja metod z obszaru humanistyki cyfrowej z socjologią jakościową przyczyni się do powstania nowych, wartościowych wniosków i odkryć, które wzbogacą naszą wiedzę o społeczeństwie i procesach społecznych.

Humanistyka cyfrowa to interdyscyplinarny obszar badawczy, który łączy metody i narzędzia informatyczne z tradycyjnymi dziedzinami humanistycznymi, takimi jak historia, literatura, sztuka i językoznawstwo (Gold, 2012). W centrum tej dziedziny znajduje się wykorzystanie technologii komputerowych do analizy, interpretacji i prezentacji danych związanych z badaniami humanistycznymi (Svensson, 2016). Zakres humanistyki cyfrowej obejmuje wiele działań, w tym analizę tekstów, cyfrowe archiwizowanie, tworzenie baz danych, wizualizację danych oraz rozwijanie interaktywnych platform edukacyjnych (Burdick i in., 2012). W analizie tekstów literackich badacze mogą wykorzystywać narzędzia do przetwarzania języka naturalnego, aby analizować style pisarskie różnych autorów lub śledzić ewolucję języka w czasie (Moretti, 2013). W historii cyfrowe archiwizowanie umożliwia tworzenie obszernych, łatwo przeszukiwalnych baz danych dokumentów historycznych, co znacznie ułatwia pracę badaczy (Rosenzweig, 2003). W dziedzinie sztuki zaawansowane techniki wizualizacji 3D pozwalają na szczegółowe badanie i konserwację dzieł sztuki bez ryzyka ich uszkodzenia (Drucker, 2013). Dzięki tym narzędziom badacze mogą przeprowadzać analizy na niespotykaną dotąd skalę, odkrywając nowe wzorce i zależności w danych, które byłyby trudne do zauważenia przy użyciu tradycyjnych metod (Jockers, 2013). Na przykład analiza dużych zbiorów tekstów literackich może ujawnić nieoczekiwane powiązania między różnymi dziełami lub autorami, co może prowadzić do nowych interpretacji w badaniach literackich (Underwood, 2017). Humanistyka cyfrowa oferuje

nieograniczone możliwości w kontekście zbierania, analizy i interpretacji danych. Stanowi ona innowacyjne podejście do badań humanistycznych i społecznych, łącząc tradycyjne metody z nowoczesnymi technologiami. Dzięki temu badacze mogą analizować dane na większą skalę, odkrywać nowe zależności i przedstawiać wyniki w bardziej przystępny sposób. To interdyscyplinarne podejście nie tylko poszerza możliwości badawcze, ale i toruje drogę do zrozumienia oraz interpretacji zjawisk kulturowych i społecznych, przyczyniając się do postępu w naukach humanistycznych i społecznych. Integracja narzędzi cyfrowych z badaniami socjologicznymi otwiera nowe perspektywy, pozwalając na bardziej szczegółową i wielowymiarową analizę zjawisk społecznych. Przeobrażenie tradycyjnych metod badawczych dzięki wykorzystaniu zaawansowanych technologii, takich jak *Big Data*, analiza tekstu czy wizualizacja danych, umożliwi socjologom odkrywanie ukrytych wzorców i zależności, które wcześniej były niedostępne.

Socjologia jakościowa to podejście badawcze skoncentrowane na zrozumieniu ludzkich zachowań, doświadczeń i interakcji społecznych. Dąży ona do uzyskania głębokiego wglądu w konteksty społeczno-kulturowe poprzez takie metody, jak wywiady pogłębione, obserwacja uczestnicząca, analiza treści oraz studia przypadków (Denzin, Lincoln, 2017). Celem tej dziedziny jest poznanie znaczeń, jakie ludzie nadają swoim działaniom i relacjom, co umożliwia holistyczne i wieloaspektowe spojrzenie na badane zjawiska. Socjologia jakościowa pozwala na zrozumienie subtelnych i często nieuchwytnych aspektów życia społecznego, które mogą być pominięte przez metody ilościowe. Badania jakościowe często koncentrują się na małych grupach lub jednostkach, co umożliwia głębszą analizę indywidualnych doświadczeń i perspektyw. W kontekście dynamicznego rozwoju technologii połączenie metod humanistyki cyfrowej z socjologią jakościową otwiera nowe możliwości badawcze i przynosi liczne korzyści. Cyfrowe narzędzia analityczne pozwalają na pracę z większymi zbiorami danych, umożliwiając ich bardziej złożoną i precyzyjną analizę (Bail, 2014). Na przykład narzędzia do analizy tekstów mogą automatycznie kodować i badać tysiące stron wywiadów czy transkrypcji, identyfikując kluczowe tematy i wzorce. To z kolei prowadzi do bardziej rygorystycznych i wiarygodnych wniosków, jednocześnie zachowując głębię analizy charakterystyczną dla badań jakościowych (Fielding, Lee, 2021). Integracja ta umożliwia również bardziej efektywne zarządzanie danymi oraz ich przechowywanie i udostępnianie. Dzięki temu badacze mogą korzystać z różnorodnych źródeł danych i tworzyć bardziej kompleksowe analizy (Brosz, Bryda, Siuda, 2017).

Jednym z kluczowych aspektów wykorzystania metod humanistyki cyfrowej w socjologii jakościowej jest możliwość eksploracji nowych źródeł informacji, takich jak media społecznościowe, fora internetowe czy blogi. Dzięki takim technikom jak analiza sentymentów czy mapowanie sieci społecznych badacze mogą zyskać świeże spojrzenie na interakcje międzyludzkie (Marres, 2017). Te nowe źródła danych pozwalają na badanie zjawisk społecznych w czasie rzeczywistym i na skalę wcześniej niedostępną dla tradycyjnych metod badawczych. Równocześnie technologie cyfrowe i informatyczne umożliwiają tworzenie interaktywnych wizualizacji, które ułatwiają prezentację wyników badań w sposób przystępny i angażujący. Takie wizualizacje mogą pomóc w komunikowaniu skomplikowanych zjawisk społecznych zarówno badaczom, jak i szerszej publiczności (Lupton, 2020). Przykładowo, interaktywna mapa może zobrazować rozprzestrzenianie się trendów kulturowych w czasie

i przestrzeni, co przyczynia się do lepszego zrozumienia dynamiki procesów społecznych. W praktyce narzędzia do analizy tekstu, takie jak NVivo, MAXQDA czy Atlas.ti, umożliwiają automatyczne kodowanie i kategoryzowanie dużych zbiorów danych tekstowych. Dzięki algorytmom do analizy sentymentu można oceniać emocjonalny ton wypowiedzi, a techniki analizy częstości pozwalają na identyfikowanie najczęściej występujących tematów i słów kluczowych. Takie narzędzia jak Gephi, Pajek czy NodeXL pozwalają na analizę relacji między jednostkami w sieciach społecznych. Dzięki wizualizacji sieci badacze mogą zidentyfikować kluczowych aktorów, struktury sieciowe i dynamikę interakcji społecznych. Narzędzia do wizualizacji danych, takie jak Tableau, Power BI czy R z biblioteką ggplot2, umożliwiają tworzenie interaktywnych wykresów, map i diagramów. Wizualizacje te pomagają w zrozumieniu i komunikowaniu złożonych danych w przystępny sposób. Wykorzystanie dużych zbiorów danych (ang. *Big Data*) z różnych źródeł – mediów społecznościowych, baz danych rządowych czy platform e-commerce – umożliwia badaczom analizę zachowań i trendów na masową skalę. Narzędzia takie jak Hadoop, Spark czy platformy chmurowe (AWS, Google Cloud) wspierają zarządzanie tymi danymi i ich analizę.

Pomimo wielu korzyści integracja metod cyfrowych z socjologią jakościową niesie ze sobą również wyzwania. Badacze muszą być świadomi potencjalnych błędów i uprzedzeń, które mogą pojawić się w danych cyfrowych, oraz etycznych implikacji korzystania z danych online (Boyd, Crawford, 2012). Kluczowe jest zachowanie krytycznego podejścia i refleksyjności charakterystycznej dla badań jakościowych, nawet przy korzystaniu z zaawansowanych narzędzi cyfrowych. Należy również pamiętać o problemach związanych z reprezentatywnością danych zbieranych online, które mogą nie odzwierciedlać pełnej różnorodności populacji. Ponadto badacze muszą być świadomi możliwości manipulacji i dezinformacji w danych cyfrowych, co może wpływać na wyniki analiz. Metody humanistyki cyfrowej oferują potężne narzędzia do badań socjologicznych, umożliwiając bardziej kompleksową i precyzyjną analizę danych. Przykłady zastosowań w literaturze przedmiotu pokazują, że metody te mogą z powodzeniem wzbogacić badania jakościowe, dostarczając nowych perspektyw i wniosków (Salganik, 2019). Integracja humanistyki cyfrowej z socjologią jakościową ma ogromny potencjał do rozwijania nowych, innowacyjnych podejść badawczych, które mogą prowadzić do głębszego zrozumienia złożonych zjawisk społecznych (Bryda, Costa, 2023). W miarę rozwoju technologii i metod analizy danych możliwości dla socjologii jakościowej będą się poszerzać, otwierając nowe drogi dla badań i odkryć w dziedzinie nauk społecznych (Bryda, 2024). Sukces tego podejścia zależy od umiejętnego łączenia tradycyjnych wartości badań jakościowych z innowacyjnymi technikami cyfrowymi. Badacze muszą stale rozwijać swoje umiejętności i być otwarci na nowe technologie, aby w pełni wykorzystać ich potencjał w badaniach socjologicznych.

Niniejszy numer specjalny „Przeglądu Socjologii Jakościowej” prezentuje artykuły szczegółowo omawiające różnorodne zastosowania metod humanistyki cyfrowej w badaniach jakościowych. Każdy z tych artykułów wnosi unikalny wkład w zrozumienie, jak cyfrowe narzędzia mogą być wykorzystywane w komputerowej analizie danych jakościowych, przedstawiając konkretne studia przypadków oraz krytyczne refleksje na temat wyzwań i możliwości. Ten numer dostarcza czytelnikom nie tylko wiedzy teoretycznej, ale także praktycznych wskazówek i inspiracji do wykorzystania metod

humanistyki cyfrowej w ich własnych badaniach socjologicznych. Prezentowane treści są ważnym źródłem informacji dla wszystkich zainteresowanych nowoczesnymi metodami badawczymi w socjologii jakościowej. Zachęcam do zapoznania się z całością numeru, który przedstawia zarówno perspektywy badawcze, jak i procedury analityczne. Jestem przekonany, że lektura pełnego wydania dostarczy cennych inspiracji i poszerzy horyzonty badawcze czytelników.

Artykuł Huberta Plisieckiego i Agnieszki Kwiatkowskiej pt. *Discovering Representations of Democracy in Big Data: Purposive Semantic Sample Selection for Qualitative and Mixed-Methods Research* przedstawia nowatorskie podejście do selekcji próbek w badaniach opartych na dużych zbiorach danych. W erze cyfrowej ogromna ilość danych tekstowych stanowi zarówno wyzwanie, jak i szansę dla badaczy społecznych. Tradycyjne metody selekcji próbek wymagają intensywnego kodowania manualnego lub wcześniejszej znajomości zbioru danych, co jest czasochłonne i subiektywne. Z kolei metody niesupercywilizowane mogą dawać niespójne wyniki, które nie zawsze są zgodne z kodowaniem opartym na teorii. Aby sprostać tym wyzwaniom, autorzy proponują celową selekcję semantyczną (ang. *purposive semantic sampling*), wykorzystującą osadzenia dokumentów (ang. *embeddings*) stworzone za pomocą ważonej średniej wektorów słów z częstotliwością termów (*tf-idf*). Metoda ta ma na celu zwiększenie skuteczności i wydajności selekcji próbek w badaniach tekstów społecznych. Artykuł ukazuje skuteczność tej metody na przykładzie analizy debat parlamentarnych dotyczących demokracji, stosując zarówno tradycyjne podejścia, jak i nowoczesne techniki oparte na embeddingach słów. Procedury analityczne obejmowały wyodrębnienie pełnego korpusu debat parlamentarnych z polskiego Sejmu w latach 1991–2020, użycie algorytmu word2vec do stworzenia embeddingów słów, testowanie pięciu różnych metod selekcji dokumentów oraz porównanie ich skuteczności za pomocą miar ilościowych i jakościowych. Demokracja, jako pojęcie wielowymiarowe, jest trudna do zdefiniowania i odmiennie interpretowana w różnych kontekstach. Metoda proponowana przez autorów pozwala na efektywne wyszukiwanie dokumentów związanych z demokracją w dużych korpusach tekstów parlamentarnych, co znacząco poprawia jakość badań. Wprowadzenie do artykułu podkreśla znaczenie ilościowej analizy treści jako metody badawczej w naukach społecznych, umożliwiającej redukcję różnorodnych komunikatów tekstowych do kluczowych składników znaczeniowych. Artykuł wskazuje na rosnące znaczenie ilościowej analizy treści w dobie cyfryzacji i dostępności ogromnych zbiorów danych tekstowych oraz na rewolucyjny postęp technologii NLP. W artykule skupiono się na modelowaniu tematycznym przy użyciu osadzonego modelu tematycznego (ang. *Embedded Topic Model – ETM*; Dieng, Ruiz, Blei, 2020). Model ETM różni się od tradycyjnych metod, takich jak LDA, ponieważ do identyfikacji tematów wykorzystuje wstępnie wytrenowane wektory znaczeniowe. W związku z tym dla każdego z podkorpusów zostały wytrenowane osobne wektory znaczeniowe, by lepiej odzwierciedlić ich specyfikę. Autorzy podkreślają znaczenie zaawansowanych metod NLP, pokazując, jak nowoczesne techniki mogą być używane do przetwarzania i analizy dużych zbiorów danych tekstowych, co jest kluczowe dla współczesnej humanistyki cyfrowej. Proponowane podejście do selekcji próbek wzmacnia badania jakościowe i ilościowe, umożliwiając głębszą analizę i eksplorację dużych korpusów danych tekstowych. Artykuł kończy się refleksją nad przyszłością metod analizy treści w naukach społecznych, sugerując, że tradycyjny podział na metody ilościowe i jakościowe może zaniknąć w obliczu rosnących możliwości algorytmicznych technologii przetwarzania danych.

Wskazuje na konieczność integracji probabilistycznych modeli statystycznych z klasycznymi metodami jakościowymi, co pozwala na efektywne badanie dużych zbiorów danych tekstowych i otwiera nowe perspektywy dla badań socjologicznych.

Artykuł Marka Troszyńskiego pt. *Badania dyskursu wspomagane korpusowo (CADS) jako wsparcie jakościowej analizy treści. Studium przypadku wykorzystania programu SketchEngine w badaniach dyskursu* przedstawia możliwość wykorzystania narzędzi lingwistyki korpusowej jako wstępnego etapu jakościowej analizy treści. Autor omawia rozwój metody badań dyskursu wspomaganych korpusem (CADS) i skupia się na funkcjach programu SketchEngine, który wspiera te badania. W tekście zawarte są liczne przykłady ilustrujące sposoby wykorzystania metod CADS i funkcjonalności SketchEngine dla analizy polskiego dyskursu prasowego. Poprzez możliwość łatwego odniesienia do tekstów źródłowych (konkordancje) SketchEngine umożliwia włączenie metod mieszanych do badań dyskursu. Badacze zajmujący się analizą treści medialnych stają przed wyzwaniem zebrania i analizy ogromnej liczby komunikatów. W obliczu rosnącego strumienia tekstów tradycyjna analiza jakościowa staje się niewystarczająca. W artykule omawiane są różne metody selekcji tekstów stosowane w badaniach mediów, takie jak doборы losowy, systematyczny i warstwowy. Autor sugeruje wykorzystanie metod lingwistyki korpusowej jako wsparcia dla jakościowej analizy treści. Podejście to pozwala na merytoryczny dobór tekstów do analizy na podstawie zdefiniowanych wcześniej słów kluczowych i wzorców językowych. Główna część artykułu skupia się na programie SketchEngine, który oferuje narzędzia do analizy korpusowej, takie jak konkordancje, kolokacje, słowa kluczowe i n-gramy. Program umożliwia tworzenie i analizowanie własnych korpusów tekstowych, co jest niezwykle przydatne w badaniach dyskursu prasowego. Dzięki takim funkcjom jak Word Sketch użytkownicy mogą uzyskać jednostronicowe podsumowanie zachowania wybranego słowa w korpusie, co ułatwia identyfikację kluczowych wzorców językowych. Konkordancje pozwalają na przeglądanie fragmentów tekstów w kontekście, co jest kluczowe dla jakościowej analizy treści. W artykule podkreślono, że metody CADS i narzędzia takie jak SketchEngine znacznie zwiększają możliwości badawcze, umożliwiając analizę dużych zbiorów danych tekstowych w sposób bardziej efektywny i zautomatyzowany. Program SketchEngine dzięki swoim zaawansowanym funkcjom pozwala na dokładne badanie dyskursu i identyfikowanie powtarzających się wzorców, co jest nieocenione w analizie medialnej. Wykorzystanie metod CADS w badaniach dyskursu medialnego demonstruje, jak narzędzia lingwistyki korpusowej mogą wspierać jakościową analizę treści, co jest kluczowe dla współczesnych badań w dziedzinie humanistyki cyfrowej. Artykuł stanowi ważny wkład do numeru specjalnego, ukazując nowe możliwości badawcze i metodologiczne, jakie otwierają się przed badaczami dzięki zastosowaniu nowoczesnych narzędzi cyfrowych.

Artykuł Krzysztofa Tomanka zatytułowany *O „subtelnościach” metod oceny wydźwięku wypowiedzi pisanych. Porównanie trzech podejść w analizie sentymentu* omawia wyniki eksperymentu metodologicznego, w którym zastosowano trzy różne metody analizy wypowiedzi tekstowych. Celem było porównanie różnic między manualnym kodowaniem tekstu, analizą półautomatyczną opartą na słowniku klasyfikacyjnym oraz metodą sztucznej inteligencji (AI), reprezentowaną przez ChatGPT w wersji 3.5. Badanie koncentrowało się na analizie sentymentu, czyli ocenie emocjonalnego wydźwięku

wypowiedzi. Pierwsza metoda – manualne kodowanie – polegała na analizie tekstów przez badaczy, którzy nadawali im oceny sentymentu. Jest to metoda najbardziej tradycyjna i uważana za rzetelną, choć subiektywną, ponieważ opiera się na intuicji i doświadczeniu koderów. Druga metoda – analiza półautomatyczna – wykorzystuje słowniki klasyfikacyjne, takie jak SentiwordNet i VADER. Te narzędzia automatycznie oceniają wydźwięk wypowiedzi na podstawie zaprogramowanych reguł i wag przypisanych poszczególnym słowom. W ostatniej metodzie – analizie za pomocą sztucznej inteligencji – algorytmy AI analizują teksty i oceniają ich sentyment, oferując interpretacje wyników. Artykuł zawiera szczegółowe porównanie wyników uzyskanych za pomocą tych trzech metod. Manualne kodowanie okazało się najbardziej subiektywne, z dużą rozbieżnością w ocenach poszczególnych koderów. Analiza półautomatyczna była bardziej spójna, ale miała ograniczenia w rozpoznawaniu kontekstu i subtelnych niuansów emocjonalnych w tekstach. Metoda AI, choć najnowsza, okazała się skuteczna w rozpoznawaniu kontekstu i sarkazmu, co czyniło ją trafniejszą w wielu przypadkach. Jednak jej działanie nie jest w pełni transparentne, co stanowi wyzwanie dla jej akceptacji jako metody naukowej. Wykorzystanie różnych metod analizy sentymentu podkreśla znaczenie zaawansowanych narzędzi cyfrowych w badaniach socjologicznych. Artykuł ukazuje, jak nowoczesne technologie, takie jak NLP i AI, mogą wspierać jakościową analizę treści, oferując nowe możliwości badawcze. Włączenie metod analizy półautomatycznej i sztucznej inteligencji do badań nad sentymentem pokazuje, jak humanistyka cyfrowa może ewoluować i jakie korzyści może przynieść badaczom. Wnioski z eksperymentu wskazują na konieczność dalszych badań nad integracją tych metod w analizach jakościowych. Artykuł zaprasza do dyskusji na temat transparentności i efektywności różnych podejść, podkreślając, że każda z metod ma swoje mocne i słabe strony.

Artykuł Piotra Cichockiego i Mariusza Baranowskiego pt. *Modelowanie tematyczne w socjologii na przykładzie dobrobytu społecznego: wyzwania metodologiczne i komponent ludzki* analizuje adekwatność socjologicznych metodologii analizy danych w kontekście dynamicznie rozwijających się technologii sieciowych i humanistyki cyfrowej. Tekst stawia pytania o sens tradycyjnego podziału na metody ilościowe i jakościowe w obliczu *Big Data*, wskazując, że ilościowe metody, takie jak probabilistyczne modele statystyczne, mogą stanowić integralną część analiz jakościowych. Artykuł ilustruje tę tezę poprzez modelowanie tematyczne (ang. *topic modelling*) oparte na metodzie LDA (ang. *Latent Dirichlet Allocation*). Przeprowadzono analizę 17 278 artykułów dotyczących dobrobytu społecznego z lat 1992–2020, publikowanych w czasopiśmie indeksowanych w bazie Web of Science. Wyniki analizy pozwoliły na wyznaczenie tematów w obrębie tych artykułów, co posłużyło do sformułowania uwag metateoretycznych na temat kohezji metod ilościowych i jakościowych w perspektywie uczenia maszynowego i przetwarzania języka naturalnego (NLP). Wprowadzenie do artykułu podkreśla znaczenie ilościowej analizy treści jako metody badawczej w naukach społecznych, umożliwiającej redukcję różnorodnych komunikatów tekstowych do kluczowych składników znaczeniowych. Historia tej metody sięga początków akademickich nauk społecznych, kiedy to w 1910 roku Max Weber zarysował projekt „socjologii prasy”. Artykuł wskazuje na rosnące znaczenie ilościowej analizy treści w dobie cyfryzacji i dostępności ogromnych zbiorów danych tekstowych oraz na rewolucyjny postęp technologii NLP. Główna część artykułu koncentruje się na modelowaniu tematycznym za pomocą LDA, które pozwala na identyfikację wspólnych tematów w korpusie tekstowym i oszacowanie ich

rozmieszczenia w dokumentach. Analizę przeprowadzono metodą *Structured Topic Modelling* w programie R. Wyniki można odnieść do metadanych (rok publikacji, liczba cytowań, klasyfikacja dyscyplinarna), co pozwala wnikliwie badać strukturę dyskursu. Autorzy podkreślają rolę narzędzi cyfrowych, takich jak NLP i AI, we wspieraniu jakościowej analizy treści w badaniach socjologicznych. Przykład analizy dobrobytu społecznego ukazuje, jak algorytmy modelowania tematycznego mogą wspierać badania nad złożonymi zjawiskami społecznymi. Artykuł kończy się refleksją nad przyszłością metod analizy treści w naukach społecznych, sugerując, że tradycyjny podział na metody ilościowe i jakościowe może zaniknąć w obliczu rosnących możliwości algorytmicznych technologii przetwarzania danych. Wskazuje na konieczność integracji probabilistycznych modeli statystycznych z klasycznymi metodami jakościowymi, co pozwala na efektywne badanie dużych zbiorów danych tekstowych i otwiera nowe perspektywy dla badań socjologicznych.

Artykuł Sławomira Mandesa i Agnieszki Karlińskiej pt. *W stronę nowej metodologii analizy treści. Podobieństwa i różnice pomiędzy modelowaniem tematycznym i jakościową analizą treści* przedstawia krytyczną refleksję na temat relacji pomiędzy jakościową analizą tematyczną a modelowaniem tematycznym (ang. *topic modeling*) – jedną z popularniejszych metod automatycznego przetwarzania tekstu. Autorzy, bazując na wynikach jakościowej i ilościowej analizy dokumentów Konferencji Episkopatu Polski, analizują wady i zalety modelowania tematycznego oraz negatywnie weryfikują tezę o zastępowalności analizy tematycznej przez modelowanie tematyczne, wskazując na konieczność połączenia podejścia jakościowego z ilościowym w ramach metodologii metod mieszanych (ang. *mixed methods*). W artykule przedstawiono rozwój cyfrowych metod analizy tekstu, które są coraz łatwiej dostępne dzięki rozwojowi takich narzędzi jak WordSmith, AntConc, SketchEngine, QDA Miner, WordStat oraz pakietów do analizy tekstów w językach programowania Python i R. Jednakże autorzy zauważają, że nauki społeczne dołączyły do tego trendu stosunkowo późno, co powoduje pewne dysproporcje w literaturze i metodologii. Głównym celem artykułu jest krytyczna ocena modelowania tematycznego i jego zastosowania w naukach społecznych. Autorzy przedstawiają wyniki analizy dokumentów Konferencji Episkopatu Polski (KEP) za pomocą jakościowej analizy tematycznej (AT) oraz modelowania tematycznego (MT) z wykorzystaniem algorytmu LDA (ang. *Latent Dirichlet Allocation*). Analiza obejmowała 181 dokumentów KEP z lat 1990–2005, które zostały podzielone na trzy typy tekstów: listy pasterskie, odezwy i komunikaty. Wyniki jakościowej analizy tematycznej wykazały, że dokumenty KEP poruszały 94 tematy główne i 294 tematy szczegółowe, z dominującymi tematami dotyczącymi religii, spraw społecznych i polityki. Wykorzystując MT, autorzy zidentyfikowali 25 tematów, które również zostały przypisane do kategorii religijnych, społecznych i politycznych. Mimo pewnych podobieństw w wynikach obu metod artykuł podkreśla różnice i ograniczenia każdej z nich. Modelowanie tematyczne pozwala na pracę z dużymi zbiorami danych i identyfikację wzorców współwystępowania słów, co może być trudne do osiągnięcia w analizie manualnej. Jednakże analiza jakościowa dostarcza głębszego kontekstu i zrozumienia materiału, co jest niezbędne do prawidłowej interpretacji wyników MT. Autorzy podkreślają, że obie metody mają swoje zalety i wady, dlatego optymalne podejście badawcze powinno łączyć obie techniki w ramach metodologii mieszanej. Artykuł kończy się propozycjami czterech sposobów integracji MT z AT: 1) wykorzystania MT do walidacji jakościowej analizy treści, 2) zastosowania MT w etapie konceptualizacji badań

eksploracyjnych, 3) wykorzystania MT do selekcji materiału do analizy oraz 4) użycia MT do generalizacji wyników badań jakościowych. W ten sposób możliwe jest wzbogacenie warsztatu badawczego i rozszerzenie zakresu badań, co przyczynia się do bardziej kompleksowej i efektywnej analizy treści. Autorzy wskazują na potencjał zaawansowanych technik cyfrowych w badaniach społecznych, podkreślając znaczenie integracji metod jakościowych i ilościowych dla osiągnięcia bardziej wszechstronnych i rzetelnych wyników badawczych.

Przyszłość metod humanistyki cyfrowej w socjologii jakościowej niesie ze sobą wiele obiecujących możliwości. Rozwój nowych technologii, interdyscyplinarność oraz większa dostępność narzędzi analitycznych mogą znacząco wzbogacić badania społeczne. Jednocześnie konieczne jest zwrócenie uwagi na wyzwania etyczne i edukacyjne, aby zapewnić odpowiednie przygotowanie badaczy do korzystania z tych metod. W miarę jak technologie będą się rozwijać, socjologia jakościowa będzie miała coraz większe możliwości do zgłębiania i zrozumienia złożonych zjawisk społecznych. Przyszłość metod humanistyki cyfrowej w socjologii jakościowej jawi się jako obiecująca głównie dzięki ciągłemu rozwojowi technologii i narzędzi analitycznych. Jednym z kluczowych kierunków rozwoju jest postępująca automatyzacja procesów badawczych. Narzędzia do analizy tekstu i *Big Data* stają się coraz bardziej zaawansowane, oferując precyzyjniejsze i bardziej szczegółowe analizy. Sztuczna inteligencja (AI) i uczenie maszynowe (ML) mają potencjał, aby zrewolucjonizować te procesy, umożliwiając identyfikację wzorców i zależności trudnych do zauważenia przy użyciu tradycyjnych metod. W najbliższych latach możemy spodziewać się wprowadzenia nowych technologii, które znacząco wpłyną na badania socjologiczne. Przykładem są narzędzia do analizy multimediów, pozwalające na automatyczne kodowanie i analizę treści wizualnych i dźwiękowych, takich jak filmy, nagrania audio czy zdjęcia. Te technologie mogą otworzyć nowe możliwości badawcze w analizie komunikacji wizualnej i mediów społecznościowych. Ponadto rozwój technologii VR (ang. *Virtual Reality*) i AR (ang. *Augmented Reality*) może przynieść innowacyjne rozwiązania dla badań jakościowych. Rzeczywistość wirtualna może być wykorzystana do symulacji i analizy interakcji społecznych w kontrolowanych warunkach, natomiast AR może pomóc w zbieraniu i analizie danych w rzeczywistych kontekstach społecznych i kulturowych. Jednym z najważniejszych aspektów przyszłości humanistyki cyfrowej w socjologii jakościowej jest interdyscyplinarność. Współpraca między różnymi dziedzinami nauki, takimi jak informatyka, psychologia, lingwistyka czy sztuka, może prowadzić do tworzenia innowacyjnych metod i narzędzi badawczych. Integracja wiedzy z różnych dyscyplin pozwoli na bardziej holistyczne podejście do badania zjawisk społecznych.

Jednakże wraz z rozwojem nowych technologii i metod cyfrowych pojawiają się nowe wyzwania etyczne. Badacze muszą zwracać szczególną uwagę na kwestie prywatności i zgody na przetwarzanie danych, zwłaszcza w kontekście *Big Data* i analizy mediów społecznościowych. Wymaga to rozwijania i aktualizacji kodeksów etycznych oraz procedur, które będą chronić uczestników badań i ich dane. Aby w pełni wykorzystać potencjał metod humanistyki cyfrowej, konieczne jest zapewnienie odpowiedniego szkolenia i edukacji dla badaczy. Programy edukacyjne na uniwersytetach i w instytucjach badawczych powinny uwzględniać kursy z zakresu analizy danych, programowania, AI oraz etyki w badaniach cyfrowych. Dzięki temu przyszłe pokolenia socjologów będą lepiej przygotowane

do korzystania z nowoczesnych narzędzi i metod badawczych. W przyszłości kluczowym aspektem będzie również zwiększenie dostępności narzędzi i technologii humanistyki cyfrowej. Obecnie wiele zaawansowanych narzędzi analitycznych jest kosztownych i wymaga specjalistycznej wiedzy. Ułatwienie dostępu do tych narzędzi poprzez tworzenie bardziej przyjaznych użytkownikowi interfejsów oraz oferowanie szkoleń i wsparcia technicznego może znacząco wpłynąć na rozwój badań socjologicznych. Przyszłość metod humanistyki cyfrowej w socjologii jakościowej niesie ze sobą zarówno ogromne możliwości, jak i wyzwania. Kluczem do sukcesu będzie umiejętne łączenie nowych technologii z tradycyjnymi metodami badawczymi przy jednoczesnym zachowaniu etycznych standardów i rozwijaniu kompetencji badaczy.

Bibliografia

Bail Christopher A. (2014), *The cultural environment: Measuring culture with big data*, „Theory and Society”, vol. 43(3–4), s. 465–482.

Boyd Danah, Crawford Kate (2012), *Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon*, „Information, Communication and Society”, vol. 15(5), s. 662–679.

Brosz Maciej, Bryda Grzegorz, Siuda Piotr (2017), *Od redaktorów: Big Data i CAQDAS a procedury badawcze w polu socjologii jakościowej*, „Przegląd Socjologii Jakościowej”, t. XIII, nr 2, s. 6–23, <https://doi.org/10.18778/1733-8069.13.2.01>

Bryda Grzegorz (2024), *Cyfrowa socjologia jakościowa. Odkrywanie językowego obrazu świata życia codziennego*, Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Bryda Grzegorz, Costa António Pedro (2023), *Qualitative Research in Digital Era: Innovations, Methodologies and Collaborations*, „Social Sciences”, vol. 12(10), 570, <https://doi.org/10.3390/socsci12100570>

Burdick Anne, Drucker Johanna, Lunenfeld Peter, Presner Todd, Schnapp Jeffrey (2012), *Digital Humanities*, Cambridge: MIT Press.

Denzin Norman K., Lincoln Yvonna S. (2017), *The SAGE handbook of qualitative research*, Thousand Oaks: Sage Publications.

Dieng Adji B., Ruiz Francisco J.R., Blei David M. (2020), *Topic modeling in embedding spaces*, „Transactions of the Association for Computational Linguistics”, vol. 8, s. 439–453.

Drucker Johanna (2013), *Graphesis: Visual forms of knowledge production*, Cambridge: Harvard University Press.

Fielding Nigel G., Lee Raymond M. (2021), *Qualitative data analysis: Integrating theory, method, and computer-assisted strategies*, Thousand Oaks: Sage Publications.

Gold Matthew K. (2012), *Debates in the Digital Humanities*, Minneapolis: University of Minnesota Press.

Jockers Matthew L. (2013), *Macroanalysis: Digital methods and literary history*, Urbana: University of Illinois Press.

Lupton Deborah (2020), *Data Selves: More-than-Human Perspectives*, Cambridge: Polity Press.

Marres Noortje (2017), *Digital sociology: The reinvention of social research*, Cambridge: Polity Press.

Moretti Franco (2013), *Distant Reading*, London–New York: Verso Books.

Rosenzweig Roy (2003), *Scarcity or Abundance? Preserving the Past in a Digital Era*, „American Historical Review”, vol. 108(3), s. 735–762, <https://dx.doi.org/10.1086/ahr/108.3.735>

Salganik Matthew J. (2019), *Bit by Bit: Social Research in the Digital Age*, Princeton: Princeton University Press.

Svensson Patrick (2016), *Big Digital Humanities: Imagining a Meeting Place for the Humanities and the Digital*, Ann Arbor: University of Michigan Press.

Underwood Ted (2017), *Distant Horizons: Digital Evidence and Literary Change*, Chicago: University of Chicago Press.

Cytowanie

Grzegorz Bryda (2024), *Wprowadzenie. Metody humanistyki cyfrowej w socjologii jakościowej*, „Przegląd Socjologii Jakościowej”, t. XX, nr 4, s. 6–17, <https://doi.org/10.18778/1733-8069.20.4.01>

The Methods of Digital Humanities in Qualitative Sociology

Abstract: Qualitative sociology is undergoing significant changes and seizing new opportunities due to technological advancements such as artificial intelligence, machine learning, and natural language processing. These technologies are revolutionizing qualitative data analysis and offering new research and methodological perspectives. Artificial intelligence and machine learning automate data analysis, manage large datasets efficiently, and lead to more precise and rigorous conclusions while preserving the depth of qualitative research. Integrating the methods of digital humanities with qualitative sociology holds great promise for gaining better insights into complex social phenomena and processes. This “Qualitative Sociology Review” special issue highlights how the methods from within digital humanities can be effectively utilized in qualitative research and data analysis, providing theoretical and practical guidance for researchers. The articles in this issue offer analytical expertise, inspiring researchers to further explore and implement innovative methods of digital humanities and technologies in qualitative sociology.

Keywords: qualitative sociology, qualitative data analysis, digital humanities, artificial intelligence, natural language processing, digital qualitative sociology