

Aleksandra Powierska

ANALIZA ZAWARTOŚCI

PROFILI SPOŁECZNOŚCIOWYCH



Skrypt opracowano przy wsparciu finansowym Priorytetowego Obszaru Badawczego Heritage w ramach Programu Strategicznego Inicjatywa Doskonałości w Uniwersytecie Jagiellońskim.

Analiza zawartości profili społecznościowych

Analiza zawartości mediów jest metodą powszechnie używaną przez badaczy medialnych ekosystemów. Jedną z najczęściej wymienianych w opracowaniach metodologicznych definicji opisuje ją bardzo ogólnie jako „technikę badawczą służącą obiektywnemu, systematycznemu i ilościowemu opisowi jawnej zawartości komunikacji”¹. Analizy zawartości (AZ) nie należy jednak utożsamiać z analizą treści, której zakres jest znacznie węższy, co zaznacza Małgorzata Lisowska-Magdziarz:

Decydując się na AZ przyjmujemy, że badaniu poddane będą nie tylko treściowe, lecz także formalne aspekty badanego materiału. Nie tylko zatem treść słów i obrazów, lecz także kompozycja formalna, struktura wewnętrzna itp.; a także sposób wydrukowania (lub wypowiedzenia), umiejscowienie w badanej gazecie lub w strukturze programu [...] ².

Celem tego skryptu nie jest jednak szczegółowe opisywanie procedur badawczych związanych z analizą zawartości ani jej historii jako metody. Zostały one już dobrze opracowane w literaturze metodologicznej, na przykład przez cytowaną powyżej autorkę. Opracowanie to ma na celu przede wszystkim zwrócenie uwagi na ważne kwestie związane z zastosowaniem tej metody w badaniach mediów społecznościowych. Serwisy społecznościowe charakteryzuje płynna natura, trudno zatem np. o jednoznaczne określenie umiejscowienia wpisu w strukturze serwisu, gdy struktura ta jest strumieniem, w którym rozkładem treści steruje algorytm. Co więcej, możliwość edycji, usuwania czy ukrywania treści w trakcie ich archiwizacji może być również dużym wyzwaniem dla badaczy, podobnie jak ogromne ilości danych i wybór najlepszej metody doboru próby. Zautomatyzowane narzędzia i programy do selekcji treści nie zawsze bowiem są metodologicznie uzasadnionym rozwiązaniem. To tylko wybrane problemy, które stawiają przed medioznawcami liczne pytania o to, jak zaprojektować analizę zawartości serwisów społecznościowych lub konkretnych profili. Odpowiedzi można szukać w doświadczeniach innych badaczy.

Dylemat doboru próby: algorytm czy człowiek?

Adekwatny dobór próby jest kluczem do sukcesu każdej analizy zawartości. Niedoszacowanie pewnych elementów lub nadreprezentacja innych mogą prowadzić do nierzetelnego

1. B. Berelson, *Content Analysis in Communication Research*, Nowy Jork 1952, s. 18.

2. M. Lisowska-Magdziarz, *Analiza zawartości mediów: przewodnik dla studentów: wersja 1.1*, „Zeszyty Wydziałowe”, nr 1, 2004, s. 15.

wnioskowania. Hwalbin Kim, S. Mo Jang, Sei-Hill Kim, Anan Wan w artykule pt. *Evaluating Sampling Methods for Content Analysis of Twitter Data*³ opisują proces doboru próby w analizie zawartości Twittera i zwracają uwagę na zalety próbkowania tradycyjnego w miejsce zautomatyzowanego – opartego na *Big Data* oraz algorytmach. To ostatnie, zdaniem autorów, wiąże się z ryzykiem pominięcia wielu ciekawych treści ze względu na ich niestandardowy charakter, który może nie wpisywać się w zdefiniowane ogólnie przez program kryteria obliczeniowe. Mając na uwadze specyfikę mediów społecznościowych, badacze dokonują ewaluacji dwóch sposobów doboru próby i testują prosty dobór losowy oraz metodę tygodnia składanego⁴.

Badania Kim i wsp. dotyczyły wpisów opublikowanych na Twitterze w 2014 roku i tematycznie nawiązujących do wyborów gubernatorskich w Karolinie Południowej. Tweety⁵ zostały dobrane na podstawie słów kluczy z bazy archiwalnej firmy zewnętrznej. To ważny aspekt badań, gdyż zbiór treści nie zawierał spamu lub postów opublikowanych przez boty. Ostatecznie populacja liczyła 1896 tweetów udostępnionych w ciągu 107 dni (autorzy zdecydowali się na wykluczenie z wybranego do badania okresu dni, w których opublikowano mniej niż 5 wpisów). Następnie dwóch koderów⁶ zakodowało cały materiał, biorąc pod uwagę m.in. takie zmienne, jak zawartość we wpisie linku do innych stron internetowych, status wpisu (oryginalny tweet czy retweet⁷), zawartość materiałów wizualnych, temat postu. Porównując odsetek obecności podstawowych zmiennych w tweetach w całej populacji, a następnie w wybranych różnymi metodami próbach, badacze mogli ocenić reprezentatywność prostego doboru losowego i metody tygodnia składanego. Z przeprowadzonego testu wynika, że do analizy zawartości Twittera bardziej efektywnym sposobem próbkowania jest losowy dobór próby. Według autorów wynika to przede wszystkim z charakteru samego serwisu, a szerzej – mediów społecznościowych⁸.

3. H. Kim, S. Mo Jang, S-H. Kim, A. Wan, *Evaluating Sampling Methods for Content Analysis of Twitter Data*, „Social Media + Society”, vol. 4 (2), April-June 2018, dok. elektr. <https://journals.sagepub.com/doi/full/10.1177/2056305118772836>, [dostęp: 5.02.2023].

4. Ibidem.

5. Tweet – nazwa krótkiego wpisu opublikowanego na Twitterze przez użytkownika serwisu. W 2023 roku maksymalna liczba znaków, jaką może wykorzystać internauta do skonstruowania wypowiedzi, wynosi 280. Do każdego wpisu można dodać materiały graficzne – obrazy lub filmy.

6. Koder to osoba, która analizuje treści według przyjętego klucza kategoryzacyjnego, co nazywa się procesem kodowania.

7. Retweet – to inaczej *tweet* „podany dalej” przez innego użytkownika, który nie jest autorem oryginalnego wpisu, a jedynie udostępnia go na swoim profilu. Retweet może zawierać dodatkowy komentarz osoby udostępniającej.

8. H. Kim i in., op. cit.

Badacze, porównując media społecznościowe z mediami tradycyjnymi, w których metoda tygodnia składanego w doborze próby jest dość często stosowana, wskazują na kilka istotnych dyferencji. Po pierwsze, serwisy społecznościowe w zasadzie nie mają żadnych ograniczeń w odniesieniu do liczby nadawców i ilości publikowanych treści (odrębną kwestią jest ich forma oraz temat, które podlegają ocenie moderatorskiej). Po drugie, trudno mówić o cykliczności mediów społecznościowych. Podczas gdy w badaniach radia, prasy czy telewizji często za jednostkę analizy obiera się jeden dzień, w wypadku platform społecznościowych takie podejście może prowadzić do błędnych analiz i wniosków. Serwisy społecznościowe są nieprzewidywalne i nie działają według planów wydawniczych. Poszczególni twórcy mogą co prawda układać strategie i harmonogram publikacji, ale nie są w stanie przewidzieć, co z zamieszczonymi treściami zrobią pozostali użytkownicy⁹. Nie oznacza to jednak, że w badaniach mediów społecznościowych metoda tygodnia składanego jest całkowicie bezużyteczna. Dobrze sprawdza się bowiem w niereprezentatywnych analizach porównawczych.

W 2015 roku podczas konferencji we Freiburgu wygłosiłam referat pt. *Polish and British News Television on Facebook: Differences and Similarities* przedstawiający wyniki analizy porównawczej komunikacji stacji TVN24 i BBC na Facebooku¹⁰. W wypadku tego badania metoda tygodnia składanego okazała się trafnym wyborem, gdyż moim celem było zbadanie, o czym informują na swoich profilach obie stacje i w jakim stopniu prezentowane treści są zbieżne, a w jakim odmiennie tematycznie. Moim celem nie była analiza konkretnego wątku czy dynamiki publikacji, a jedynie porównanie narracji, typów i liczby zamieszczanych wpisów. To uzasadnia wybór dnia jako jednostki analizy. W badaniach treści rozproszonych, czyli zamieszczanych na różnych profilach i w różnych interwałach czasowych, dzień nie powinien być jednak jednostką analizy, gdyż może dojść do zawyżenia lub niedoszacowania skali obserwowanych zjawisk, w szczególności gdy próba jest niewielka.

Wybór badanej populacji i dobór próby w analizie zawartości mediów społecznościowych będzie zawsze problematyczny. Idealna sytuacja to taka, gdy badacz ma pełny dostęp do badanych profili i może stworzyć listę opublikowanych treści, z których losowo wybierze jednostki analizy do próby. To jednak nie zdarza się zbyt często. Nawet w wypadku wyżej opisanego badania dotyczącego Twittera, nie można mieć pewności, że baza tweetów była wyczerpująca – istnieje prawdopodobieństwo, że wśród postów dotyczących wyborów w Karolinie Południowej znajdowały się i takie wpisy, które nie zawierały zdefiniowanych słów kluczowych, a tym samym nie zostały włączone przez program do zbioru (populacji), z którego następnie dobrano próbę. Z uwagi na to istotne jest,

9. Ibidem.

10. Zob. International Conference Freiburg, 1–4.12.2015, dok. elektr. <https://freiburg2015.sched.com/list/descriptions/>, [dostęp: 8.02.2023].

aby w raporcie z badań szczegółowo opisywać ogół zastosowanych procedur i przede wszystkim wykazywać ich ograniczenia.

Analiza zawartości – jedna metoda czy kilka metod?

Podstawą klasycznej analizy zawartości jest proces kodowania. Im bardziej szczegółowy klucz kategoryzacyjny, tym większa szansa na wskazanie mniej oczywistych korelacji między zdefiniowanymi zmiennymi. Podstawą jego stworzenia jest prawidłowe poznanie materiału badawczego. Pod tym względem analiza zawartości jest podobna do netnografii – bardzo ważne jest rzetelne „rozpoznanie terenu”. Bez zrozumienia internetowych gatunków i konwencji komunikacyjnych, stylów wypowiedzi konkretnych grup, a także mechaniki samego serwisu, z którego pozyskujemy treści do badania, ułożenie wyczerpującego klucza kategoryzacyjnego może okazać się problematyczne. Tym samym istnieje ryzyko, że część zjawisk związanych z analizowanym fenomenem cyberkultury zostanie pominięta lub błędnie zinterpretowana. Jest to szczególnie istotne, kiedy nie kodujemy zebranego materiału samodzielnie, ale decydujemy się na współpracę z innymi koderami. Z jednej strony, takie rozwiązanie stwarza warunki większej obiektywności oraz rzetelności badania. Z drugiej, wymaga bardzo dobrego przygotowania i przeszkolenia osób kodujących.

Brandon Miller przeprowadził analizę zawartości ośmiu profili transgenderowych influencerów na YouTube¹¹. Badaczka interesował przede wszystkim zakres tematyczny publikowanych filmów oraz ich walor edukacyjny. Z uwagi na fakt, że nie istnieje żadna oficjalna lista transgenderowych youtuberów, autor zastosował arbitralny dobór próby, wybierając konta do analizy na podstawie trzech artykułów internetowych, w których wymieniono najbardziej wpływowych twórców. Spośród 12 przywołanych profili do próby trafiło 8 – liczących najwięcej subskrybentów. Następnie z każdego kanału wybrano 15 filmów z najwyższą liczbą wyświetleń. Do procesu kodowania Miller zrekrutował dwie osoby, które zostały przeszkolone w zakresie pojawiającej się w analizowanych materiałach terminologii. Na tej podstawie stworzono księgę kodową. Jednostkę analizy stanowił jeden film. Warto jednak zaznaczyć, że zamieszczane na YouTube posty wideo to często dłuższe i rozbudowane formy wypowiedzi, w których pojawia się wiele wątków. Z uwagi na to badacz zdecydował się na bardzo szczegółowe zdefiniowanie kategorii kodowania motywów pojawiających się w badanych filmach. Dla przykładu, wśród sprecyzowanych

11. B. Miller, *YouTube as Educator: A Content Analysis of Issues Themes, and the Educational Value of Transgender-Created Online Videos*, „Social Media + Society”, vol. 3 (2), April-June 2017, dok. elektr. <https://journals.sagepub.com/doi/full/10.1177/2056305117716271>, [dostęp: 9.02.2023].

zmiennych dotyczących kwestii transgenderowych znalazły się m.in. takie kategorie tematyczne, jak coming out, randki, rodzina, religia, zdrowie psychiczne, stereotypy¹².

Można zastanawiać się, na ile badanie Millera jest analizą zawartości, a na ile zbliża się swym charakterem do analizy treści. Autor nie bierze bowiem pod uwagę dyskusji, które toczą się pod zamieszczonymi przez influencerów materiałami wideo. Nie bada również konwencji filmowych czy czasu, w którym zostały one opublikowane. Dla porównania, analiza kampanii „Man up”, przygotowana przez australijski zespół badawczy, prezentuje znacznie szerszy zakres interpretacji¹³. Przedmiotem badań zespołu była kampania społeczna zachęcająca mężczyzn do szukania pomocy w wypadku problemów natury psychicznej. W ramach działań nagłaśniających temat wyemitowano 3-częściowy dokument TV oraz prowadzono działania narracyjne w Internecie (na oficjalnej stronie internetowej projektu oraz w mediach społecznościowych). Głównym celem analizy było zbadanie odbioru kampanii przez widownię, a przede wszystkim charakterystyka dyskusji toczonych wokół publikowanych materiałów oraz sposobu postrzegania przez odbiorców kwestii związanych z męskością i samobójstwami. By uzyskać odpowiedzi na postawione pytania badawcze, naukowcy przeprowadzili analizę danych pozyskanych z oficjalnego profilu kampanii zamieszczonego w serwisie Facebook. Treści (posty oraz komentarze) zostały zarchiwizowane za pomocą narzędzia informatycznego. Jednostką doboru próby był post, a jednostką analizy – komentarze, które pojawiły się pod danym wątkiem. Tematy dyskusji, będące podstawą kodowania, zidentyfikowano na drodze analizy motywów pojawiających się w głównych wpisach i założenia, że komentarze użytkowników będą ich kontynuacją oraz podczas wstępnego kodowania, obejmującego 5% próby, przeprowadzonego w celu identyfikacji nowych wątków. Podczas kodowania każdy wpis użytkownika mógł zostać przypisany maksymalnie do dwóch kategorii tematycznych¹⁴. Po zakodowaniu całego materiału przeprowadzono analizę ramową¹⁵, czyli uporządkowano poszczególne wątki w nadrzędne ramy: ogólne opinie na temat kampanii „Man up”, dyskusje dotyczące definiowania męskości w kulturze współczesnej oraz wątek samobójstwa i zdrowia psychicznego. Wobec tych trzech obszarów narracyjnych

12. Ibidem.

13. M. Schlichthorst i in., *Using Social Media Network to Engage Men in Conversations on Masculinity and Suicide: Content Analysis of Man Up Facebook Campaign Data*, „Social Media + Society”, vol. 5 (4), October-December 2019, dok. elektr. <https://journals.sagepub.com/doi/full/10.1177/2056305119880019>, [dostęp: 9.02.2023].

14. Ibidem.

15. Temat analizy ramowej w medioznawstwie ciekawie rozwija Krzysztof Wasilewski, ukazując jej różne zastosowania i uwzględniając przemiany tekstów medialnych pod wpływem rozwoju nowych technologii. Zob. K. Wasilewski, *Framing i analiza ramowa – stan badań we współczesnym medioznawstwie. Przegląd stanowisk badawczych*, „Media – Kultura – Komunikacja Społeczna”, t. 1, nr 14, 2019, s. 91–106.

określono również wydźwięk komentarzy w kategoriach pozytywny / negatywny / konstruktywny¹⁶.

Powyższa analiza obejmuje szeroki zakres danych, bada nie tylko sam temat komentarzy, ale także ich wymowę, na podstawie których można wnioskować o postawach odbiorców. Analiza zawartości jest metodą, która otwiera wiele możliwości interpretacyjnych. Często też łączona jest z innymi metodami lub stanowi ich uzupełnieniem. Jak każda metoda ma jednak swoje ograniczenia. Na te ostatnie zwraca uwagę Agnieszka Calek, która porównuje instagramowe profile Anny Lewandowskiej i Roberta Lewandowskiego – pary polskich celebrytów cieszących się dużą popularnością w mediach społecznościowych. Autorka wychodzi od badania netnograficznego, które pozwala jej zapoznać się ze specyfiką i dynamiką badanych kont, następnie przeprowadza analizę zawartości, którą w kolejnym kroku uzupełnia analizą dyskursu¹⁷. Włączenie tej ostatniej metody do procesu badawczego Calek argumentuje ograniczeniami, jakie niesie ze sobą analiza zawartości:

[...] analiza zawartości pozwoliła na dość precyzyjne opisanie materiałów zamieszczanych na badanych profilach. Dzięki jej wynikom można ustalić ogólny obraz, wskazać działania najbardziej typowe i popularne. Gdyby jednak ograniczyć analizę materiałów tylko do badania tą metodą, jej słabością w tym kontekście okazałoby się, że działania mniej typowe i wyjątkowe kodowane byłyby zwykle w kategorii „inne” w celu uniknięcia nadmiernego rozbudowania klucza. A bez nich obraz badanego przypadku nie będzie pełny. Stąd ważne jest, aby takie właśnie przypadki materiałów wizualnych (tj. treści skrajne i nietypowe) poddać analizie dyskursu, której wyniki dopełnią zakres uzyskanych wcześniej rezultatów¹⁸.

Efektywność analizy zawartości będzie zawsze zależała od tematu i celu badania oraz zakresu dostępnych danych. Netnografia może być etapem wstępnym badań, ponieważ pozwala badaczowi nie tylko zapoznać się z materiałem do analizy, ale także zrozumieć, w jaki sposób on powstaje i czym kierują się jego twórcy. Umożliwia również rozpoznanie kodów kulturowych danej społeczności internetowej i podjęcie decyzji o ewentualnym włączeniu innych metod do procesu badawczego. Wszystko to jest ważne przy tworzeniu klucza kategoryzacyjnego i pozwala na zdefiniowanie wyczerpujących kategorii kodowania. Dotychczas opisywane badania koncentrowały się na analizach, w których

16. M. Schlichthorst i in., op. cit.

17. A. Calek, *Analiza obrazu. Studium przypadku Anny i Roberta Lewandowskich – pierwszej pary Instagrama*, [w:] *Metody badań medioznawczych i ich zastosowanie*, A. Szymańska, M. Lisowska-Magdziarz, A. Hess (red.), Kraków 2018, s. 361–378.

18. Ibidem, s. 372.

materiał opracowywali koderzy, gdyż jego ilość nie przekraczała ludzkich możliwości. Przy ogromnych zbiorach danych badacze coraz częściej sięgają po zautomatyzowane metody, które znacznie przyspieszają zarówno zbieranie, jak i późniejszą analizę danych.

Lena Clever i in. wybrali multimodalną zautomatyzowaną analizę zawartości opartą na studium przypadku, aby opisać, w jaki sposób islamscy ekstremiści wykorzystują Instagram do działań propagandowych¹⁹. W pierwszej kolejności za pomocą aplikacji Instamancer badacze stworzyli bazę treści opublikowanych na instagramowym profilu „Generation Islam” obejmującą trzy lata aktywności. Następnie w obrębie zebranych postów wydzielili trzy zbiory: bazę hashtagów, bazę obrazów oraz bazę tekstów (w tym tekstów wyodrębnionych z obrazów dzięki użyciu specjalistycznego oprogramowania). Z uwagi na taki podział sama analiza również składała się z trzech etapów: analizy sieci hashtagów, wizualnej analizy sentymentu oraz analizy tekstu. Do wszystkich metod wykorzystano odpowiednie oprogramowanie analityczne. Ostatnim etapem badania było wyliczenie korelacji pomiędzy poszczególnymi zmiennymi²⁰. Badania Clever i wsp. pokazują zasadność wykorzystywania zautomatyzowanych narzędzi w wypadku obszernych baz danych. Ręczne kodowanie zawartości profilu na Instagramie z trzech lat zajęłoby dużo czasu lub wymagałoby powołania licznego zespołu przeszkolonych koderów. Decydując się na wykorzystanie narzędzi informatycznych, trzeba mieć jednak na uwadze, że nawet najlepsze aplikacje mogą natrafić na problem z rozpoznaniem ironii czy sarkazmu lub sytuacji, w których użytkownicy tekst traktują subwersywnie.

Podsumowanie

Analiza zawartości mediów społecznościowych w centrum stawia teksty kultury, które są różnorodne zarówno pod względem technicznym, jak i semiotycznym. Ich znaczenie będzie wyprowadzane nie tylko z samej struktury, ale także z miejsca i kontekstu publikacji. W mediach społecznościowych wybrzmiewa to w sposób szczególny. Jeden mem może mieć wiele interpretacji w zależności od kontekstu, w którym zostanie opublikowany i komentarzy, którymi opatrzą go użytkownicy. Nie jest to oczywiście żadna nowość. Umieszczenie programu telewizyjnego w ramówce również wpływa na jego odczytanie przez widzów, jednak w mediach społecznościowych wszystkie procesy są znacznie bardziej dynamiczne i płynne, a każdy tekst może bardzo szybko stać się tekstem multimodalnym. Małgorzata Lisowska-Magdziarz, pisząc o wielomodalności zauważa, że

19. L. Clever i in., *Behind Blue Skies: A Multimodal Automated Content Analysis of Islamic Extremist Propaganda on Instagram*, „Social Media + Society”, vol. 9 (1), January-March 2023, dok. elektr. <https://journals.sagepub.com/doi/full/10.1177/20563051221150404>, [dostęp: 9.02.2023].

20. Ibidem.

współczesne teksty medialne „składają się z połączonych różnymi zależnościami obrazów nieruchomych i ruchomych, wizualnej komunikacji nieobrazowej, tekstu pisanego, mówionego, dźwięków, muzyki”²¹.

Wydaje się zatem, że największym wyzwaniem analizy zawartości jest zdolność badanych tekstów do obrastania kolejnymi warstwami wskutek łatwości ich edycji. Dwie analizy zawartości przeprowadzone od podstaw na tym samym materiale (czyli z uwzględnieniem ponownej archiwizacji w wypadku drugiego kodowania), ale w określonym odstępie czasowym, mogą przynieść zupełnie inne wyniki, bo w czasie między jednym a drugim badaniem użytkownicy dodadzą nowe komentarze, udostępnią analizowane teksty innym użytkownikom, usuną lub zmienią swoje wypowiedzi. Drugą istotną kwestią jest fakt, że w wielu przypadkach analiza zawartości obejmuje tylko treści widoczne dla badacza. W przygotowywanej ręcznie bazie danych nie znajdują się zatem wpisy ukryte²² lub usunięte przez administratorów profili oraz te, o których losie zdecydował algorytm. Niekompletność bazy danych, lub raczej brak pewności co do jej kompletności, to niewątpliwie jedno z największych ograniczeń analizy zawartości w mediach społecznościowych, które należy uwzględnić już na etapie planowania procesu badawczego. W gruncie rzeczy analiza zawartości mediów społecznościowych to analiza widocznej i dostępnej zawartości mediów społecznościowych.

W praktyce

1. Zdefiniuj cel badawczy i pytania badawcze.
2. Zastanów się, jak obszerny materiał jest potrzebny do analizy, która przyniesie odpowiedzi na postawione pytania badawcze.
3. Przemyśl, czy analiza zawartości będzie metodą wystarczającą do osiągnięcia celu badawczego (rozpocznij „rozeznanie terenu”, czyli wstępne badania netnograficzne w celu zapoznania się z materiałem, oraz wypisz wszystkie możliwe ograniczenia; rozważ też, czy i w jaki sposób możesz je zminimalizować poprzez włączenie innych metod).
4. Sporządź bazę danych w najkrótszym możliwym czasie, aby w trakcie jej tworzenia uniknąć znaczących zmian w obrębie danego serwisu, które mogą wpłynąć

21. M. Lisowska-Magdziarz, *Badanie wielomodalnych przekazów w mediach masowych*, [w:] *Metody badań medialnożnanawczych i ich zastosowanie*, A. Szymańska, M. Lisowska-Magdziarz, A. Hess (red.), Kraków 2018, s. 145.

22. Facebook umożliwia administratorom stron biznesowych ukrywanie wybranych komentarzy fanów. Po ukryciu wpis jest nadal widoczny dla jego autora oraz jego znajomych, ale nie widzą go pozostali użytkownicy.

na wcześniej zarchiwizowane treści (np. dodanie nowej reakcji na Facebooku może spowodować, że użytkownicy zaczną jej używać również przy starszych postach, które wyświetli im algorytm).

5. Zdecyduj, czy będziesz kodować samodzielnie czy za pomocą narzędzi analitycznych (pamiętaj o księdze kodowej – nawet jeśli kodujesz samodzielnie, dokładne zdefiniowanie kategorii w kluczu jest niezbędne).
6. Przeprowadź badania pilotażowe i sprawdź poprawność arkusza kategoryzacyjnego, a następnie skoryguj ewentualne błędy oraz przystąp do badania właściwego.