

Analiza zmiany postrzegania Youtuberów w związku z #PandoraGate

Opis problemu

Afera #PandoraGate umownie rozpoczęta filmem opublikowanym 3.10.2023 dotycząca podejrzenia zachowań pedofilskich w środowisku największych polskich twórców YouTube wstrząsnęła polskimi mediami społeczeństwymi. Jak uległ zmianie wizerunek zaangażowanych w aferę osób?



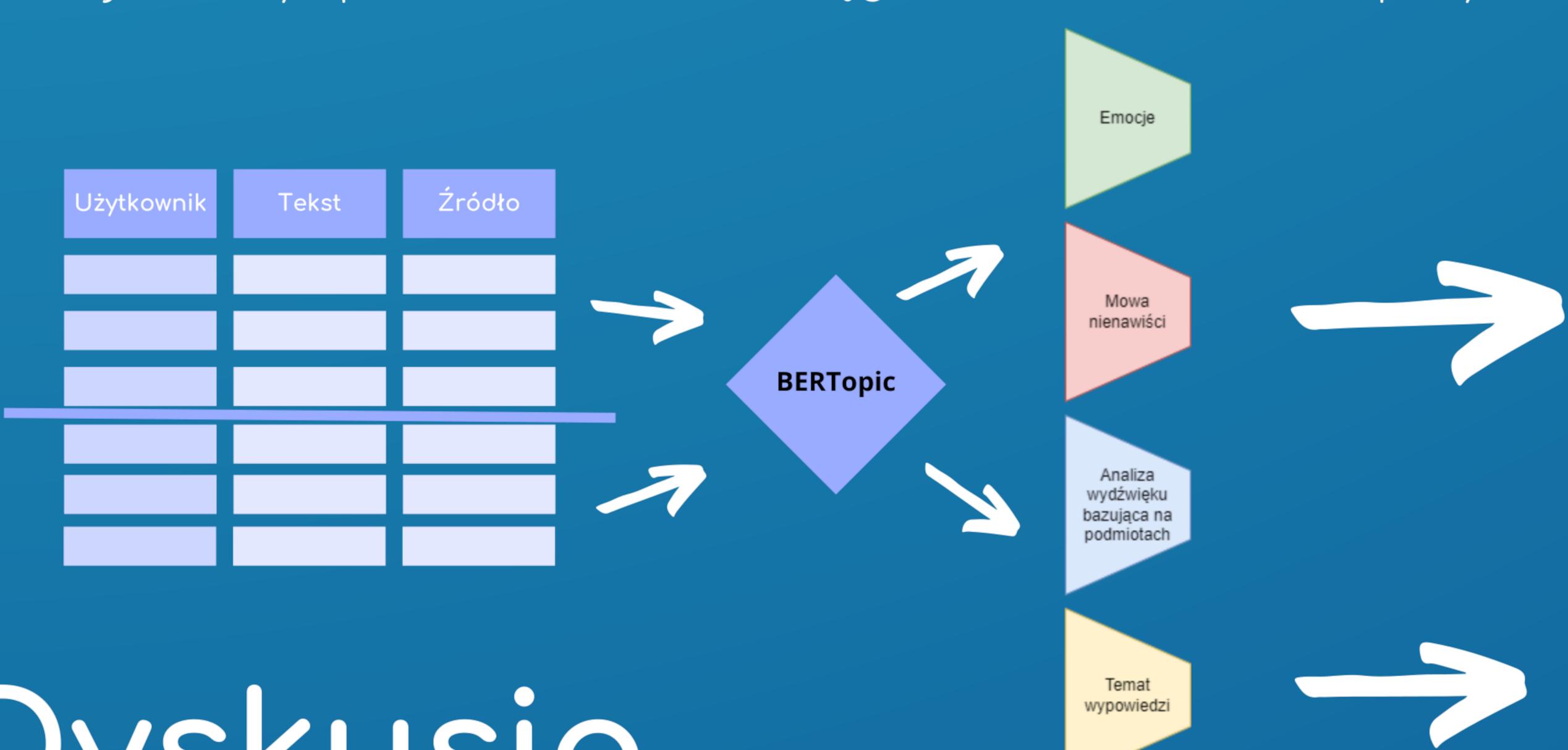
Metoda

Oba zestawy danych zostały poddane analizie pod kątem tematu, emocji, mowy nienawiści i analizy wydźwięku w stosunku do podmiotów.

Wykorzystane zostały modele:

- bardsai/twitter-emotion-pl-base
- dkleczek/Polish-Hate-Speech-Detection-Herbert-Large
- OpenAI GPT-4
- BERTopic

GPT-4 został wykorzystany do analizy sentymentu w oparciu o wyekstrahowane podmioty. Następnie dokonaliśmy selekcji podmiotów do dalszej analizy spośród: Dubiel, Wardęga, Stuu, Boxdel, Konopsky.



Dyskusja

Dane w mediach społecznościowych zawierają bardzo różnorodne informacje. Wiele z zebranych wpisów nie dotyczy tematu lub jest dyskusją między użytkownikami mediów. W związku z tym zastosowane rygorystyczne odfiltrowanie wiadomości typu spam oraz wpisów niezwiązanych z tematem analizy spowodowało znaczne zmniejszenie objętości danych i zwiększenie możliwości ich badania. Zastosowanie analizy w czasie mogłoby znacząco podnieść wartość projektu, jednakże byłaby ona nieistotna statystycznie przy posiadanej próbce danych. Ponadto istotnym kierunkiem badawczym byłaby analiza sieci społecznych i grup w nich występujących - w szczególności zmian cech użytkowników względem badanych podmiotów przed i po #PandoraGate. Dodatkowo modele analizy emocji i mowy nienawiści nie spełniły założonych kryteriów jakościowych, zatem wyniki uzyskane przy jego pomocy nie miały wpływu na analizę.

Wnioski

W analizie możemy wyróżnić dwie grupy postaci - "dziennikarzy": Wardęga i Konopsky, którzy opublikowali materiały obciążające "oskarżonych"****: Stuu, Dubiel, Boxdel.

Wpisy w zebranych danych z mediów społecznościowych przeważająco charakteryzują się wydźwiękiem negatywnym oraz neutralnym. Wpisy neutralne w dużej części posiadają charakter informacyjny, nie zawierają opinii autora.

Przesunięcie wydźwięku w kierunku neutralnego i pozytywnego dla dziennikarzy pozwala przypuszczać, że w wyniku ich działań, nastąpiła znacząca pozytywna zmiana w polaryzacji społeczności w stosunku do nich.

Na podstawie otrzymanych wyników nie można stwierdzić o zmianie negatywnej względem Stuu oraz Boxdela. Jedyną postacią wobec której polaryzacja zmieniała się negatywnie, w niewielkim stopniu, jest Dubiel.

Technologie

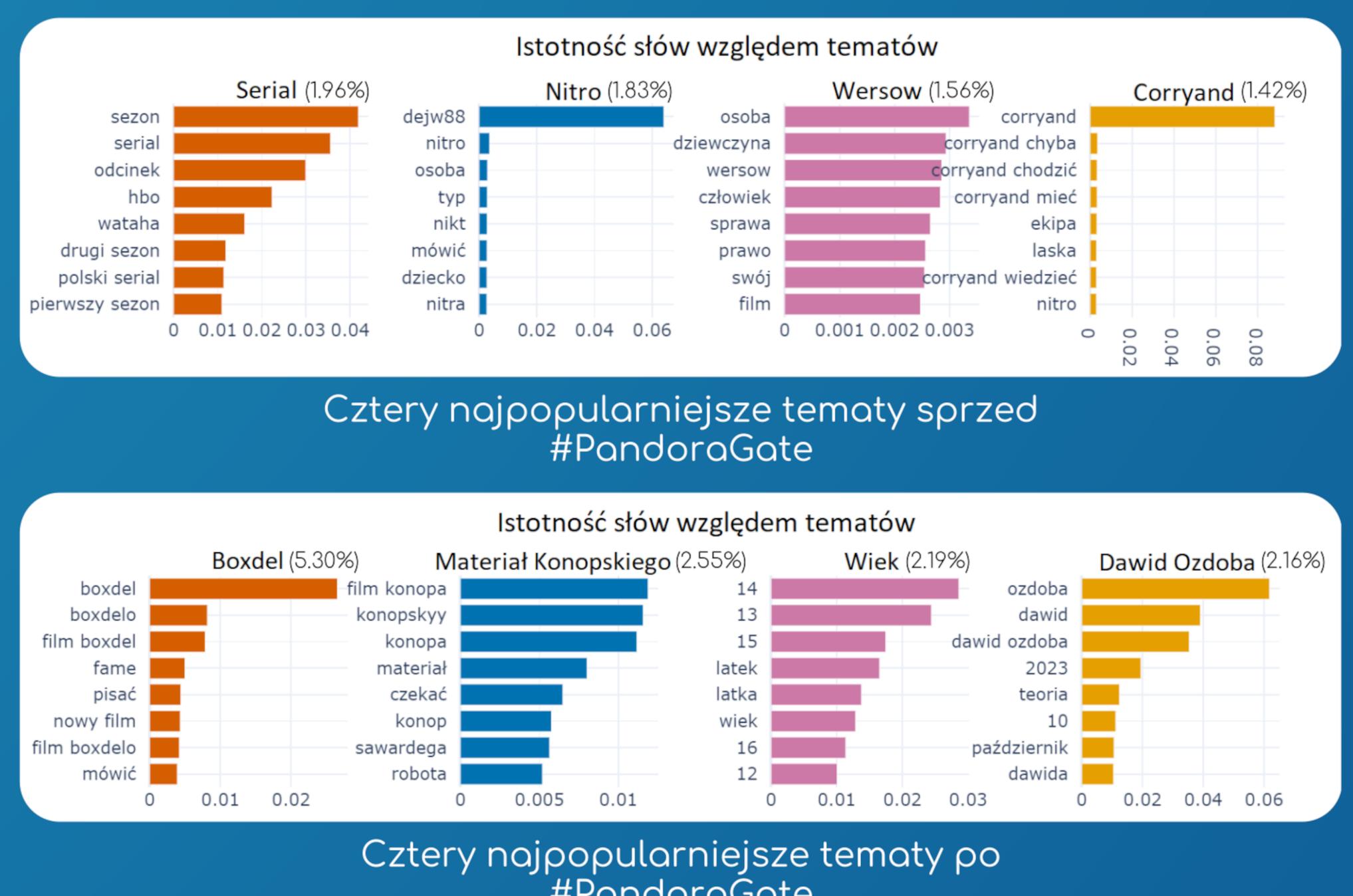


Dane

Zebraliśmy wpisy i komentarze osób pod filmami bezpośrednio związane z #PandoraGate oraz powiązanymi tagami w mediach: YouTube, Twitter oraz Wykop. Analizowaliśmy je jako dwa zestawy - posty/komentarze/tweety opublikowane:

1. Przed 3.10.2023 - przed filmem #PandoraGate (67333 wpisów).
2. Po 3.10.2023 - po filmie #PandoraGate (70329 wpisów).

BERTopic użyty do modelowania tematycznego umożliwił wyłonienie populacji danych do dalszej analizy: 41.8% próbek spośród danych sprzed #PandoraGate i 51.3% próbek spośród danych po #PandoraGate



Wyniki

Zbadaliśmy zmianę wydźwięku w stosunku do badanych podmiotów względem wydarzenia #PandoraGate



*Na moment wykonywania projektu nie zostały ogłoszone żadne prawomocne wyroki w tej sprawie.

**Analiza i wyciągnięte wnioski mogą zawierać obciążenia wynikające z jakości analizy automatycznej i wykorzystanych modeli.

***Na chwilę produkcji plakatu formalne postępowanie sądowe toczy się jedynie wobec Stuarta "Stuu" Bartona.