

# Jakie emocje wzbudzają polskie serwisy informacyjne na Twitterze?

Magda Boruch, Jakub Dziwiński, Katarzyna Jabłońska, Dominika Kunc  
*Wydział Informatyki i Zarządzania, Politechnika Wrocławska, Wrocław, Polska*

## Streszczenie

Portale mediów informacyjnych w Internecie cieszą się dużym zainteresowaniem wśród Polaków. Są one źródłem najnowszych informacji zarówno z kraju, jak i ze świata. Czytając takie artykuły, oczekujemy przede wszystkim braku stronniczości oraz rzetelnego przekazania informacji. W Polsce istnieje wiele portali informacyjnych, jednak są takie, które wyróżniają się większym zainteresowaniem czytelników niż inne. Każde z nich ma zarówno swoich zwolenników, jak i przeciwników. Zdecydowaliśmy się porównać cztery, naszym zdaniem najbardziej popularne portale informacyjne pod kątem tego, jak nacechowane są publikowane przez nie treści oraz jakie emocje wywołują u swoich czytelników.

## 1 Wstęp

Media oraz serwisy społecznościowe są nieodłączną częścią świata, który nas otacza. Można pokusić się o stwierdzenie, że w dzisiejszych czasach te dwa elementy są jednymi z głównych czynników kształtujących zdania ludzi na różne tematy. Jest to tak zwana teoria porządku dziennego (z angielskiego *Agenda-setting*). Niestety w natłoku informacji panującym w internecie może stać się tak, że treści docierające do użytkowników są uzależnione od tego, jakie profile dany użytkownik śledzi.

W idealnym świecie media informacyjne powinny przedstawiać informacje o danych wydarzeniach w sposób obiektywny. Aby zachować wiarygodność informacji, media te nie powinny stawiać po żadnej stronie w konfliktach, a treści dostarczane przez nie powinny być opisane możliwie dokładnie, tak aby nie pominąć istotnych szczegółów. Mimo neutralnego stylu pisania i przedstawiania informacji, zagrożenie niedoinformowaniem może wywodzić się też z rodzaju publikowanych treści - nie każde media informują o wszystkich wydarzeniach.

Zjawisko to nazywane jest filtrowaniem dostępu informacji (z angielskiego *gatekeeping*). Polega ono na wyborze informacji, które będą udostępniane przez dane źródło. Przede wszystkim jest to konieczne ze względu na ograniczoną przepustowość kanałów informacyjnych - np. czas antenowy, czy koszty produkcji tradycyjnych mediów. W XXI wieku nie jest to ogromnym ograniczeniem, ponieważ informacje mogą być publikowane w internecie, który nie ma takich ograniczeń. Czy jednak sprawia to, że media nie dokonują selekcji materiałów, która może być oparta na subiektywnych założeniach danego portalu? Odpowiedź na to pytanie może przynieść analiza artykułów, które media udostępniają na swoich profilach na portalu [Twitter](#).

Większość profili informacyjnych udostępnia artykuły z własnych stron internetowych za pomocą linków w tweetcie

(poście na Twitterze). Spośród najpopularniejszych serwisów informacyjnych w Polsce postanowiliśmy wybrać cztery - [TVN24](#), [TVP Info](#), [Polsat News](#) oraz [Wirtualna Polska](#) i przeanalizować artykuły, które publikują na Twitterze oraz odpowiedzi odbiorców danego profilu do tych postów.

Do analizy wykorzystane zostały dane od dnia 1 września do 22 listopada 2020 roku. W ogólności 2020 rok był burzliwy na całym świecie, jednak w tym okresie w kontekście polskich mediów i poruszenia społecznego działo się jeszcze więcej. W tym okresie pojawiały się informacje dotyczące pandemii koronawirusa, wyborów w Stanach Zjednoczonych, Strajku Kobiet, weta w Unii Europejskiej czy też Święta Niepodległości. Dodatkowymi krokami podjętymi podczas analizy było zgrupowanie artykułów w podgrupy tematyczne, a następnie wybranie artykułów dotyczących tego samego wydarzenia, co umożliwiło porównanie reakcji emocjonalnych odbiorców różnych profili.

## 2 Dane

Zebranie dużej ilości danych z czterech różnych źródeł byłoby bardzo czasochłonne, gdybyśmy musieli zrobić to ręcznie, dlatego zdecydowaliśmy się na automatyzację tego procesu. Zauważyliśmy, że każdy z wybranych przez nas portali posiada konto na Twitterze, oraz regularnie publikuje posty, które zawierają odnośniki do artykułów na stronie tego portalu, oraz jego krótki opis. Dzięki ogólnodostępnemu narzędziu [Twint](#) byliśmy w stanie pobrać te informacje. Największe znaczenie dla nas, miały zawarte w postach odnośniki do stron www. Mając te informacje, stworzyliśmy cztery osobne programy do pobierania danych ze stron internetowych (ang. *scraper*). Ostatecznie proces polegał więc na pobieraniu danych ze stron portali, do których odnośniki zawarte były w postach z Twittera. Co więcej, aby zbadać reakcję użytkowników Twittera na publikowane posty, również przy użyciu narzędzia Twint zebraliśmy odpowiedzi do poszczególnych tweetów. Liczność zebranych danych przedstawia Tabela 1.

Portal	Liczba tweetów	Liczba odpowiedzi
TVN24	4 681	47 970
TVP Info	7 609	234 701
Polsat News	7 117	46 134
Wirtualna Polska	2 850	4 769
Łącznie	22 257	333 574

Tablica 1: Liczność zebranych danych

## 2.1 Grupowanie artykułów

Aby przeprowadzenie założonych analiz było możliwe, należało dopasować artykuły publikowane na różnych kontach tak, aby dotyczyły tych samych wydarzeń. W tym celu wszystkie zebrane posty z Twittera zawierające artykuły zostały wstępnie pogrupowane na kategorie tematyczne: koronawirus, wybory w Stanach Zjednoczonych, weto w Unii Europejskiej, Strajk kobiet, Święto Niepodległości oraz kategoria inne, dla artykułów z mniej obszernych i łatwo wyróżnialnych tematów. Liczność tych kategorii w zależności od portalu przedstawia Tabelka 2.

Kategoria	TVN24	TVP Info	Polsat News	Wirtualna Polska
Koronawirus	1 046	1 570	1 601	1 071
Weto UE	143	240	116	36
Wybory w USA	331	541	352	97
Święto Niepodległości	6	71	24	32
Strajk Kobiet	196	544	338	260
Inne	2 796	4 184	3 090	1 379

Tabelka 2: Liczność kategorii w zależności od portalu.

Następnie dzięki użyciu specjalistycznych narzędzi do analizy języka naturalnego dostarczonych przez CLARIN-PL (Topic Modelling i ReSpa) artykuły z każdej kategorii zostały wstępnie pogrupowane przy użyciu dwóch oddzielnych algorytmów, aby finalnie jako docelowe zgrupowanie wybrać wyłącznie takie grupy, które zostały wyłonione przez obydwa algorytmy.

Algorytm oparty na Topic Modelling dla treści artykułów z danej kategorii wybiera tematy, których mogą one dotyczyć i każdemu artykułowi przypisuje prawdopodobieństwo przynależenia do danego tematu. Następnie dla każdej pary artykułów z tego samego dnia obliczone zostały odległości cosinusowe między wektorami przynależności do tematów. W tym wariancie artykuły dotyczyły jednego wydarzenia, jeżeli ich miara cosinusowa przekraczała dany próg będący liczbą rzeczywistą z przedziału od 0 do 1. W praktyce progi te miały wartości od około 0.7 do 0.95 w zależności od różnorodności artykułów w danej kategorii. Konieczne było dodatkowo empiryczne ocenienie, czy grupowanie działa prawidłowo.

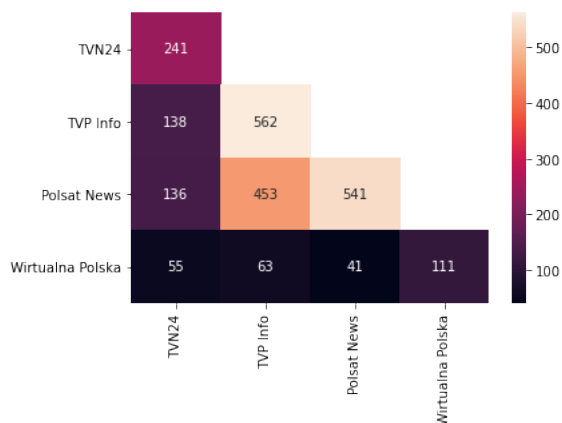
Algorytm ReSpa w ogólności służy do wyznaczania słów kluczowych w tekście. W naszym przypadku wysoka powtarzalność słów kluczowych w artykułach była jednoznaczna z przynależnością do tej samej grupy tematycznej (opisującej to samo wydarzenie).

Warto zaznaczyć, że głównym celem nie jest analiza wszystkich takich artykułów, więc skupiliśmy się na jakości, a nie ilości uzyskanych grup. Finalnie otrzymaliśmy 677 grup z 1865 postami z Twittera. Łącznie dało to 27 931 odpowiedzi użytkowników do zgrupowanych postów. Liczność grup i odpowiedzi użytkowników dla każdej z kategorii zaprezentowane zostały w Tabeli 3.

Kategoria	Liczba grup	Liczba odpowiedzi
Koronawirus	158	5670
Weto UE	31	1809
Wybory w USA	91	2502
Święto Niepodległości	13	1663
Strajk Kobiet	54	7488
Inne	330	8799
Łącznie	667	27 931

Tabelka 3: Liczność grup i odpowiedzi użytkowników dla każdej kategorii.

Liczba analizowanych artykułów dla źródeł TVN24 oraz TVP Info oscylowała w okolicy 3 tysięcy, dla Polsat News było to ponad 4 tysiące. Uwzględniając jedynie artykuły, które zostały przydzielone do grup tematycznych, proporcje te były zupełnie inne. Na rysunku 1 widać, że ilość artykułów pogrupowanych i pochodzących z TVN24 jest dwukrotnie mniejsza od pozostałych dwóch źródeł (Polsat News i TVP Info). Tak znaczna różnica, może oznaczać, że źródło to w jakiś sposób się wyróżniało. Mogło być to spowodowane innym doбором tematyki swoich artykułów albo odmiennego sposobu opisu wydarzeń. Rysunek ten oprócz liczebności artykułów danego źródła w grupach przedstawia także informację o tym, jak często analizowane media pisały o tym samym temacie. Artykuły pochodzące ze źródeł TVP Info oraz Polsat News pokrywały się aż w 80% wyodrębnionych grup. Jeśli chodzi o stację Wirtualna Polska, to pozostałe źródła równie często występowały wraz z nią w tej grupie.



Rysunek 1: Liczba artykułów o tym samym wydarzeniu dla danych źródeł. Wartość na przekątnej oznacza ilość wspólnych artykułów z którymkolwiek innym źródłem.

## 2.2 Sentyment i emocje

Kolejnym krokiem analiz było zajęcie się sentymentem oraz emocjami panującymi w odpowiedziach użytkowników. Sentyment oznacza nacechowanie emocjonalne treści - czy jest ona pozytywna, neutralna czy negatywna. Do oceny sentymentu skorzystaliśmy z modelu sentimentPL. Teksty z Twittera potrafią zawierać wiele mało istotnych informacji z punktu widzenia oceny sentymentu. Dlatego przed przetworzeniem zostały one poddane obróbce. Usunięto z nich

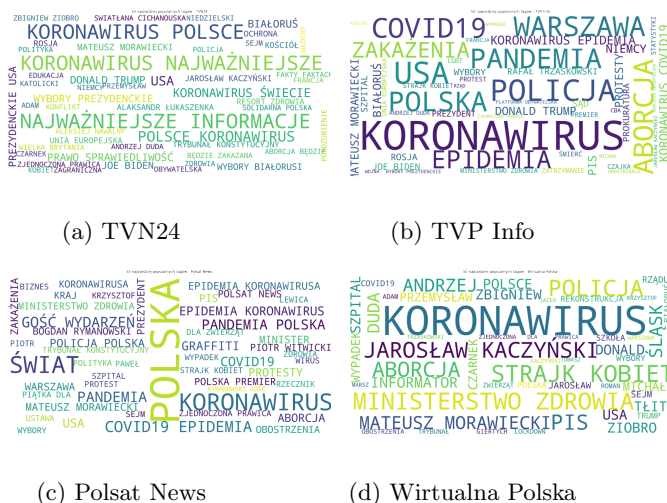
Z kolei jeśli chodzi o emocje zawarte w tekście, to w przeprowadzanych analizach rozważane jest 8 emocji: radość, zaskoczenie, cieszenie się na coś oczekiwanego, strach, smutek, zaufanie, złość i wstręt. Tymi emocjami oznaczane są polskie słowa w [Słowsieci](#) z adnotacjami emocjonalnymi. Każda odpowiedź została zlematyzowana (każde słowo zostało zastąpione swoim lematem, czyli formą słownikową), a następnie została przetworzona pod kątem występujących w niej emocji, w efekcie czego każda odpowiedź powiada średnią wartość punktową dla każdej z 8 emocji na podstawie emocjonalności słów użytych w odpowiedzi. Dodatkowym czynnikiem wspomagającym wyliczanie emocji było zawarte w Słowsieci wartościowanie (np. słaby pozytywny/negatywny, mocny pozytywny/negatywny, neutralny), co umożliwiło skalowanie intensywności emocji.

Ponieważ artykuły zostały zawężone do tych udostępnianych na Twitterze, warto przybliżyć podstawowe informacje dla kont związanych z poszczególnymi stacjami.

Źródło	L. obserwujących	Śr. tweetów	Śr. polubień	Śr. udostępnień	Śr. odpowiedzi
TVN24	1 453 914	57	29	6	8
TVP Info	709 764	92	30	10	26
PolSAT News	93 602	86	13	4	6
Wirtualna Polska	70 333	34	3	1	2

W Tabeli 4 można zauważyć, że wielkości kont na portalu Twitter różnią się zarówno rozmiarem, jak i aktywnością. Najwięcej obserwujących ma TVN24, natomiast to nie to źródło udostępnia największą ilość artykułów dziennie. To TVP Info oraz Polsat News dominują w tej kategorii. Warto tutaj zaznaczyć, że pomimo niewielkiej różnicy w średniej liczbie tweetów na dzień, różnica w liczbie obserwujących jest znaczna. Dodatkowo można zauważyć, że co prawda TVN24 oraz TVP Info mają największą ilościową liczbę polubień - nie ma się jednak co tutaj dziwić, skoro są to dwa najpopularniejsze media z analizowanych. TVP Info wysuwa się również na prowadzenie, jeśli pod uwagę weźmiemy liczbę udostępnień ich postów oraz odpowiedzi do nich.

Jak można zauważyć na Rysunku 2, najbardziej popularne tagi przypisane do artykułów różnią się w zależności od konkretnego portalu. Widać jednak że wszystkie z nich udostępniały sporo newsów dotyczących tych samych tematów. Bez większego zaskoczenia tag koronawirus, wraz z innymi bezpośrednio z nim związanymi, dominuje we wszystkich stacjach. Dodatkowo dosyć powszechne we wszystkich stacjach były artykuły związane ze strajkiem kobiet, wyborami prezydenckimi w USA oraz partią rządzącą, czy z politykami bezpośrednio z nią związanymi.



Rysunek 2: 50 najpopularniejszych tagów w artykułach dla każdego z portali

Rozkład tagów związanymi z konkretnymi tematami nieco różni się w zależności od portalu, ale śmiało, można uznać, że ogólnie piszą one o tych samych lub chociażby powiązanych ze sobą wydarzeniach.

## 4.1 Analiza danych w czasie

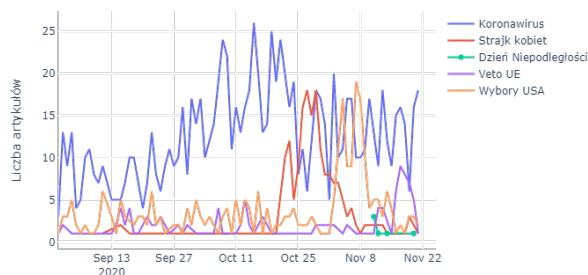
#### 4.1.1 Analiza publikowanych artykułów

Wykresy przedstawione na Rysunku 3 pokazują ile artykułów na dany, ogólnie pojęty temat publikowały poszczególne media. Na pierwszy rzut oka można by uznać, że krzywe reprezentujące konkretne kategorie są dosyć analogiczne dla wszystkich źródeł. Jest to oczywiście wynikiem tego, że gdy ma miejsce jakieś fascynujące wydarzenie, to powstaje wiele kolejnych artykułów o nim, wraz z nowymi informacjami. Przyglądając się powyższym grafikom nieco dokładniej, jesteśmy w stanie powiedzieć, kiedy miejsce miały właśnie takie zdarzenia.

Bez zaskoczenia najwięcej artykułów dotyczyło koronawirusa, a krzywe związane z tym tematem mają podobny kształt dla wszystkich portali. Największy wzrost widać na przełomie września i października, kiedy to było coraz więcej chorych i wprowadzane były coraz to bardziej restrykcyjne obostrzenia.

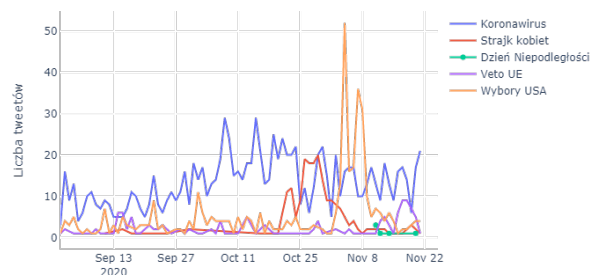
Z kolei na rysunku 4 przedstawiono, ile raz dany portal udostępnił na swoim profilu na Twitterze artykuł z danej kategorii. Niektóre portale częściej udostępniały artykuły z danej kategorii, ale w większości przypadków trendy w udostępnianiu i publikowaniu były podobne dla każdej stacji. Jedyнным wyróżniającym się zjawiskiem była liczba artykułów udostępnionych na Twitterze TVN24 dotyczących wyborów w Stanach Zjednoczonych - znacznie więcej udostępnień w porównaniu do innych tematów. Może to wynikać z tego, że portal ten co jakiś czas udostępniał artykuł aktualizowany o nowe dane związane z wyborami. Różnice te nie zostały wychwycone na wykresie publikowanych artykułów.

Liczba artykułów na dany temat - TVN24



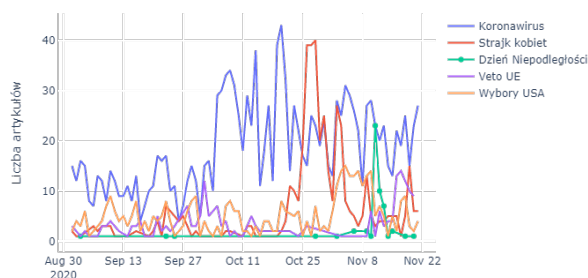
(a) TVN24

Liczba tweetów z artykułami na dany temat - TVN24



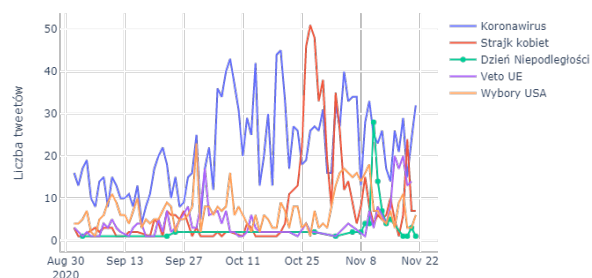
(a) TVN24

Liczba artykułów na dany temat - TVP Info



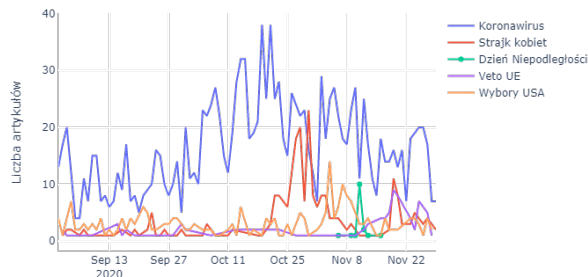
(b) TVP Info

Liczba tweetów z artykułami na dany temat - TVP Info



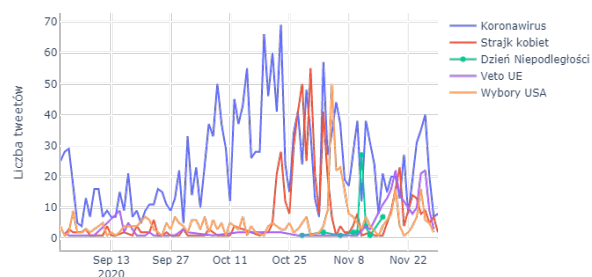
(b) TVP Info

Liczba artykułów na dany temat - Polsat News



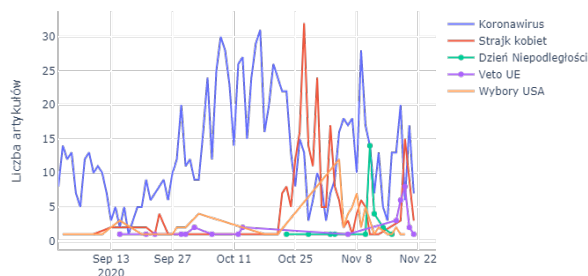
(c) Polsat News

Liczba tweetów z artykułami na dany temat - Polsat News



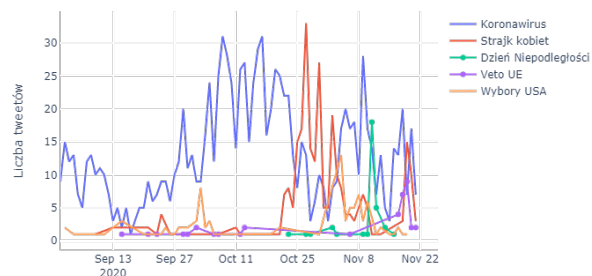
(c) Polsat News

Liczba artykułów na dany temat - Wirtualna Polska



(d) Wirtualna Polska

Liczba tweetów z artykułami na dany temat - Wirtualna Polska



(d) Wirtualna Polska

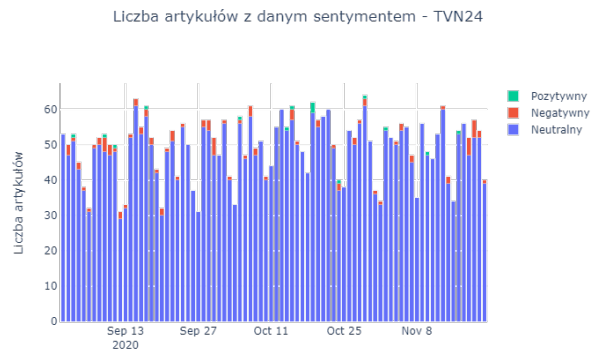
Rysunek 3: Liczba artykułów z danej kategorii w zależności od daty

Rysunek 4: Liczba tweetów z artykułami z danej kategorii w zależności od daty



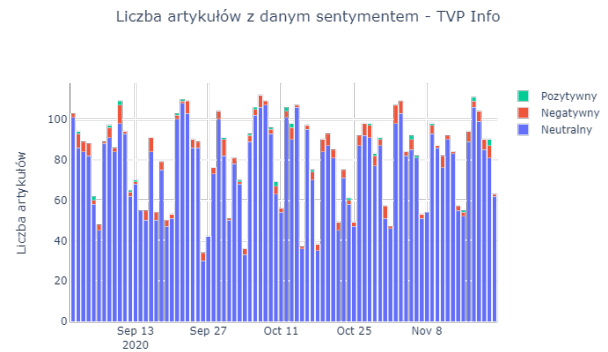
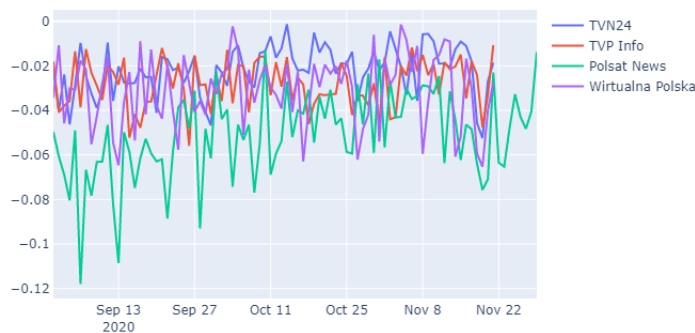
ponieważ w momencie pobierania artykułów każdy z nich zawierał wszystkie zaktualizowane informacje.

Ponadto przeanalizowany został sentyment artykułów publikowanych przez każdy z portali. Jak widać w większości przypadków artykuły, mają neutralny sentyment, czyli nie są nacechowane emocjonalnie. Tego właśnie oczekuje się od dziennikarstwa. Jednak można zauważyć, że pojawia się również niewiele artykułów z nacechowaniem pozytywnym i trochę więcej z nacechowaniem negatywnym. W publikowaniu ujemnie spolaryzowanych treści wybija się Polsat News, co można zauważyć poprzez zwiększoną widoczność koloru czerwonego na Rysunku 6c.



(a) TVN24

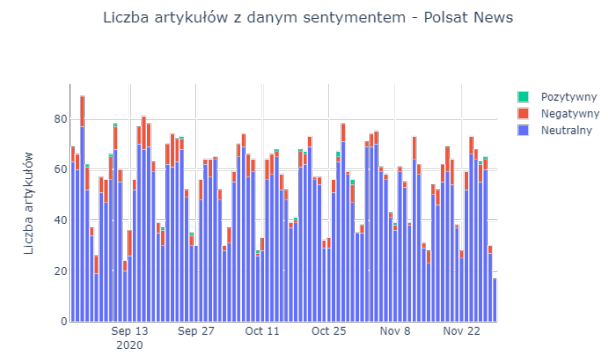
Średni sentyment artykułów



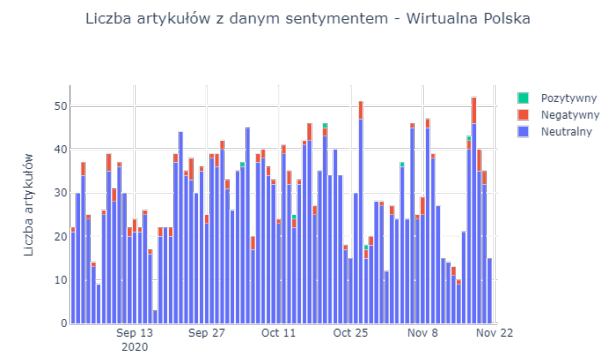
(b) TVP Info

Rysunek 5: Średnia wartość sentymentu artykułów w zależności od daty

Z kolei najwięcej neutralnych treści publikuje TVN24, co potwierdza również wykres przedstawiony na rysunku 5. Widoczne jest również odchylenie średniej dla artykułów z portalu Polsat News w stronę intensywniej negatywnego (ale nadal dosyć neutralnego) nacechowania na początku rozważanego okresu, aby pod koniec zbliżyć się do pozostałych źródeł.



(c) Polsat News



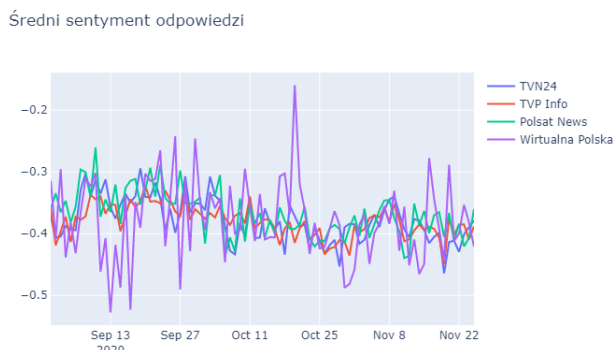
(d) Wirtualna Polska

Rysunek 6: Liczba artykułów z danym sentymentem w zależności od daty

#### 4.1.2 Analiza odpowiedzi odbiorców

Aby zobaczyć, co odbiorcy danych profili wyrażali w swoich odpowiedziach, postanowiliśmy przeanalizować jakie wartości przyjmował sentyment i emocje zawarte w nich względem dni.

Jako pierwszy sprawdzony został średni sentyment odpowiedzi dla każdego z profili (rysunek 7). Średnie dzienne sentymenty mają wartości poniżej zera, co oznacza, że są nacechowane negatywnie. Na wykresie wyróżnia się łamana dla profilu Wirtualna Polska - może wynikać to z tego, że profil ten jest najmniej popularny spośród zestawianych profili, za czym idzie mniejsza liczba odpowiedzi - wystarczy kilka mocno nacechowanych odpowiedzi i średnia przyjmuje bardziej skrajne wartości. Ze względu na różnice w liczności odpowiedzi dla poszczególnych profili zostały zaprezentowane wykresy z wartościami liczbowymi wystąpień odpowiedzi o konkretnym sentymencie. Widoczne wykresy potwierdzają wcześniejsze stwierdzenie - większość odpowiedzi jest nacechowana negatywnie.



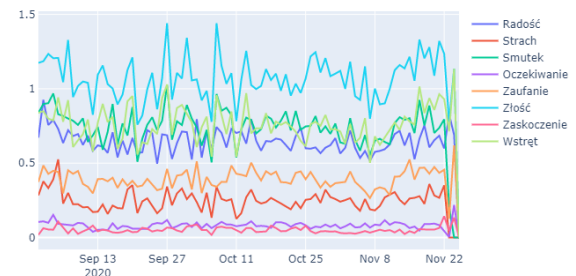
Rysunek 7: Średnia wartość sentymentu w zależności od daty

W przypadku średnich wartości dla emocji w odpowiedziach (rysunek 8) rozkłady wyglądają dość podobnie - występuje najwięcej złości, wstrętu i smutku. Dopiero po nich można dostrzec radość, następnie zaufanie i strach. Oczekiwanie i zaskoczenie występuje najrzadziej dla każdego z profili. Na wykresach dla profilu TVN24 oraz TVP Info po prawej stronie wykresu następuje wyraźny wzrost średniej wartości wstrętu. Może to być spowodowane przez publikację reportażu o sprawie ukrywania skandali w Kościele oraz niedużą liczbą silnie nacechowanych odpowiedzi.

Na rysunku 10 przedstawiono licznosc odpowiedzi z danymi emocjami. Spora część z odpowiedzi jest neutralna, aczkolwiek złość, smutek, wstręt, a także radość i zaufanie są dość wyraźnie widoczne w dziennych zestawieniach.

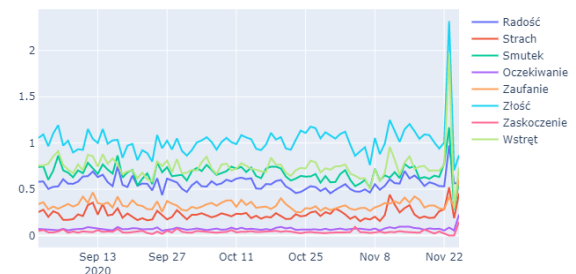
Wspólną charakterystyką odpowiedzi do poszczególnych profili (rysunki 9 i 10) widać miejsca z największą liczbą emocjonalnych odpowiedzi. Dla TVN24, TVP Info oraz Wirtualnej Polski maksimum jest osiągnięte 30 października, a dla Polsat News ten dzień jest drugim najliczniejszym. W tym dniu odbywał się największy protest w ramach strajku kobiet „Marsz na Warszawę”, skąd może wynikać taka liczność nacechowanych odpowiedzi do profili.

Średnia wartość nacechowania emocjonalnego - TVN24



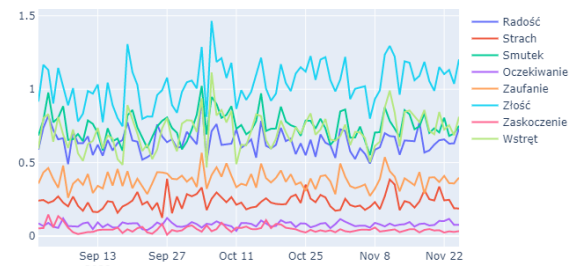
(a) TVN24

Średnia wartość nacechowania emocjonalnego - TVP Info



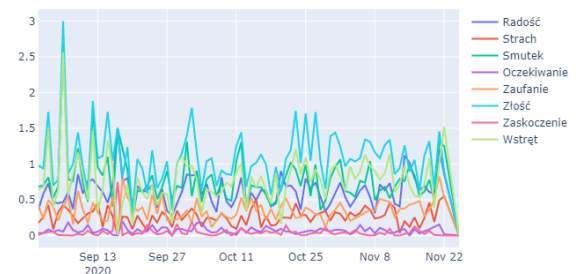
(b) TVP Info

Średnia wartość nacechowania emocjonalnego - Polsat News



(c) Polsat News

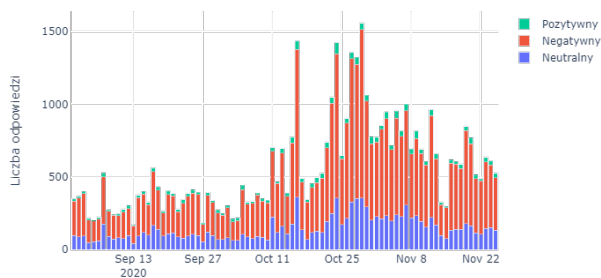
Średnia wartość nacechowania emocjonalnego - Wirtualna Polska



(d) Wirtualna Polska

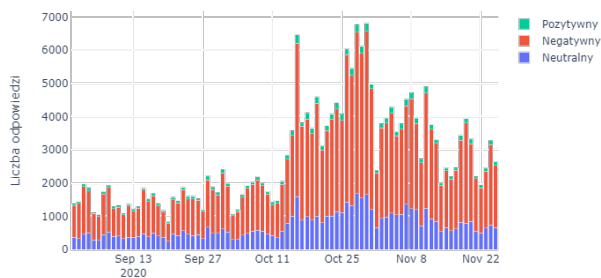
Rysunek 8: Średnia wartość nacechowania dla poszczególnych emocji w zależności od daty

Liczba odpowiedzi z danym sentymentem - TVN24



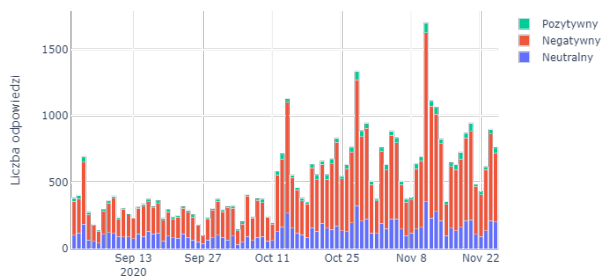
(a) TVN24

Liczba odpowiedzi z danym sentymentem - TVP Info



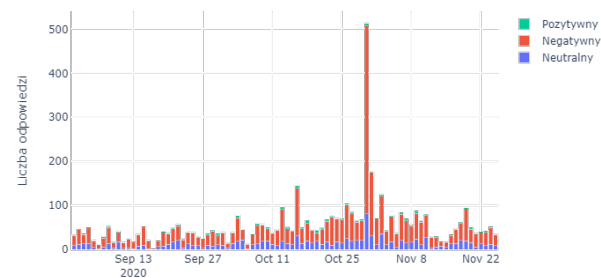
(b) TVP Info

Liczba odpowiedzi z danym sentymentem - Polsat News



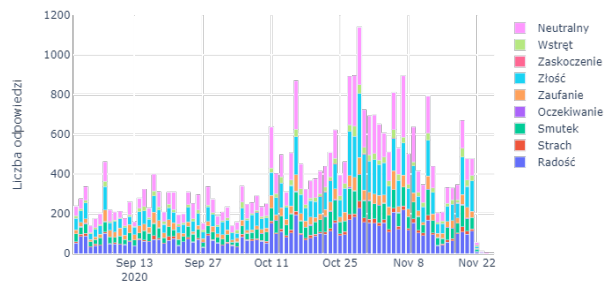
(c) Polsat News

Liczba odpowiedzi z danym sentymentem - Wirtualna Polska



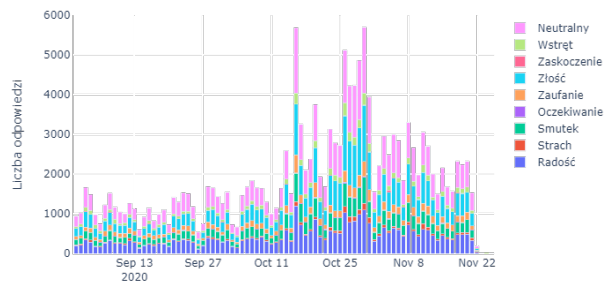
(d) Wirtualna Polska

Liczba odpowiedzi z nacechowaniem emocjonalnym - TVN24



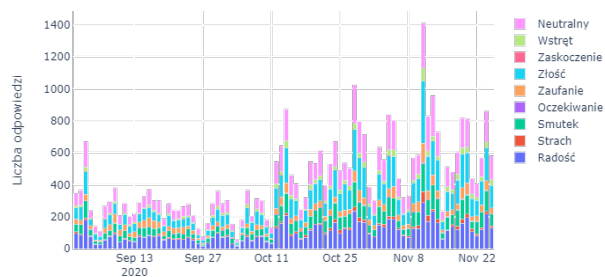
(a) TVN24

Liczba odpowiedzi z nacechowaniem emocjonalnym - TVP Info



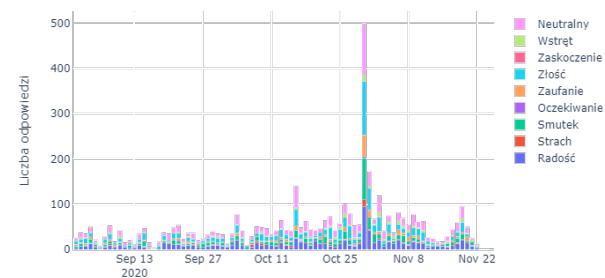
(b) TVP Info

Liczba odpowiedzi z nacechowaniem emocjonalnym - Polsat News



(c) Polsat News

Liczba odpowiedzi z nacechowaniem emocjonalnym - Wirtualna Polska



(d) Wirtualna Polska

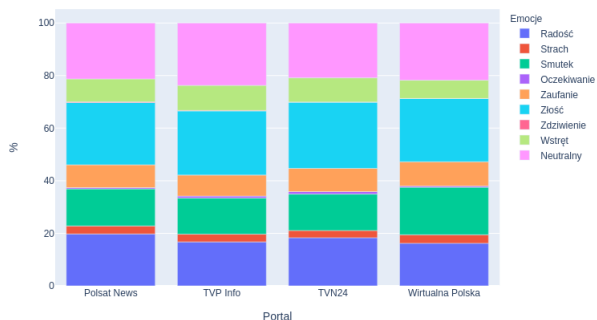
Rysunek 9: Liczba odpowiedzi z danym sentymentem w zależności od daty

Rysunek 10: Liczba odpowiedzi z danym nacechowaniem emocjonalnym w zależności od daty

## 4.2 Ogólne wyniki dla wszystkich odpowiedzi

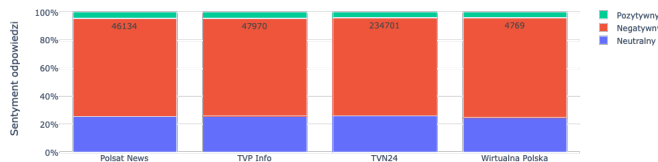
Rozkład emocji zawarty w odpowiedziach (niekoniecznie do postu z artykułem) dla każdego z portali przedstawiony został na Rysunku 11. Można zauważyć, że jest on podobny dla każdej stacji, jednak trzeba mieć na uwadze, że są to statystyki ogólne. Podobnie jest w przypadku sentymentów wszystkich odpowiedzi do danego konta - również są do siebie bardzo podobne, co widać na rysunku 12. Około 70% wszystkich odpowiedzi do źródeł z rozważanego okresu ma sentyment negatywny, niecałe 25% stanowią odpowiedzi neutralne, a odpowiedzi o pozytywnym wydźwięku stanowią zaledwie niespełna 5% wszystkich odpowiedzi. Różne stacje piszą na różne tematy, inne sytuacje mogą bawić i złościć odbiorców. Dokładniejsze analizy, z podziałem na konkretne tematy przedstawione są poniżej. Wybrane zostały te tematy, w których naszym zdaniem najbardziej widoczne są różnice emocjonalne pomiędzy czytelnikami portali.

Emocje we wszystkich odpowiedziach



Rysunek 11: Rozkład emocji w odpowiedziach dla każdego z portali

Sentyment w odpowiedziach - w zależności od źródła



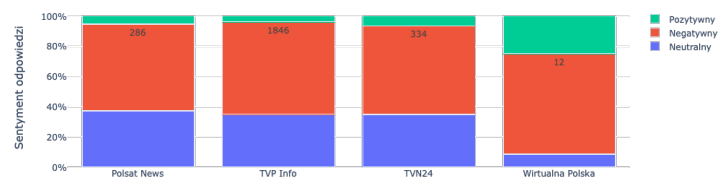
Rysunek 12: Rozkład sentymentów w odpowiedziach dla każdego z portali

## 4.3 Wyniki dla danych podzielonych na kategorie

W przypadku podziału na kategorie brane pod uwagę były odpowiedzi do postów zawierających link do artykułu. W tym przypadku średni odsetek odpowiedzi negatywnych dla wszystkich kategorii i wszystkich źródeł wynosił około 67%, neutralnych niecałe 29%, a pozytywnych troszkę ponad 4%.

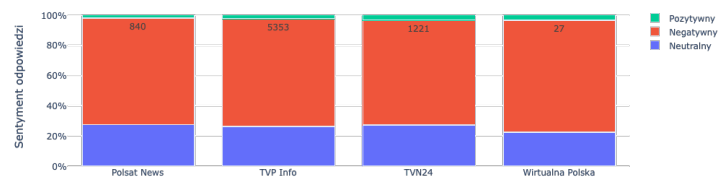
Dla kategorii koronawirus, Święto Niepodległości, weto UE i inne rozkład sentymentu dla każdej ze stacji był zbliżony do wartości średnich. Wartym zaznaczenia mogą być odchylenia mające miejsce w pozostałych trzech grupach. Najwyższy odsetek pozytywnych odpowiedzi został zaobserwowany w kategorii wybory w Stanach Zjednoczonych (rysunek 13 - ponad 10% odpowiedzi miało wydźwięk na plus. Z kolei kategoria strajk kobiet okazała się zawierać niemal 72% negatywnie nacechowanych odpowiedzi. Jak widać na rysunku 14 rozkład sentymentów pozostawał mniej więcej taki sam. Jedynie warto zaznaczyć większy odsetek (około 1 p.p. różnicy) pozytywnych odpowiedzi w TVN24 i Wirtualnej Polsce względem Polsat News i TVP Info.

Wybory w USA



Rysunek 13: Rozkład sentymentów w kategorii wybory w Stanach Zjednoczonych dla każdego z portali

Strajk Kobiet



Rysunek 14: Rozkład sentymentów w kategorii strajk kobiet dla każdego z portali



## 4.4 Wyniki dla danych o konkretnym wydarzeniu

Oprócz ogólnych statystyk interesowało nas również to, jak media wypowiadały się na temat konkretnych wydarzeń. A także to jakie emocje wzbudzały te wydarzenia wśród internautów. Wybraliśmy więc najgorętsze tematy z każdej kategorii. Skupiliśmy się jedynie na sprawach, na które, wypowiedziały się przynajmniej trzy analizowane przez nas źródła. W dalszej części opisaliśmy kilka wybranych tematów. Wybór ten bazował na ilości odpowiedzi do każdego z wydarzeń oraz różnorodności reakcji.

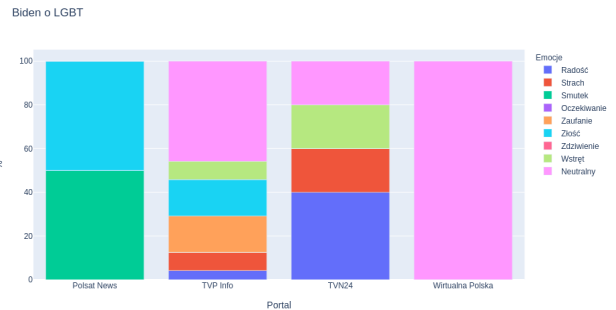
### 4.4.1 Joe Biden o strefach wolnych od LGBT

21 września 2020 roku ówczesny kandydat na prezydenta Stanów Zjednoczonych wypowiedział się na temat społeczności LGBTQ+ i stref wolnych od LGBT zaznaczając, że prawa tej społeczności to prawa człowieka, oraz że ani w Unii Europejskiej, ani gdziekolwiek na świecie, nie ma miejsca na coś takiego jak „strefa wolna od LGBT”. Na Rysunku 15 zamieszczono oryginalny post zamieszczony na portalu Twitter.

Na Rysunku 16 przedstawiono emocje panujące w odpowiedziach do postów na ten temat. Każdy udostępniony artykuł na ten temat miał neutralny sentyment. W porównaniu zostały uwzględnione TVN24, TVP Info oraz Polsat News. Jak można zobaczyć z tych trzech źródeł najwięcej radości znalazło się w odpowiedziach do profilu TVN24, co może być powiązane z preferencjami politycznymi i lewicowymi poglądami. W przypadku Polsat News emocje podzieliły się na dwie części - smutek i złość. Z kolei wśród odbiorców TVP Info najwięcej było odpowiedzi neutralnych, a zaraz po nich znalazły się emocje „zaufanie” oraz „złość”. Znając szerszy kontekst „zaufanie” i „złość” można odnieść do poglądów, z którymi utożsamia się znaczna część odbiorców mediów prawicowych, więc mogą one wynikać z zaufania krajowej partii rządzącej i złości, że zagraniczne głosy próbują ingerować w wartości przez nią reprezentowane.



Rysunek 15: Post autorstwa Joe Bidena dotyczący stref wolnych od LGBT, źródło: twitter.com, 23.01.2021.



Rysunek 16: Emocje wśród odpowiedzi do postów z artykułami na temat stanowiska Joe Bidena w sprawie stref wolnych od LGBT.

### 4.4.2 Trzaskowski o Marszu Niepodległości

Prezydent Warszawy Rafał Trzaskowski nie wyraził zgody na zorganizowanie marszu z okazji Święta Niepodległości. Mimo wszystko marsz jednak odbył się, a Prezydent po zdarzeniach, jakie na nim zaszły (bójki, wandalizm, brak dystansu społecznego) skomentował sprawę negatywnie, opisując środowiska prawicowe.



Rysunek 17: Post autorstwa Prezydenta Warszawy Rafała Trzaskowskiego dotyczący marszu Niepodległości, źródło: facebook.com, 24.01.2021.

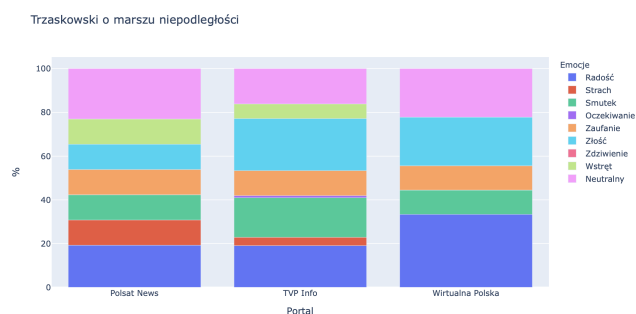
O tym wydarzeniu napisały 3 z rozważanych mediów - Polsat News, TVP Info oraz Wirtualna Polska. Jak wcześniej było zaznaczone, znaczna większość artykułów ma neutralny sentyment. W tym przypadku jednak TVP Info jako jedyne napisało artykuł z sentymentem negatywnym. Może to wynikać, chociażby z wzajemnego negatywnego nastawienia Rafała Trzaskowskiego do TVP czy też Prawa i Sprawiedliwości, które wywołał w swoim poście.

W przypadku reakcji odbiorców danego profilu rozkład emocji w odpowiedziach został przedstawiony na Rysunku 19. Jak widać w odpowiedziach do profilu Polsat News, emocje rozłożyły się równomiernie z delikatną przewagą wypowiedzi radosnych oraz neutralnych. Dla profilu TVP Info

dominującą emocją była złość, następnie smutek i dopiero po nim radość i neutralny, co może być spowodowane negatywnym sentymentem artykułu opublikowanego przez to medium. W przypadku Wirtualnej Polski dominuje radość, a trochę dalej znajdują się emocje neutralna i złość.



Rysunek 18: Sentyment artykułów dotyczących wypowiedzi Rafała Trzaskowskiego na temat Marszu Niepodległości.



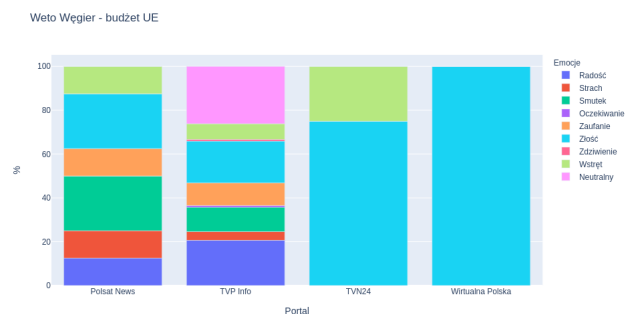
Rysunek 19: Emocje w odpowiedziach do postów zawierających artykuły dotyczące wypowiedzi Rafała Trzaskowskiego na temat Marszu Niepodległości.

#### 4.4.3 Weto Węgier

Weto budżetu unijnego miało dotyczyć przyjęcia mechanizmu warunkowości, który mógłby umożliwić zawieszenie transferów budżetowych w przypadkach, kiedy brak praworządności zagraża interesom finansowym Unii Europejskiej. Polska i Węgry chciały mieć pewność, że rozporządzenie to nie ma charakteru arbitralnego oraz nie wprowadzi małżeństw homoseksualnych. Wszystkie z rozważanych mediów udostępniły artykuły na temat oficjalnego oświadczenia premiera Węgier Viktora Orbana na temat zawetowania budżetu UE, a każdy z nich miał neutralny sentyment.

Wśród odbiorców Wirtualnej Polski wzbudziły wyłącznie złość. W przypadku TVN24 do znacznie przeważającej złości doszedł wstęt. Odbiorcy Polsat News byli najbardziej zezłoszczeni i smutni. Odpowiedzi do tweetów TVP

Info jako jedyne były w największej części neutralne, a na drugim miejscu uplasowała się radość, tuż za nią złość. Z niemalże pełną pewnością można stwierdzić, że odbiorcy TVN24 i Wirtualnej Polski nie popierali zgłaszanego weta, a ponad 20% odbiorców TVP Info ucieszyła się z tych wydarzeń.



Rysunek 20: Emocje wśród odpowiedzi do postów z artykułami na temat weta Polski i Węgier w sprawie budżetu Unii Europejskiej.

## 5 Wnioski

Rozważane portale informacyjne powielają schematy działania w przypadku publikacji artykułów na portalu Twitter. Artykuły, które udostępniają, w większości są neutralne, jednak przedstawione zostały przypadki, w których widać było mocniejszą polaryzację, zazwyczaj w stronę wartości negatywnych.

Ogólne rozkłady sentymentów i emocji, zarówno dla publikowanych artykułów, jak i dla odpowiedzi do nich są bardzo zbliżone dla każdego źródła. Jednak analizując dane w kontekście kategorii tematycznych, czy konkretnych wydarzeń można było zobaczyć, jak zmieniają się emocje. Widoczne były również różnice w emocjach panujących w odpowiedziach kierowanych do każdego źródła, szczególnie w przypadku rozważania pojedynczych wydarzeń.

Wyciągnięcie wniosków na temat każdej z prezentowanych zależności chcielibyśmy zostawić każdemu czytelnikowi do indywidualnej oceny. Polecamy również zapoznanie się z materiałem „Media a poglądy polityczne”, którego autorem jest dr hab. Michał Wenzel, profesor Uniwersytetu SWPS. Ta krótka lektura być może pokaże zależności wcześniej nieznane, lub też utwierdzi w obecnych przekonaniach, ale z pewnością pomoże lepiej zrozumieć przedstawione w tym artykule analizy.

**Podziękowania:** Chcielibyśmy bardzo podziękować osobom, bez których nie udałoby się ten projekt:

mgr inż. Bartoszowi Perzowi - za pomoc w przetwarzaniu dużej ilości danych,

mgr inż. Krzysztofowi Rajdzie - za mentoring i przekazanie wiedzy na temat tajników analizy mediów społecznościowych,

mgr inż. Arkadiuszowi Janzowi - za mentoring i przekazanie wiedzy z zakresu analizy języka naturalnego.