

Wyższa Szkoła Informatyki i Zarządzania z siedzibą w Rzeszowie
Kolegium Mediów i Komunikacji Społecznej

R o z p r a w a d o k t o r s k a

**Badanie zmediatyzowanej komunikacji politycznej
przy pomocy analizy statystycznej
i algorytmów predykcyjnych sztucznej inteligencji**

Arkadiusz Gawel

Promotor

dr hab. Andrzej Adamski, prof. WSiIZ

Promotor pomocniczy

dr Małgorzata Gosek

Rzeszów 2023

Spis treści

Wstęp.....	4
Idea i powody podjęcia pracy	4
Stan aktualnych badań	12
Cel pracy	16
Problemy badawcze	17
Hipotezy badawcze	17
Metodologia i techniki badawcze.....	18
Struktura pracy	21
Rozdział 1. Podstawy teoretyczne pojęć użytych w pracy	23
1.1. Media – definicja i podział.....	23
1.2. Mediatyzacja	36
1.3. Modele komunikacji.....	42
1.4. System społeczny, polityczny i medialny jako środowisko dla systemu komunikowania politycznego	52
1.5. Model systemu komunikowania politycznego i jego uczestnicy	59
1.6. Język.....	66
1.7. Big Data	77
1.8. Kluczowe pojęcia i definicje	81
1.9. Podsumowanie rozdziału	83
Rozdział 2. Mediatyzacja polskiej polityki	84
2.1. Przemiany społeczno-gospodarcze w Polsce w latach po 1989 roku.....	84
2.2. Polityka w latach od 1989 do aktualnego stanu (2021).....	85
2.3. Sytuacja społeczna po 1989 roku	88
2.3. Rozwój mediów w Polsce po 1989 roku.....	102
2.4. Przemiany komunikacji politycznej pod wpływem mediów cyfrowych i rola mediów cyfrowych	114
2.5. Podsumowanie rozdziału	125
Rozdział 3. Część Praktyczna – Metody analizy tekstu	126
3.1. Opis danych.....	126
3.2. Oczyszczenie i przygotowanie danych	127
3.3. Liczba wystąpień według posłów i partii	128
3.4. Liczność słów oraz podział na klastry.....	132
3.5. Porównanie dendrogramów oraz pomiar indeksem Jaccarda	148
3.6. Analiza koszykowa zwrotów występujących razem	159
3.7. Analiza ngramów	185

3.8.	Analiza sentymentów	200
3.9.	Dane z Twittera	222
3.10.	Sentymenty w tweetach.....	246
3.12.	Podsumowanie rozdziału	251
Rozdział 4. Część Praktyczna – Modele predykcyjne.....		253
4.1.	Przygotowanie danych	255
4.2.	Model z wykorzystaniem drzew decyzyjnych – R.....	258
4.3.	Model z wykorzystaniem metody najbliższego sąsiada kNN – R	261
4.4.	Model z wykorzystaniem lasów losowych – R.....	263
4.5.	Model z wykorzystaniem algorytmu wzmacniającego GBM Boosting – R	266
4.6.	Model z wykorzystaniem maszyn wektorów nośnych (SVM) – R.....	268
4.7.	Model z wykorzystaniem algorytmu sieci neuronowych – Python.....	270
4.8.	Porównanie modeli.....	274
4.9.	Testowanie na kilku przypadkach	275
4.10.	Podsumowanie rozdziału	295
5.	Zakończenie	297
5.1.	Weryfikacja hipotez	297
5.2.	Rekomendacje	301
6.	Bibliografia	304
6.1.	Źródła	304
6.2.	Opracowania	307
7.	Streszczenie.....	317

Wstęp

Idea i powody podjęcia pracy

Proces mediatyzacji jest jednym z najbardziej nasilonych procesów współczesności. Wraz z rozwojem Internetu, nieuniknione było, że media staną się coraz bardziej wykorzystywane i przystosowane do instytucji społecznych oraz politycznych. Tradycyjne media, takie jak telewizja, radio czy prasa drukowana nadal istnieją, ale w wielu przypadkach w innych nowych formach, takich jak telewizja na smartfonach, wersja online gazety i strumieniowe radio¹.

Mediatyzacja jest koncepcją odnoszącą się do zmian wprowadzanych przez media do każdego aspektu naszego życia, w tym do polityki. Jesper Strömbäck wyróżnia 4 fazy mediatyzacji polityki²:

1. Pierwsza faza mediatyzacji osiągana jest, gdy środki masowego przekazu stanowią najważniejsze źródło informacji i są kanałem komunikacji między obywatelami, a instytucjami politycznymi i aktorami, takimi jak partie polityczne, czy agencje rządowe.
2. W drugiej fazie mediatyzacji media stają się bardziej niezależne od rządów lub innych organów politycznych, a w konsekwencji mogą być regulowane zgodnie z logiką mediów, a nie według dowolnej logiki politycznej.

¹ K. S. Nie, C. P. Keea, A. L. Ahmad, *Mediatization: A Grand Concept or Contemporary Approach?*, „Procedia - Social and Behavioral Sciences”, 2014, nr 155, s. 362-367.

² J. Strömbäck, *Four Phases of Mediatization: An Analysis of the Mediatization of Politics*, „The International Journal of Press/Politics”, 2008, nr 13(3).

3. W trzeciej fazie aktorzy społeczni muszą dostosować się do mediów, a nie na odwrót do logiki politycznej. Znaczenie mediów w trzeciej fazie wykracza poza pojedyncze interakcje z aktorami politycznymi, próbującymi wpłynąć na wiadomości.
4. Czwarta faza mediacji jest zatem osiągnięta, gdy aktorzy społeczni i polityczni nie tylko dostosowują się do logiki medialnej i dominujących wartości wiadomości z mediów, ale także internalizują je mniej lub bardziej świadomie, pozwalając aby logika mediów stała się integralną częścią procesów rządzenia. Lata po roku 2010, gdzie nastąpił rozkwit mediów społecznościowych przynoszą dynamiczne zmiany i rysujące się problemy związane z procesem mediatyzacji:
 - Mediatyzacja postępuje – przyspieszyła ją pandemia Covid-19, która spowodowała przeniesienie większości aktywności do Sieci, także tych związanych ze sferą komunikacji publicznej. Wiele osób i instytucji publicznych używało kanałów internetowych do informowania obywateli o rekomendacjach ekspertów, przekazywania praktycznych porad o zachowaniu w razie kontaktu z osobami chorymi, telefonach do obiektów opieki zdrowotnej lub porad w zakresie zachowania zasad higieny i dezynfekcji. Praktyczne porady z zaufanych i aktualnych źródeł pomagały informować opinię publiczną o zaleceniach lokalnych władz administracyjnych.³
 - Proces mediatyzacji polityki również w polskich realiach przybiera na sile. *Twitterowe* konto Ministerstwa Zdrowia (@MZ_GOV_PL) w kwietniu 2022 obserwowało ponad 600 tys. osób, przekazując codziennie informacje o liczbie zakażonych na Covid-19 i innych aktualnych tematach. Constance Ducombe wskazała, jak ważne jest komunikowanie za pośrednictwem tej platformy również w polityce⁴. W Polsce występował największy odsetek *tweetów* politycznych wraz z wzmiankami politycznymi, na drugim miejscu znalazła się Hiszpania, a na trzecim Kanada⁵. Polityczne sprawy bieżące również są głośno komentowane w mediach jak np. protest organizowany przeciwko orzeczeniu Trybunału

³ T. Goban-Klas, *The Media in the "Liquid" Coronavirus Pandemic of 2020. Communication During Lockdown: Separate but Together*, „Studia Medioznawcze”, 2020, nr 4(83).

⁴ C. Ducombe, *The Politics of Twitter: Emotions and the Power of Social Media*, Oxford Academic, 2019.

⁵ C. Ducombe, *The Politics...*, dz. cyt.

Konstytucyjnego z 22 października 2020, które delegalizuje w Polsce wykonywanie aborcji z powodu wad płodu⁶. Wykorzystanie nowych technologii informacyjnych i komunikacyjnych – ICT (*ang. Information and Communication Technologies*) w kampaniach wyborczych jest od co najmniej dwóch dekad codziennością⁷. Nie tylko rząd, ale także polscy politycy korzystają z internetowych mediów. Według raportu agencji *Grayling* polscy parlamentarzyści na tle Europy i UE wykazują się większą aktywnością w mediach pod względem średniej liczby opublikowanych postów na osobę. Politycy częściej publikują na Twitterze, trochę inaczej niż ich wyborcy, którzy raczej preferują Facebooka⁸. Na podstawie badania można stwierdzić, że liczba publikowanych treści na Twitterze jest ponad dwukrotnie wyższa niż na platformie Facebook – Facebook stanowi: 31% treści, Twitter: 66% treści. Udział Instagrama w komunikacji polityków pozostaje na dość niskim poziomie – 2,7 % wszystkich treści wygenerowanych przez posłów w social media⁹. Wraz z wyraźnym wzrostem popularności mediów społecznościowych w Polsce proporcjonalnie widoczne jest większe zainteresowanie polityków¹⁰. Media społecznościowe początkowo doceniane głównie przez liderów oraz partie polityczne cieszące się niewielkim poparciem (nieprzekraczającymi progu wyborczego) wraz rosnącym zainteresowaniem wśród potencjalnych wyborców (w tym dziennikarzy) stały się niezbędnym elementem komunikacji polityków w Polsce¹¹. Partia Prawo i Sprawiedliwość posiadała na dzień 15.06.2022 ponad 277 tys. obserwujących na platformie Twitter, koło Konfederacja ponad 151 tys., Lewica 111 tys., PSL 44 tys., Koalicja Obywatelska ponad 27 tys., a mówiąc o liderach partii to np. konto Donalda Tuska na Twitterze obserwowało 1.4 mln osób.

⁶ J. Trzcińska, *Women and Women's Rights Narratives on Polish Twitter in Light of the Abortion Issue*, *Global Media Journal*, 2021, nr 11(2).

⁷ D. Czyżowski, L. Porębski, *Media społecznościowe w kampanii wyborczej. Wykorzystanie Twittera w polskich wyborach parlamentarnych w roku 2015*, „STUDIA I ANALIZY” *Studia politologiczne*, 2017, nr 45, s. 165.

⁸ Grayling's report on how politicians across Europe are using social media, 2021, https://bpcc.org.pl/uploads/ckeditor/attachments/19242/GPol_-_European_survey_of_SM_use_amongst_MPs_Poland_highlighted_.pdf, 20.04.2022.

⁹ *Grayling's report...*, dz. cyt., 20.04.2022.

¹⁰ M. Adamik-Szysia, *Social media as a tool of political permanent campaign on the example of the activity of polish politicians*, „SWS Journal of Social Sciences and Art”, 2019, nr 2, s. 2.

¹¹ M. Adamik-Szysia, *Social media...*, dz. cyt., s. 2.

- Postępująca cyfryzacja powoduje gwałtowny przyrost danych, których przetwarzanie w celu zamiany na dane użyteczne staje się wyzwaniem, ale i szansą. Nabiera na znaczeniu analiza wielkich zbiorów danych (*ang. Big Data*). Analizy *Big Data* mają na celu najczęściej w zależności od rodzaju analityki odpowiedzieć na zapotrzebowanie, opisać, zoptymalizować bądź przewidzieć zjawiska. W tym kontekście pojawia się termin *rafinacja* dużych zasobów sieciowych¹², który odnosi się do procesu pozyskiwania ze zbiorów danych cennych informacji – niekiedy wymaga to wielu przekształceń i połączeń zbiorów danych. Rafinacja danych ma zastosowanie w każdym obszarze aktywności człowieka, zarówno biznesie, jak i sferze publicznej, także w obszarze polityki, np. może być przydatna w prognozie poparcia partii politycznych¹³.
- Na bazie coraz większej ilości dostępnych danych, coraz silniej można obserwować trend kwantyfikacji i danetyzacji. Termin kwantyfikacji odnosi się do ilościowego ujmowania problemów, doświadczeń oraz obserwacji, aby stały się mierzalne. Miary są to fakty pokazane w sposób ilościowy, ale jest to również podstawa nauki¹⁴. O kwantyfikacji świata mówią Viktor Meyer-Schonberger i Kenneth Cukier w publikacji *Big Data. Rewolucja, która zmieni nasze myślenie*¹⁵. Pojawienie się komputerów umożliwiło wdrożenie narzędzi do pomiaru oraz przechowywania danych, co ulepszyło proces danetyzacji (*ang. danetization*) czyli kwantyfikowania świata i wszystkiego co otacza człowieka¹⁶. Na przestrzeni lat termin stał się bardzo popularny i często stosowany w różnych dziedzinach życia, co jest powiązane ze wzrostem zainteresowania stosowaniem analizy z udziałem *Big Data*.
- Złożoność procesów współczesności i wzajemne ich oddziaływanie sprawia, że pojawia się potrzeba szukania nowych metod i narzędzi badawczych. Znaczenia

¹² W. Gogołek, *Rafinacja dużej skali zasobów sieciowych — Big Data. Dziennikarskie źródło informacji*, 2016; B. Czechowska-Derkacz, D. Chomik, J. Wojsław (red.), *Potencjał informacyjny rafinacji Big Data. Biznes i Media*, [w:] *Media Biznes Kultura. Media w procesie zmian – etyka i komunikacja*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk 2018.

¹³ W. Gogołek, *Rafinacja dużej...*, dz. cyt.

¹⁴ C. J. McKeen, F. Livingston, *Physical and mental measurements of the students of Columbia University*, „The Psychological Review”, 1986, nr 6, s. 618-648.

¹⁵ V. Meyer-Schonberger, K. Cukier, *Big Data. Rewolucja, która zmieni nasze myślenie*, MT Biznes, 2014.

¹⁶ V. Meyer-Schonberger, K. Cukier, *Big Data. Rewolucja...*, dz. cyt., s. 103-132.

nabierają badania interdyscyplinarne, gdzie połączenie różnych dyscyplin naukowych niesie ze sobą większą wartość i wyciąganie cennych wniosków¹⁷.

- W kwestiach językowych pojawił się termin *politolingwistyka* rozumiany jako *ogół badań o charakterze językoznawczym, dotyczących sfery życia politycznego*¹⁸. Aktorzy polityczni coraz śmielej posługują się językiem polityki nie jako środkiem służącym do komunikacji, lecz narzędziem rywalizacji¹⁹. Janina Fras wyróżnia trzy podgrupy języka polityki: język propagandy (cele są dalekosiężne), język agitacji (ma cele doraźne, teksty są napastliwe i emocjonalne) oraz język dyplomacji (język ogólnikowy, nieokreślony)²⁰. Niektórzy badacze ujmują język polityki nie jako osobną grupę, lecz jako mieszankę różnych stylów. Kazimierz Ożóg twierdzi, że *język polityki jest wygodnym terminem i znaczy po prostu język tekstów politycznych. Teksty te łączą w sobie różne odmiany polszczyzny, rejestry leksykalne, style, odmienne sposoby mówienia*²¹. Podobnie jak Jerzy Bralczyk, który uważa że: *Można zresztą wątpić, czy da się obronić publicystycznie atrakcyjną tezę o istnieniu obecnie odmiany, którą dałoby się określić jako język polityki*²².
- Z ogólnego pojęcia dyskursu opisywanego jako *zbiór zdarzeń komunikacyjnych, mających na celu negocjowanie znaczeń przez uczestniczących w nich aktorów* wyłania się oprócz odmian dyskursu potocznego, naukowego, biznesowego, dyskurs polityczny²³. Marek Czyżewski proponuje rozróżnienie trzech głównych odmian dyskursu: dyskurs publiczny, dyskurs polityki i dyskurs polityczny²⁴. Dyskurs publiczny pełni rolę nadrzędną i jest opisywany jako ten zawierający

¹⁷ T. Goban-Klas, *Nauki o mediach – baza czy nadbudowa nowych teorii nauk społecznych?*, Uniwersytet Jagielloński, Kraków 2012, s. 3.

¹⁸ S. Sojda, *Słowacka i polska politolingwistyka*, [w:] *Z warsztatu współczesnego słowacysty. Studia słowacystyczne dedykowane pamięci profesor Marii Honowskiej*, (red.) H. Mieczkowska, A. Hudymač, Z. Babik, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2010, s. 95.

¹⁹ A. Balczyńska-Kosman, *Język dyskursu publicznego w polskim systemie politycznym*, „Środkowo-europejskie Studia Polityczne”, 2013, nr 2, s. 143.

²⁰ J. Fras, *Język propagandy politycznej*, [w:] B. Dobek-Ostrowska, J. Fras, B. Ociepka, *Teoria i praktyka propagandy*, Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław 1997, s. 88.

²¹ K. Ożóg, *Uwagi o języku polskiej polityki po roku 1989*, [w:] *Przekazy polityki*, (red.) A. Siewierska-Chmaj, Konsorcjum Akademickie, Kraków–Rzeszów–Zamość 2009, s. 84-85.

²² J. Bralczyk, *O języku propagandy i polityki*, Wydawnictwo Trio, Warszawa 2007, s. 293.

²³ A. Balczyńska-Kosman, *Język dyskursu...dz. cyt.*, s. 144.

²⁴ M. Czyżewski, *Rytualny chaos. Studium dyskursu publicznego*, (red.) M. Czyżewski, S. Kowalski, A. Piotrowski, Wrocław 2004, s. 21.

wszelkie przekazy, które są dostępne w sposób publiczny: dyskursy instytucjonalne (usługi publiczne, jak szkoły czy urzędy), dyskursy społeczne (dyskurs biznesowy, dyskurs literacki) oraz dyskurs medialny. Dyskurs polityki stanowi część dyskursu publicznego i definiowany jest jako wypowiedzi ludzi reprezentujących elity władzy. Czasami dyskurs publiczny poprzez tabloidyzację mediów koncentruje się na życiu prywatnym polityków, co znacznie splota obraz życia publicznego²⁵. Dyskurs polityczny ujęty jest szerzej jako *proces komunikacji elit symbolicznych* jako osoby i grupy sprawujące władzę w środkach komunikacji masowej włączając dziennikarzy oraz w głównej mierze polityków – niektórzy z nich mogą reprezentować podwójną rolę również elit władzy.²⁶ Marek Czyżewski, Sergiusz Kowalski, Andrzej Piotrowski twierdzą, że w przypadku dyskursu na scenie politycznej w jednej z czterech faz opisywanych szerzej w swojej pracy, po 1989 r. możemy zaobserwować zjawisko rytualnego chaosu (w fazie II). Ich zdaniem, rytualny chaos polega na tym, że *świadek sporów na scenie politycznej lub w środkach masowego przekazu odnosi wrażenie, że porozumienie między przeciwstawnymi stronami jest z gruntu niemożliwe, a stronom tym nie chodzi o nic więcej, tylko nagłaśnianie własnych racji*²⁷. W takich warunkach do odbiorcy komunikatów dociera głównie chaos oraz cały dyskurs ulega rytualizacji i teatralizacji, a konflikt polityczny zaczyna stawać się dwubiegunowy, co często powoduje społeczną bierność²⁸.

- Katarzyna Kłosińska w swojej pracy analizując programy, przemówienia liderów, manifesty oraz ulotki partii politycznych po 1989 r. wyróżnia dwa typy dyskursu, które dominują w Polsce. Są to: (I) Dyskurs etyczny, gloryfikujący moralność i uczciwość, uwidacznia negatywne kwestie rzeczywistości poprzez użycie słów o nacechowaniu negatywnym (np. hańba, zdrada), negatywnych metafor (wirus zakażenia, lewicowa gangrena), zepsucia i brudu (zgnilizna moralna, brudne

²⁵ A. Balczyńska-Kosman, *Język dyskursu* ..., dz. cyt., s. 153.

²⁶ M. Czyżewski, *Rytualny chaos*..., dz. cyt., s. 21.

²⁷ M. Czyżewski, S. Kowalski, A. Piotrowski, *Rytualny chaos. Studium dyskursu publicznego*, Areus, Kraków 1997, s. 53.

²⁸ K. Podemski, *Badania polskiego dyskursu publicznego II RP, PRL i III RP Przegląd zagadnień*, „Kultura i Społeczeństwo: Polska Akademia Nauk Komitet Socjologii Instytut Studiów Politycznych”, 2013, nr 2, s. 40.

interesy).²⁹ Nadawcy aspirują do bycia myślącymi o przyszłych pokoleniach przewodnikami czy misjonarzami, a politycy *ojcami narodu*³⁰. Dyskurs ten, wg autorki, dzieli się na dwie grupy: *romantyczną* (odwołujący się do narodu i *sacrum*, związany z patriotyzmem, tradycją, ojczyzną, rodziną, katolicyzmem); *socjalistyczną* (odwołujący się do pracy, człowieka, krzywdy, podkreślający wolność od biedy i bezrobocia)³¹. (II) Drugi typ dyskursu w ocenie autorki to dyskurs pragmatyczny, który podkreśla pozytywne strony rzeczywistości, w małym stopniu posługuje się metaforami, a polityk należący do wyznawców tego dyskursu staje się fachowcem, który realizuje konkretne działania³². K. Kłosińska uznaje partię Platforma Obywatelska (PO) za należącą do dyskursu pragmatycznego, a partię Prawo i Sprawiedliwość (PiS) za należącą do dyskursu romantycznego³³. Należy zwrócić uwagę, że dyskurs polityczny ciągle się zmienia.

- Kontynuując, w teorii dyskurs powinien mieć charakter demokratyczny, pluralistyczny, powszechny oraz permanentny. J. Fras wykląda, że model demokracji dyskursywnej, charakteryzuje w debatach publicznych duży udział mediów masowych, co powinno kształtować obywateli do uczestniczenia we wspólnocie komunikacyjnej³⁴. W społeczeństwie demokratycznym, gdzie dyskursy polityczne są zgodne z zasadami retoryki i logiki mediów tworzy się zjawisko mediatyzacji języka polityki³⁵. Alina Balczyńska-Kosman twierdzi, że dzisiejszy język dyskursu politycznego charakteryzuje duża potoczność, negatywna ekspresja, lekceważenie i brutalizacja przeciwnika, zamiast konstruktywnych argumentów często stosuje się silne emocje, a działanie mediów narzuca sposób prezentowania tematów politycznych z użyciem m.in. *primingu* (eksponowanie i częste prezentowanie z użyciem autorytetów popularnych treści) i *framingu* (nadawanie ram, wydłużanie jednych informacji i pomijanie innych – określony sposób narracji)³⁶.

²⁹ K. Kłosińska, *Etyczny i pragmatyczny. Polskie dyskursy polityczne po 1989 roku*, Narodowe Centrum Kultury, Warszawa 2012.

³⁰ K. Podemski, *Badania polskiego...*, dz. cyt., s. 47.

³¹ K. Kłosińska, *Etyczny i...*, dz. cyt.

³² K. Podemski, *Badania polskiego...*, dz. cyt., s. 47.

³³ K. Kłosińska, *Etyczny i...*, dz. cyt.

³⁴ J. Fras, *Komunikacja polityczna. Wybrane zagadnienia gatunków i języka wypowiedzi*, Wrocław 2005, s. 101-102.

³⁵ A. Balczyńska-Kosman, *Język dyskursu...*, dz. cyt., s. 145.

³⁶ A. Balczyńska-Kosman, *Język dyskursu...*, dz. cyt., s. 152-153.

Warto zwrócić uwagę na język polityki po stronie polityków oraz uczestników komunikowania politycznego w Internecie.

Opisane powyżej powody skłaniają do podjęcia badań w obszarze nauki o mediach wzbogaconych o metody sztucznej inteligencji (*ang. Artificial Intelligence*) szczególnie ukierunkowane na analizę treści oraz przetwarzanie języka naturalnego (*ang. Natural Language Processing (NLP)*) wraz budowaniem i procesem testowania modeli predykcyjnych.

Z tych właśnie przyczyn, ideą podjęcia pracy było zastosowanie kompleksowego podejścia opartego na metodach badawczych różnych dyscyplin naukowych w celu udowodnienia, że dane tekstowe mają wartość nie tylko w analizach jakościowych, ale przede wszystkim w ujęciu ilościowym. Poglądy, które są akcentowane w wypowiedziach niosą ze sobą wartość możliwą do skwantyfikowania i na podstawie wykorzystania elementów z różnych dyscyplin pozwalają wyciągnąć wnioski potrzebne do stworzenia modeli predykcyjnych, wskazujących na przynależność partyjną na podstawie użytych treści. Analiza treści wykonana została z wykorzystaniem metod i narzędzi *Big Data* na danych pochodzących z mediów społecznościowych oraz publicznych danych ze stenogramów sejmowych, co stanowi interdyscyplinarne podejście – aplikację do dyscypliny nauki o komunikacji społecznej i mediach algorytmów przetwarzania języka naturalnego, elementów sztucznej inteligencji oraz pomocnych elementów wizualizacji danych. W myśl Tomasza Goban-Klasa *Praktyka badawcza zmierza ku rozszerzeniu pola zainteresowań i zakresu współpracujących dyscyplin*³⁷. Odpowiednio stosowane metody i narzędzia będą dostarczały cennej wiedzy w zakresie analizy treści mediów politycznych, kwantyfikacji tej treści oraz całego procesu mediatyzacji polityki, stanowiącego nieodłączną część postępu technologicznego mającego wpływ na społeczeństwo i wybory polityczne.

³⁷ T. Goban-Klas, *Nauki o mediach i komunikacji społecznej jako nowa dyscyplina nauk humanistycznych*, „Studia Medioznawcze”, 2008, 2(33), s. 14.

Stan aktualnych badań

Biorąc pod uwagę tematykę pracy należy zwrócić uwagę na kilka obszarów badań.

Rdzeniem pracy jest proces mediatyzacji i wykorzystywanie nowych mediów jako środka przekazu informacji. O teorii mediatyzacji piszą w swoich publikacjach m. in. Stig Hjarvard³⁸, Winfried Schulz³⁹ i Knut Lundby⁴⁰, Friedrich Krotz⁴¹, Andreas Hepp⁴² oraz wielu innych. Według S. Hjarvarda pojęcie to było używane w różnych kontekstach, ale nie włożono odpowiednio dużo pracy, aby zdefiniować termin i rozwinąć teoretyczny koncept⁴³. Temat nowych mediów podejmuje w swoich pracach m. in. Paul Levinson⁴⁴, Magdalena Szpunar⁴⁵. Sonia Livingstone, Leah Lievrouw mówią o złożoności tego pojęcia i łączą kwestie informatycznych technologii komunikowania wraz z społecznymi kwestiami i trzema elementami, jak artefakty i urządzenia techniczne, działania i praktyki, układy i organizacje społeczne⁴⁶. Według Denisa McQuaila nowe media cechują się wzajemnym powiązaniem, interaktywnością, wszechobecnością, delokalizacją, otwartością oraz dostęp do indywidualnych użytkowników⁴⁷. O wpływie mediatyzacji na politykę jako jeden z pierwszych wspominał Kent Asp który mówił o mediatyzacji życia politycznego jako procesie⁴⁸. Gianpietro Mazzoleni i Winfried Schulz odnieśli koncepcję mediatyzacji do wpływu mediów na politykę⁴⁹. O mediatyzacji polityki i relacjach

³⁸ S. Hjarvard, *The mediatization of culture and society*, Routledge, London and New York 2013;
S. Hjarvard, *The Mediatization of Society. A Theory of the Media as Agents of Social and Cultural Change*, Nordicom Review 29, 2008, nr 2.

³⁹ W. Schulz, *Reconstructing Mediatization as an Analytical Concept*, „European Journal of Communication”, 2004, nr 1.

⁴⁰ K. Lundby, *Mediatization. Concept, changes, consequences*, Peter Lang, New York 2009.

⁴¹ F. Krotz, *Mediatisierung: Fallstudien zum Wandel von Kommunikation*, Wiesbaden: VS Verlag für Sozialwissenschaften, 2007.

⁴² A. Hepp, *Transforming Communications. Media-related Changes in Times of Deep Mediatization*, „COMMUNICATIVE FIGURATIONS”, 2017, nr 16.

⁴³ S. Hjarvard, *The mediatization of culture and society*, Routledge, London and New York 2013, s. 8.

⁴⁴ P. Levinson, *Nowe nowe media*, WAM, Kraków 2010; P. Levinson, *Telefon komórkowy. Jak zmienił świat najbardziej mobilny ze środków komunikacji*, Muza, Warszawa 2006.

⁴⁵ M. Szpunar, *Nowe stare medium. Internet między tworzeniem nowych modeli komunikacyjnych a reprodukowaniem schematów komunikowania masowego*, Warszawa 2012.

⁴⁶ S. Livingstone, L. Lievrouw, *The Handbook of New Media*, London 2002.

⁴⁷ D. McQuail, *Teoria komunikowania masowego*, Wydawnictwo Naukowe PWN, Warszawa 2008.

⁴⁸ K. Asp, *Måktiga massmedier: Studier i politisk opinionsbildning [Powerful mass media: studies in political opinion-formation]*, Akademilitteratur, Stockholm 1986.

⁴⁹ G. Mazzoleni, W. Schulz, *Mediatization of Politics: A Challenge for Democracy?*, *Political Communication*, 1999, nr 16, s. 247-261.

między systemem politycznym, a medialnym pisał Zbigniew Oniszcuk⁵⁰. O typologii partii politycznych pisze Krzysztof Kowalczyk⁵¹. Biorąc pod uwagę budowę strategii komunikacyjnych przez partie na przykładzie jednej z nich pisze Mariusz Kolczyński wskazując przy tym jak jest to ważny aspekt w celu sukcesywnego budowania poparcia politycznego⁵². W przypadku systemów komunikowania można wyróżnić m.in. społeczny, publiczny, masowy, organizacyjny, które opisuje T. Goban-Klas⁵³. Bogusława Dobek-Ostrowska opisuje teorię i pojęcie procesu komunikowania politycznego⁵⁴. Rafał Leśniczak przedstawia jak wyglądało w kontekście mediatyzacji komunikowanie polityczne Konferencji Episkopatu Polski podczas kampanii prezydenckich i parlamentarnych⁵⁵.

Wpływem mediów na społeczeństwo zajmuje się m.in. T. Goban-Klas⁵⁶ czy też Andrzej Adamski⁵⁷, pokazując przemianę oraz konsekwencje ewolucji mediów. O specyfice zarządzania sytuacji kryzysowych w mediach społecznościowych pisze Monika Kaczmarek-Śliwińska⁵⁸. Tematem rozwoju społeczeństwa Sieci zajmują się m. in. Derrick de Kerckhove⁵⁹ i Jan van Dijk⁶⁰. Gudmund Hernes ukuł twierdzenie „społeczeństwo pokręcone w mediach” (ang. *media-twisted society*)⁶¹. Wiele prac odnosi się do kwestii

⁵⁰ Z. Oniszcuk, *Mediatyzacja polityki i polityzacja mediów. Dwa wymiary wzajemnych relacji*, „Studia Medioznawcze”, 2011, nr 4 (47), s. 11-22.

⁵¹ K. Kowalczyk, *Typologia polskich partii politycznych według kryterium programowego*, 10.14746/ssp, nr 1, 2014, s. 73-99.

⁵² M. Kolczyński, *Strategia komunikacyjna Konfederacji w wyborach w 2019 roku – próba rekonstrukcji*, „Studia Politologiczne”, 2020, nr 55, s. 259-276.

⁵³ T. Goban-Klas, *System informacji w państwie, a system komunikacji społecznej*, Przekazy i Opinie, 1989, nr 2, s. 16.

⁵⁴ B. Dobek-Ostrowska, *Komunikowanie polityczne i publiczne*, PWN, Warszawa 2006.

⁵⁵ R. Leśniczak, *Komunikowanie polityczne Konferencji Episkopatu Polski podczas kampanii parlamentarnej 2019 r. w kontekście procesów mediatyzacji*, [w:] (red.) G. Habrajska, *Teorie i praktyki komunikacji*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2020, s. 499-524; R. Leśniczak, *Komunikowanie polityczne Konferencji Episkopatu Polski po zakończonych wyborach prezydenckich 2015 r. Priming i framing – analiza tygodników opinii „Newsweek Polska” i „Polityka”*, Annales UMCS, 2017, nr 24(2).

⁵⁶ T. Goban-Klas, *Cywilizacja medialna. Geneza, ewolucja, eksplozja*, Wydawnictwa Szkolne i Pedagogiczne, Warszawa 2005; T. Goban-Klas, *Zarys historii i rozwoju mediów. Od malowideł naskalnych do multimediów*, Wydawnictwo Naukowe AP, Kraków 2001; T. Goban-Klas, *Media i komunikowanie masowe. Teorie i analizy prasy, radia, telewizji i Internetu*, Wydawnictwo Naukowe PWN, Warszawa 2004.

⁵⁷ A. Adamski, *Media w analogowym i cyfrowym świecie. Wpływ cyfrowej rewolucji na rekonfigurację komunikacji społecznej*, Dom Wydawniczy ELIPSA, Warszawa 2012.

⁵⁸ M. Kaczmarek-Śliwińska, *Specyfika zarządzania sytuacją kryzysową w przestrzeni mediów społecznościowych w perspektywie typologii SCCT W. T. Coombsa*, „Studia Medioznawcze”, nr 20(4), s. 318-332; M. Kaczmarek-Śliwińska, *Hejt jako źródło sytuacji kryzysowych podmiotów w przestrzeni społecznej*, 2020, nr 8(2), s. 75-87.

⁵⁹ D. de Kerckhove, *Inteligencja otwarta. Narodziny społeczeństwa sieciowego*, Mikom, Warszawa 2001.

⁶⁰ J. van Dijk, *Społeczne aspekty nowych mediów*, Wydawnictwo Naukowe PWN, Warszawa 2010.

⁶¹ G. Hernes, *Det mediavridde samfunn [The media-twisted society]*, in G. Hernes (red.) *For-handlingsekonomi og blandningsadministrasjon*. Bergen: Universitetsforlaget, 1978.

towarzyszących mediatyzacji – Małgorzata Molęda-Zdziech zajmuje się zjawiskiem celebrytyzacji⁶², Małgorzata Lisowska-Magdziarz – realizacjami medialnymi⁶³.

Jeśli chodzi o język to znamienne są prace m.in. Renata Grzegorzczukowej⁶⁴, Mariana Bugajskiego⁶⁵. Język w Internecie to m.in. prace Jana Grzenia⁶⁶ oraz Moniki Worsowicz⁶⁷. Biorąc pod uwagę język polityki nie sposób nie przytoczyć prac Kazimierza Ożoga⁶⁸, Bożeny Sieradzkiej-Baziur⁶⁹ oraz analiz różnych aspektów dyskursu publicznego i politycznego autorstwa M. Czyżewskiego⁷⁰. Językowy obraz świata opisuje m. in. Jerzy Bartmiński, Ryszard Tokarski⁷¹, Marta Kopińska⁷², Jolanta Maćkiewicz⁷³, a badaniem tego zjawiska z kwantyfikacją zajmował się m. in. Michał Ulidis⁷⁴. Wartościowanie w języku opisują prace Jadwigi Puzyniny⁷⁵ oraz Ewy Kapeli⁷⁶. Polskim dyskursem publicznym i politycznym zajmują się m.in. M. Czyżewski, S. Kowalski, A. Piotrowski⁷⁷, K. Kłosińska⁷⁸, K. Podemski⁷⁹. Prace opisujące użycie języka w polityce, zdaniem autora wymagają aktualizacji po często zmieniającej się scenie politycznej po 1989 do nowych warunków panujących w dzisiejszej polityce, gdzie od wyborów parlamentarnych w 2005 roku prym wiodą dwie główne partie polityczne. Dodatkowym aspektem jest coraz

⁶² M. Molęda-Zdziech, *Czas celebrytów. Mediatyzacja życia publicznego*, Difin, Warszawa 2013.

⁶³ M. Lisowska-Magdziarz, *Feniksy, łabędzie, motyle. Media i kultura transformacji*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2012.

⁶⁴ R. Grzegorzczukowa, *Wstęp do językoznawstwa*, PWN, Warszawa 2007.

⁶⁵ M. Bugajski, *Język w komunikowaniu*, Warszawa 2006.

⁶⁶ J. Grzenia, *Komunikacja językowa w Internecie*, Warszawa 2006, s. 64

⁶⁷ M. Worsowicz, *Jak pisać do Internetu? Retoryczne aspekty poradnictwa warsztatowego*, „Retoryka w komunikacji specjalistycznej” (red.) M. Załęska, Polskie Towarzystwo Retoryczne, 2015.

⁶⁸ K. Ożóg, *O języku współczesnej polityki*, „Polityka i Społeczeństwo”, 2007, nr 4, s. 103-111; K. Ożóg, *Język współczesnej polskiej polityki*, „Stylistyka”, 2008, nr 17, s. 161-184.

⁶⁹ B. Sieradzka-Baziur, *Język polityki - dotychczasowe prace i perspektywy badawcze*, „Horyzonty Polityki”, 2011, nr 2(2).

⁷⁰ M. Czyżewski, *Rytualny chaos...*, dz. cyt.

⁷¹ J. Bartmiński, R. Tokarski, *Językowy obraz świata a spójność tekstu*, [w:] *Teoria tekstu. Zbiór studiów*, (red.) T. Dobrzyńska, Wrocław, 1986.

⁷² M. Kopińska, *Język jako narzędzie interpretacji rzeczywistości – językowy obraz świata*, 2011.

⁷³ J. Maćkiewicz, *Kategoryzacja a językowy obraz świata* [w:] *Językowy obraz świata* (red.) J. Bartmiński, Lublin 1999.

⁷⁴ M. Ulidis, *Blog a językowy obraz świata partii politycznej. Analiza kwantytatywno-korpusowa na przykładzie komunikacji Platformy Obywatelskiej*, „Nowe media studia i rozprawy”, 2013, nr. 4, <http://dx.doi.org/10.12775/NM.2013.005>, 2013.

⁷⁵ J. Puzynina, *Problemy wartościowania w języku i w tekście*, „Etnolingwistyka”, 2004, nr 16.

⁷⁶ E. Kapela, *Językowe środki wartościowania w wywiadach z politykami* [w:] A. Charciarek, H. Fontański, J. Lubocha-Kruglik (red.), *Jednostki języka w systemie i mowie*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2015, s. 157-167.

⁷⁷ M. Czyżewski, S. Kowalski, A. Piotrowski, *Rytualny chaos...*, dz. cyt.

⁷⁸ K. Kłosińska, *Etyczny i pragmatyczny...*, dz. cyt.

⁷⁹ K. Podemski, *Badania polskiego...*, dz. cyt.

częstsze użycie mediów społecznościowych oraz zmiana barw politycznych, co może skutkować przeniesieniem językowych naleciałości przez polityków do innych partii.

O wykorzystaniu analizy danych z zasobów sieciowych dla danych politycznych pisze W. Gogołek⁸⁰, Paweł Kuczma⁸¹, Dariusz Jaruga⁸² oraz biorąc pod uwagę anglojęzyczne publikacje dotyczące social mediów w polityce m. in.: Andreas Jungherr⁸³ oraz Elena Bellio, Giulia Crestini, Alessandra Arkoudas⁸⁴. Juliana Schwarz omawia wyniki trzech różnych podejść do przewidywania ideologii politycznej w danych komentarzy na YouTube⁸⁵. Mohit Iyyer, Peter Enns, Jordan Boyd-Graber, Philip Resnik zastosowali algorytmy sztucznej inteligencji do określenia stanowiska politycznego wyrażonego w zdaniach (klasyfikując przynależność do poglądów republikan i demokratów)⁸⁶. Podobnie Elanor Colleoni, Alessandro Rozza, Adam Arvidsson na podstawie danych z mediów społecznościowych klasyfikują poglądy polityczne⁸⁷. Rob Schumaker, Hsiu-chin Chen przedstawiają rolę artykułów z wiadomościami finansowymi, używając algorytmów sztucznej inteligencji tworzą modele przewidujące ceny akcji firm⁸⁸. Fouad Zablith, Ibrahim H. Osman z wykorzystaniem uczenia maszynowego stworzyli modele oceny

⁸⁰ W. Gogołek, *Rafinacja informacji sieciowej*, [w:], A. Jastriebow, M. Raczyńska, *Informatyka w dobie XXI wieku, Nauka, Technika, Edukacja a nowoczesne technologie informatyczne*, Radom 2011, s. 229-238.; W. Gogołek, *Potencjał informacyjny nowych mediów – rafinacja informacji sieciowej*, [w:] 350 lat polskiej prasy, mat. pokonf. Uniwersytet Warszawski, Warszawa 2011, s. 253-263.; W. Gogołek, i inni, *Z badań nad wykorzystaniem rafinacja informacji sieciowej. Wybory prezydenckie i parlamentarne 2015*, „Studia Medioznawcze”, Instytut Dziennikarstwa UW, Warszawa, 2015, nr 3(62), s. 31-43, http://studia-medioznawcze.pl/article.php?date=2015_3_62&content=gogolek&lang=pl, , 10.08.2021.

⁸¹ P. Kuczma, W. Gogołek, *Web Information Potential on the Basis of the 2010 Presidential Elections*, Wyd. Skorpion, Warszawa 2011, s. 241-263.

⁸² W. Gogołek, D. Jaruga, *Z badań nad ...*, dz. cyt.

⁸³ A. Jungherr, *Twitter in Politics: A Comprehensive Literature Review*, University of Konstanz - Department of Politics and Public Administration, 2014.

⁸⁴ E. Bellio, G. Crestini, A. Arkoudas, *Twitter and politics: Evidence from the US presidential elections 2016*, „Journal of Marketing Communications”, 2018.

⁸⁵ J. Schwarz, *Detecting political ideology in Youtube comments using machine learning*, „Seminar Social Media and Business Analytics”, 2009.

⁸⁶ M. Iyyer, P. Enns, J. Boyd-Graber, P. Resnik, *Political Ideology Detection Using Recursive Neural Networks*, Computer Science, Linguistics, iSchool, and UMIACS University of Maryland, 2014; *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Baltimore, Maryland, USA, June 23-25, 2014, s. 1113-1122.

⁸⁷ E. Colleoni, A. Rozza, A. Arvidsson, *Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data*, „Journal of Communication”, 2014, nr 64(2).

⁸⁸ R. Schumaker, H. Chen, *Textual Analysis of Stock Market Prediction Using Financial News Articles*, „Americas Conference on Information Systems (AMCIS)”, 2006, nr 27(2).

produktów i usług⁸⁹. W zakresie wykorzystania *Big Data* warto zwrócić uwagę na prace Andrew McAfee, Erik Brynjolfsson⁹⁰, Alexandros Labrinidis, H. V. Jagadish⁹¹. *Big Data* również ma zastosowanie w obszarze polityki, piszą o tym m.in.: David W. Nickerson, Todd Rogers,⁹² Alon Peled⁹³.

Cel pracy

Głównym celem pracy jest odpowiedź na pytanie: czy, w jaki sposób i w jakim stopniu analiza treści w komunikowaniu politycznym i narzędzia sztucznej inteligencji mogą służyć do budowy modeli predykcyjnych dopasowania do grupy politycznej⁹⁴. Na podstawie celu głównego określono cele szczegółowe, które odnoszą się bezpośrednio do kilku kwestii:

1/ Pokazanie, w jaki sposób i w jakim stopniu rozwój nowych mediów i technologii zmienia komunikowanie polityczne.

2/ Diagnoza sposobu komunikowania polityków z poszczególnych partii, w szczególności charakterystycznych dla nich słów, zwrotów i emocji.

3/ Zbadanie treści wypowiedzi użytkowników mediów społecznościowych dotyczących partii politycznych oraz porównanie z wypowiedziami polityków.

4/ Zbudowanie i weryfikacja modeli, które będą w stanie przewidzieć na podstawie słownictwa uczestników komunikowania politycznego ich dopasowania do grupy politycznej oraz wybór optymalnego modelu.

⁸⁹ F. Zablith, I. H. Osman, *ReviewModus: Text classification and sentiment prediction of unstructured reviews using a hybrid combination of machine learning and evaluation models*, „Applied Mathematical Modelling”, 2019, nr 71, s. 569-583.

⁹⁰ A. McAfee, E. Brynjolfsson, *Big Data: The Management Revolution*, „Harvard Business Review”, 2012, nr 90(10).

⁹¹ A. Labrinidis, H. V. Jagadish, *Challenges and Opportunities with Big Data*, „Proceedings of the VLDB Endowment”, 2012, nr 5(12).

⁹² D. W. Nickerson, T. Rogers, *Campaigns and Big Data*, „HKS Faculty Research Working Paper Series”, 2013.

⁹³ A. Peled, *The Politics of Big Data A Three-Level Analysis*, European Consortium of Political Research (ECPR) General Conference, 2013.

⁹⁴ Grupy polityczne rozumiane w charakterze partii oraz koalicjantów, które weszły do Sejmu z list poszczególnych komitetów wyborczych (Prawo i Sprawiedliwość (jako PiS), Koalicja Obywatelska (jako KO), Koalicja Polska (jako KP), Konfederacja, Lewica).

Problemy badawcze

Problem badawczy główny został sformułowany w postaci pytania: Czy oraz w jakim stopniu analiza treści wypowiedzi uczestników komunikowania politycznego może być wykorzystana do określania ich dopasowania do grupy politycznej?

Problemy badawcze szczegółowe, na które należy zwrócić uwagę to:

1/ Jakie są współczesne kanały komunikowania politycznego, technologie z obszaru *Big Data* i jakie niesie to nowe możliwości wykorzystania w zakresie analizy treści w komunikowaniu politycznym?

2/ Czy poszczególne ugrupowania polityczne mają pewne cechy wspólne, jeśli chodzi o przekaz i dobór słów oraz zwrotów?

3/ W jaki sposób zbadać wypowiedzi uczestników komunikowania politycznego, w szczególności:

- w jaki sposób dobrać metodę pomiaru nacechowania oraz emocji, w tym interwały czasowe i wskaźniki pomiaru?
- w jaki sposób zmierzyć podobieństwo tekstu (dla danych z social mediów i dla wypowiedzi polityków)?
- jakich miar użyć do weryfikacji podobieństwa tekstu?

4/ Jak zbudować i skonfigurować modele na bazie algorytmów predykcyjnych służące do przewidywania przynależności do grupy politycznej wykorzystując wyniki analizy treści wypowiedzi uczestników komunikowania politycznego i które z algorytmów będą bardziej wydajne i dadzą lepsze wyniki klasyfikacji?

Hipotezy badawcze

W odniesieniu do celu badawczego głównego, sformułowana została hipoteza główna: Wypowiadane słowa polityków oraz użytkowników mediów społecznościowych komentujących wydarzenia polityczne niosą ze sobą wartość informacyjną i po procesie kwantyfikacji oraz rafinacji danych mogą posłużyć do analizy treści i budowy modeli predykcyjnych dopasowania do grupy politycznej.

Po ustaleniu hipotezy głównej ustalono hipotezy pomocnicze:

1/ Wraz z postępującą mediatyzacją polityki zwiększa się wykorzystanie mediów cyfrowych w komunikowaniu politycznym oraz zwiększa się możliwość analizy treści komunikowania politycznego.

2/ Poszczególne grupy polityczne mają indywidualne słowniki słów oraz zwrotów wykorzystywane w komunikacji.

3/ Słowa, zwroty oraz ich nacechowanie i emocje uczestników komunikowania politycznego (polityków i użytkowników mediów społecznościowych) zmieniają się w interwałach czasowych.

4/ Istnieją podobieństwa między grupami polityków oraz użytkowników mediów społecznościowych wypowiadających się na tematy polityczne.

5/ Słowa wypowiediane przez uczestników w wypowiedziach komunikowania politycznego będą wskazywały do jakich partii pasują najbardziej.

6/ Charakterystyka algorytmu, w tym jego wydajność, będzie miała wpływ na stopień dopasowania użytkownika do partii politycznej na podstawie jego wypowiedzi.

Metodologia i techniki badawcze

Dla weryfikacji postawionych hipotez zastosowane zostały różne metody i techniki badawcze.

W pierwszym etapie dokonana została analiza zebranego materiału (*ang. desk research*) dotyczącego, mediów, procesu mediatyzacji, opis najważniejszych modeli komunikacji. Dodatkowo opisano systemy i zakresy systemów: społecznego, politycznego i medialnego. Dokonano analizy wraz z opisem elementów dotyczących zarówno kwestii językowych, jak i technicznych (informatycznych) oraz dokonano syntezy pozwalającej na sformułowanie kluczowych diagnoz w badanych obszarach.

W drugim etapie również wykorzystano analizę zebranego materiału (*ang. desk research*) pod kątem mediatyzacji polskiej polityki oraz dokonano opisu przemian społeczno-gospodarczych w Polsce po 1989 roku do stanu aktualnego (Sejmu IX kadencji). Zwrócono uwagę na równoległy proces rozwoju mediów w Polsce i przemian w komunikacji politycznej w tym okresie. Zbadano również wykorzystanie nowych mediów do komunikacji w społeczeństwie zarówno po stronie obywateli, jak i po stronie polityków.

W przypadku metodologii części trzeciej wykorzystano dane poszczególnych posłów aktualnej kadencji Sejmu z mównicy sejmowej, które zostały pobrane i przygotowane dzięki dedykowanemu oprogramowaniu napisanemu przez autora języku *Python* do automatyzacji procesu zbierania danych. Wykorzystano również dane z serwisu Twitter użytkowników komentujących wydarzenia polityczne. Dane zostały pobrane do analiz z wykorzystaniem narzędzia *bsi-sentiment*⁹⁵ skonfigurowanego w języku *Python*. Zastosowano proces oczyszczania danych, po którym wykonano analizę pokazującą frekwencję, ngramy (elementy sekwencji wyrazów), klasteryzację (grupowanie) poszczególnych słów w grupach danych dotyczących polityków oraz danych z social mediów. Porównano wypowiedzi indeksem podobieństwa⁹⁶. Do klasteryzacji posłużono się dystansem Euklidesa z zastosowaniem algorytmu hierarchicznego klasteringu, aby pogrupować stopniowo słowa o podobnych zakresach wystąpień⁹⁷. Następnie klastry zbadano pod względem użytych konfiguracji⁹⁸. Wykonano analizę sentymentów pokazującą poziom nacechowania wypowiedzi⁹⁹. Dane podzielono na okresy i zbadano trendy. Oceniono każdą wypowiedź w stopniu nacechowania. Zbadano emocje w słowach uwzględniając również okresy. Zbudowano słowniki słów i wyrażeń używanych przez polityków oraz użytkowników social mediów.

W przypadku części czwartej oczyszczone dane odpowiednio przygotowano i wykonano modele predykcyjne metodami klasyfikacyjnymi posługując się algorytmami uczenia maszynowego jak: *kNN*¹⁰⁰, *GBM*¹⁰¹, *SVM*¹⁰², *Decision Trees*¹⁰³, *Random Forest*¹⁰⁴,

⁹⁵ Bsi-sentiment, <https://pypi.org/project/bsi-sentiment/>, GNU General Public License v3 (GPLv3), 05.05.2021.

⁹⁶ S. Niwattanakul, J. Singthongchai, E. Naenudornand, S. Wanapu, *Using of Jaccard Coefficient for Keywords Similarity*, Proceedings of the International MultiConference of Engineers and Computer Scientists, 2013.

⁹⁷ S. C. Johnson, *Hierarchical Clustering Schemes*, Psychometrika, 1967, nr 32(3).

⁹⁸ P. H. A. Sneath, R. R. Sokal, *Numerical Taxonomy: The Principles and Practice of Numerical Classification*, Freeman, San Francisco 1973, s. 278.

⁹⁹ M. D. Devika, C. Sunitha, A. Ganesh, *Sentiment Analysis: A Comparative Study On Different Approaches*, „Procedia Computer Science”, 2016, nr 87.

¹⁰⁰ Z. Zhang, *Introduction to machine learning: k-nearest neighbors*, „Ann Transl Med”, 2016, nr 4(11).

¹⁰¹ H. Lu, S. P. Karimireddy, N. Ponomareva, V. Mirrokni, *Accelerating Gradient Boosting Machine*, Cornell University, 2020.

¹⁰² C. Cortes, V. Vapnik, *Support-vector networks*, „Machine Learning”, 1995, nr 20 (3), s. 273-297.

¹⁰³ J. R. Quinlan, *Simplifying decision trees*, „International Journal of Man-Machine Studies”, 1987, nr 27 (3), s. 221-234.

¹⁰⁴ T. Kam Ho, *Random Decision Forests*, Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August, 1995, s. 278-282.

ANN¹⁰⁵ z wykorzystaniem elementów *Deep Learning*¹⁰⁶. Ze względu na różnicę klas w poszczególnych grupach polityków zastosowano dwa podejścia: *oversampling*¹⁰⁷ oraz *downsampling*¹⁰⁸ – wyniki porównano. Wykorzystano język programowania R oraz Python z narzędziem *TensorFlow*. Wyniki porównano odpowiednimi współczynnikami oraz dokonano rekomendacji rozwiązań optymalnych.

Dobór metod badawczych do weryfikacji poszczególnych hipotez oraz podział na etapy badań przedstawiony został w tabeli 1.

Tabela 1. Hipotezy badawcze a etapy badań

lp	etap badań	badane dane	przydatność badania do weryfikacji hipotezy						cel
			H 1	H 2	H 3	H 4	H 5	H 6	
1	Badanie procesu mediatyzacji (w tym mediatyzacji polityki), wykorzystanie mediów przez uczestników komunikowania politycznego	Publikacje naukowe, oficjalne dane dotyczące użytkowania mediów (w tym mediów społecznościowych) przez polityków	*						Diagnoza w jakim stopniu politycy oraz komentujący wydarzenia polityczne korzystają z nowych mediów oraz jak proces mediatyzacji zmienił się na przestrzeni ostatnich lat.
2	Analiza deskryptywna polskiej sceny politycznej (zbadanie słowników, klastrów, podobieństwa, ngramów, analiz koszykowych, sentymentów)	Dane z mównicy sejmowej (steno-gramy sejmowe)		*	*	*	*		Zbadanie odpowiednio przygotowanego zbioru danych przeznaczonego do dalszych etapów prac. Rezultat to analiza zbioru danych z podziałem na poszczególne grupy polityczne.

¹⁰⁵ W. McCulloch, W. Pitts, *A Logical Calculus of Ideas Immanent in Nervous Activity*, „Bulletin of Mathematical Biophysics”, 1943, 5 (4), s. 115-133.

¹⁰⁶ Jest to rodzaj uczenia maszynowego, Głębokie uczenie automatyzuje znaczną część procesu wyodrębniania cech, eliminując część wymaganej ręcznej interwencji człowieka. Źródło; Deep learning, <https://www.ibm.com/cloud/learn/deep-learning>, 12.05.2021; J. Schmidhuber, *Deep Learning in Neural Networks: An Overview*, „Neural Networks”, 2015, nr 61, s. 85-117.

¹⁰⁷ N. V. Chawla, K.W. Bowye, L. O. Hall, W. P. Kegelmeyer, *SMOTE: Synthetic Minority Over-sampling Technique*, „Journal of Artificial Intelligence Research”, 2002, nr 16.

¹⁰⁸ D. Dumitrescu, C. Boianiu, *A Study of Image Upsampling and Downsampling Filters*, „MDPI”, 2019, nr 8(2).

3	Analiza deskryptywna osób komentujących wydarzenia polityczne (zbadanie słowników, ngramów, korelacji, sentymentów)	Dane z serwisów społecznościowych (Twitter)		*	*	*		Zbadanie zbioru danych dla osób komentujących wydarzenia polityczne z podziałem na wspomniane grupy polityczne z wykorzystaniem danych z serwisu Twitter.
4	Opracowanie modeli predykcyjnych	Dane z mównicy sejmowej (steno-gramy)		*			*	Stworzenie konfiguracji modeli, które będą klasyfikowały na podstawie określonych słów przynależność do partii politycznych
5	Ocena i testowanie modeli predykcyjnych	Dane z mównicy sejmowej (steno-gramy)					*	Weryfikacja i pomiar wydajności poszczególnych modeli z rekomendacją

Struktura pracy

Praca została podzielona na cztery główne części. Pierwsza część jest poświęcona zarysowi teoretycznemu. W początkowej części skupiono się na opisie mediów oraz ewolucji mediów. Następnie zwrócono uwagę na proces mediatyzacji. Kolejne części w rozdziale pierwszym dotyczą modeli komunikacji a następnie komunikacji politycznej i jej uczestników. Kolejne rozdziały poświęcone są *Big Data*, a całość rozdziału kończy słownik najważniejszych pojęć.

Druga część jest poświęcona mediatyzacji polskiej polityki. Opisane zostały przemiany społeczno-gospodarcze w Polsce po 1989 roku oraz polska scena polityczna (zarys zmian po 1989 roku i stan aktualny). Kolejno przedstawiono rozwój mediów w Polsce w tym okresie, a następnie przemiany komunikacji politycznej pod wpływem mediów cyfrowych i rolę mediów cyfrowych. Zbadano język, jego funkcje i podstawowe struktury, uwzględniając język polityki. Zwrócono uwagę na występujące dyskursy oraz występowanie zjawiska wartościowania w języku oraz pojęcia *językowego obrazu świata*.

Trzecia część odnosi się do aspektów praktycznych, uwagę skupiono na analizie deskryptywnej danych politycznych oraz danych z social mediów. Dokonano analiz statystycznych pokazujących częstości słów oraz zwrotów w postaci *bigramów* i *trigramów*.

Wykonano klasteryzację słów oraz dokonano porównania podobieństwa w poszczególnych grupach.

Czwarta część koncentruje się na budowie modeli predykcyjnych na bazie poprzednich części analiz. W tej części skupiono się na konfigurowaniu modeli oraz ich parametrów. Dokonano porównania wyników oraz rekomendacji wyboru optymalnego modelu.

Rozdział 1. Podstawy teoretyczne pojęć użytych w pracy

W rozdziale pierwszym omówione zostały podstawowe pojęcia używane w pracy, teorie związane z mediami, procesem mediatyzacji, modele komunikacji (w tym komunikowania politycznego). Rozdział opisuje również rolę, jaką pełni język w komunikowaniu. Wprowadzone zostało także pojęcie *Big Data*, wraz z krótkim rysem podejścia do badania dużych zbiorów danych.

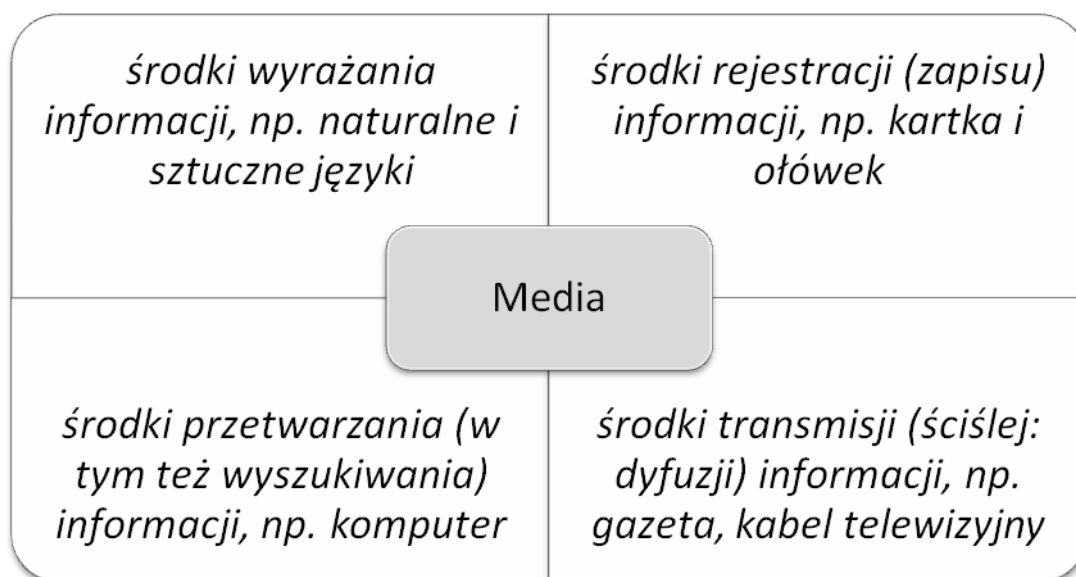
1.1. Media – definicja i podział

Pojęcie *media* ma wiele znaczeń. Jedną z nich prezentuje Walery Pisarek:

Media, środki komunikowania – mimo terminologicznego charakteru nazwa używana (także w literaturze naukowej) w wielu znaczeniach; m. nazywane są: 1) języki narodowe; 2) wszelkie systemy znaków pozwalające zapisywać wiadomości, oparte na różnych rozpoznawalnych sygnałach takich jak brzęczenie, stukanie, pulsowanie; 3) wypowiedzi w gazetach i czasopismach, powieści, słuchowiska itp.; 4) materiał, na którym te wypowiedzi są utrwalane, a więc papier, taśma magnetyczna, płyta kompaktowa; 5) urządzenia, które umożliwiają przekazywanie, jak mikrofon czy kamera; 6) urządzenia, które umożliwiają odbiór, jak telewizor czy radioodbiornik; 7) nośniki informacji, jak gazety, czasopisma, książki, filmy, radio, sieci informatyczne, a także 8) instytucje, które organizują rozpowszechnianie informacji, a więc agencje, wydawnictwa, korporacje radiowe i telewizyjne itp. Najważniejszym sposobem redukcji wieloznaczności terminu m. jest ich podział na: 1) m., czyli środki wyrażania informacji, np. naturalne i sztuczne języki; 2) m., czyli środki rejestracji (zapisu) informacji, np. kartka i ołówek, magnetofon; 3) m., czyli środki transmisji (ściślej: dyfuzji) informacji, np. gazeta, kabel telewizyjny; 4) m., czyli środki przetwarzania (w tym też wyszukiwania) informacji, np. komputer.¹⁰⁹

¹⁰⁹ W. Pisarek, *Słownik terminologii medialnej*, Towarzystwo autorów i wydawców prac naukowych Universitas, Kraków 2006, s. 117.

Rysunek 1. Media i ich podział. Opracowanie własne



Źródło: W. Pisarek, *Słownik terminologii medialnej*, Towarzystwo autorów i wydawców prac naukowych Universitas, Kraków 2006, s. 117.

Współcześnie media są jednocześnie częścią tkanki społeczeństwa i kultury oraz niezależną instytucją, która stoi między innymi instytucjami kulturalnymi i społecznymi i koordynuje ich wzajemne oddziaływanie¹¹⁰. Termin *medium* przeistoczył się w media i nastąpiła prawdopodobna zbieżność trzech znaczeń: (1) *starego ogólnego znaczenia jako pośredniczenia lub materii pośredniczącej*; (2) *znaczenia fachowego, w rozróżnieniu na media drukowane, wizualne i dźwiękowe*; (3) *wyspecjalizowanego znaczenia w kapitalizmie, w którym usługa emisji, gazeta, coś co już istnieje postrzegane jest jako medium dla czegoś innego, na przykład reklamy*¹¹¹. T. Goban-Klas twierdzi, że jako polska kalka językowa *mass media* stały się mediami masowymi (lub massmediami), a w mowie potocznej po prostu mediami¹¹².

¹¹⁰ S. Hjarvard, *The mediatization...*, dz. cyt., s. 8.

¹¹¹ R. Williams, *Keywords*, London 1976, s. 169-170.

¹¹² T. Goban-Klas, *Rwący nurt mediów mediocen – nowa fala mediatyzacji życia społecznego pisma z lat 2012–2020*, Wyższa Szkoła Informatyki i Zarządzania w Rzeszowie, Kraków–Rzeszów 2020, s. 33; T. Goban-Klas, *Termin „media” – krytyka, obrona i objaśnienie, Między Klio a Themis Księga dedykowana Profesorowi Jackowi Sobczakowi*, J. W. Adamowski (red.) T. Wallas (red.), K. Kakarek (red.), Wydawnictwo Naukowe SILVA RERUM, 2016, s. 310-322.

Zanim do tego doszło człowiek w swojej historii komunikowania się przeszedł sześć „epok” – od ery znaków i sygnałów, przez powstanie języka, erę pisma, druku, telekomunikacji, aż po erę komputera (telekomputera)¹¹³.

Pojawienie się druku przyniosło rozwój komunikowania masowego. Ruchoma czcionka zmieniła Europę pod wieloma względami, zarówno aspekty religijne, kulturalne, jak i społeczne. Książki, ulotki polityczne, a następnie gazety rozpoczęły nowy etap komunikowania publicznego. Mimo, że środki komunikowania masowego są znane od prawie 400 lat, to samo pojęcie mediów masowych upowszechniło się w niektórych krajach Europy i Stanach Zjednoczonych dopiero w drugiej połowie XIX wieku. Po prasie, kolejnymi mediami masowymi były radio i telewizja. Media masowe zmieniały się z czasem, a wraz z nimi ich wykorzystanie i społeczeństwo¹¹⁴.

Kolejne lata przynosiły rozwój mediów i zmianę ich charakteru. Znalazło to odbicie w podziale zaproponowanym przez D. McQuaila na „stare media” (książka i biblioteka, media drukowane, film, radio i telewizja, muzyka fonograficzna) i „nowe media” których główną bazą jest Internet¹¹⁵.

Media stare

Termin *stare media* zwykle używany jest w kontraście do nowych mediów. *Stare media* są analogowe, statyczne i liniowe (narzucające sposób czytania). Dodatkowo oparte są o główną instytucję nadawczą, zwykle o jednokierunkowym publicznym przekazie dla biernej mało zróżnicowanej grupy odbiorców.¹¹⁶ Tradycyjne platformy medialne uważały odbiorców lub użytkowników za pasywnych konsumentów produktów medialnych; odbiorcy ideologii medialnej praktycznie byli bez możliwości uzyskania informacji zwrotnej¹¹⁷. Stare media i nowe media są komplementarne ze sobą i są częścią tego

¹¹³ T. Goban-Klas, *Media i komunikowanie i media masowe. Teorie i analizy prasy, radia, telewizji i Internetu*, Wydawnictwo Naukowe PWN, Warszawa 2004, s. 16-17.

¹¹⁴ T. Goban-Klas, *Media i...*, dz. cyt., s. 17-26.

¹¹⁵ D. McQuail, *Teoria komunikowania masowego*, Wydawnictwo Naukowe PWN, Warszawa 2008, s. 43-63.

¹¹⁶ A. Adamski, *Media w...*, dz. cyt., s. 92-93.

¹¹⁷ R. B. Rozario, *New Media and the Traditional Media Platforms: Introspection on the Differences in Technical and Ideological Factors and Audience-integration Patterns between New Media and Traditional Media*, doi.org/10.12724/ajss.26.3, 2013, s. 50.

samego ekosystemu¹¹⁸. Stare media nie są wypierane – ich funkcje i status są raczej zmieniane przez wprowadzanie nowych technologii¹¹⁹. Istnieją też znaczne problemy jeśli chodzi oddziaływanie *mediów starych* w stosunku do *nowych mediów*. P. Levinson wskazuje na przykładzie ogłoszenia informacji o śmierci Tima Russerta, że w przypadku działania telewizji i dziennikarzy *NBC* oraz innych stacji jest zgoła inny niż działania *nowych mediów*. Stacje telewizyjne czekały na ogłoszenie tej wiadomości po zaakceptowaniu dyrektora (stare media zawierają element dozorczy) o konkretnej godzinie, podczas gdy w internetowej encyklopedii *Wikipedia* informacja została zaktualizowana od razu bez żadnej odgórnej konsultacji¹²⁰.

Z punktu widzenia tematu niniejszej pracy, warto spojrzeć na te media masowe, które odegrały szczególną rolę w komunikowaniu politycznym.

Książka drukowana jest uznawana za wynalazek rewolucyjny, lecz na samym początku wykorzystywana była do reprodukcji tekstów, jako usprawnienie kopiowania ręcznego.¹²¹ Konsekwencją używania tego medium było upowszechnienie i powstanie książek tłumaczonych na narodowe języki. Za pierwszą książkę drukowaną uważa się *Biblię Gutenberga* (lata 50. XV wieku). W okresie średniowiecza książkę traktowano jako magazyn mądrości, a nie jako środek komunikacji¹²². Biblioteka jako zbiór książek stanowi potwierdzenie idei książki jako zapisu danych¹²³. Książkę jako medium, według D. McQuaila charakteryzuje: *technologia ruchomej czcionki, zszywane strony, forma kodeksu, wiele kopii, towarowy charakter, różnorodna (świecka) treść, indywidualne wykorzystani, dążenie do wolności wydruku, indywidualne autorstwo*¹²⁴.

Media drukowane i gazeta. Gazeta wyłoniła się z ulotek i biuletynów końcem XVI i XVII wieku. Na początku stanowiła przekaz w odpowiedniej formie dla sfery publicznej. Wczesną formę gazety cechowała nieregularność ukazywania, orientacja handlowa, mnogość celów (wiadomości, plotek i rozrywki). W XVII wieku biuletyn stanowił

¹¹⁸ J. Walker, *Old Media and New Media: Like It or Not, They're Partners*, Reason, September 15, 2004.

¹¹⁹ H. Jenkins, *Convergence Culture. Where Old and New Media Collide*, NEW YORK UNIVERSITY PRESS, 2006, s. 14.

¹²⁰ P. Levinson, *Nowe nowe...* dz. cyt., s. 149-151.

¹²¹ D. McQuail, *Teoria komunikowania...*, dz. cyt., s. 45.

¹²² Tamże, s. 45.

¹²³ Tamże, s. 45.

¹²⁴ Tamże, s. 45.

kompilację wykonaną przez drukarza (i wydawcy zarazem). Wydawnictwa występowały w formie oficjalnej, publikowały informacje rządowe, czy monarchów i miały podobne cech co biuletyn, a przede wszystkim były głosem władzy. Biuletyn handlowy ukształtował pojęcie gazety oraz miał swój wpływ na komunikację i jej odbiór – gazeta zaczęła być kierowana do anonimowych czytelników, a nie jak wcześniej wykorzystywana tylko jako narzędzie propagandowe. Prasa drukarska spowodowała rozwój innych form publikacji takich, jak piosenki, komiksy, raporty, mapy, plakaty, ulotki oraz wiele innych. Gazetę jako medium charakteryzuje według D. McQuaila: *regularne częste ukazywanie się, towarowy charakter, treść związana z aktualnymi wydarzeniami, funkcje w sferze publicznej, miejska, świecka społeczność odbiorców, względna swoboda*.¹²⁵ W pracy *Wymyślanie wrogów i inne teksty okolicznościowe* Umberto Eco twierdzi, że nowa rola prasy (obecnie już stosowana) polega na tym, że *zamiast notować istotne informacje (kiedyś władza decydowała, które informacje zasługują na uwagę, wypowiadały wojny, obniżały kurs walut i podpisywały porozumienia), gazety dokonują autonomicznego wyboru tych, które powinny uchodzić za znaczące, oraz tych wymagających przemilczenia, czasem wręcz konsultując (do czego już doszło) z elitami politycznymi, które «tajemnice» należy wyjawiać, a których nie tykać*¹²⁶. Ciekawym przypadkiem jest relacja prasy z Kościołem katolickim gdzie medialna doktryna Kościoła od dawna charakteryzowała się nieufnością. Było to widoczne zwłaszcza w XIX wieku gdzie było to spowodowane wieloma czynnikami, na m.in. że prasa (już w nowoczesnej wersji) była często wykorzystywana przez wrogów Kościoła¹²⁷.

Kolejnym medium był film, którego powstanie datuje się na koniec XIX wieku. Cechował go szybki rozwój i docieranie do znacznych grup osób. Medium nie jest ściśle powiązane z niniejszą pracą, dlatego szczegółowy opis pominięto. Po filmie, kolejne media, które rozwinęły się na dużą skalę to radio i telewizja. Obydwie formy powstały z innych technologii: telegrafu, telefonu, fotografii i nagrań dźwiękowych. W 1943 Sąd Najwyższy Stanów Zjednoczonych przyznał prawa patentowe dla Nikoli Tesli po jego śmierci – jego patent przesyłania fal radiowych był gotowy 1900 roku. Powszechnie za

¹²⁵ Tamże, s. 47.

¹²⁶ U. Eco, *Wymyślanie wrogów*, Rebis, 2011, s. 335; M. Palczewski, *Umberto Eco o mediach – w esejach, rozprawach i w powieści*, „Media, Kultura, Społeczeństwo”, 2014, nr 9-10, s. 19-29.

¹²⁷ A. Adamski, B. Przywara, S. Przybyło, *The Church 'S Attitude To The Means And Philosophy Of Communication In The 9th And 16th Centuries: Courage And Trust Versus Escape And Censorship*, „Konstantinove listy/Constantine s Letters”, nr 13(2), 2020, s. 111-123.

twórcę radia również uznaje się Guglielmo Marconiego.¹²⁸ Utworzenie w Wielkiej Brytanii 1922 roku rozgłośni BBC (*British Broadcasting Company*), zdaniem T. Goban-Klasa można uznać za rozpoczęcie ery radia jako masowej jednokierunkowej łączności bezprzewodowej¹²⁹. Następnie rozwój radiofonii postępował w bardzo szybkim tempie i nie wyhamowało go nawet zawieszenie wielu prywatnych radiostacji przez rząd USA na czas I wojny światowej¹³⁰. Telewizja, według McQuaila ma status „najbardziej umasowionego z mediów ” mimo tego, że czas spędzany przed telewizorem jest zróżnicowany w poszczególnych państwach¹³¹. Za datę powstania telewizji amerykańskiej uznaje się rok 1928. W Polsce w 1952 odbyła się pierwsza po wojnie transmisja telewizyjna. Jako cechy radia McQuail uznaje: *elastyczna i tania produkcja, wygoda korzystania, zróżnicowanie treści, względna swoboda, zindywidualizowany użytek, potencjał współuczestnictwa*, telewizji zaś: *bardzo duży zasięg i publiczność, przekaz audiowizualny, złożona technika i organizacja, publiczny i szeroki zakres regulacji, narodowy i międzynarodowy charakter, zróżnicowane formy przekazu*¹³². W Polsce telewizja postrzegana jest jako jeden z aktorów politycznych podlegających ocenie społecznej w zakresie jej wiarygodności i zaufania¹³³. Giovanni Sartori, który uważa telewizję za jeden z najpotężniejszych instrumentów kształtowania opinii: *Dzisiaj suwerenny naród zazwyczaj «wyraża sądy» w zależności od tego, jaką opinię podpowie mu telewizja kierując zaś opiniami, telewizja doprawdy rozciąga swą władzę nad centrum, gdzie zachodzą wszystkie procesy współczesnej polityki*¹³⁴.

Według Katarzyny Maciejewskiej-Mieszkowskiej *telewizja jest nie tylko zasadniczym kanałem przepływu informacji, lecz także w określonych warunkach może stać się niebezpiecznym instrumentem wpływu na nie zawsze świadomych manipulacji obywateli*¹³⁵. Zdaniem autorki, zjawiskiem, które może niepokoić współcześnie jest

¹²⁸ Rozprawa sądowa przebiegała po śmierci wynalazcy, w niektórych publikacjach za twórcę radia uznaje się G. Marconiego, mimo że przyznał się do wykorzystania wcześniejszych prac N. Tesli w zbudowaniu swojej wersji radia.

¹²⁹ T. Goban-Klasa, *Zarys historii i rozwoju mediów. Od malowideł naskalnych do multimedialnych*. Wydawnictwo Naukowe Akademii Pedagogicznej w Krakowie, Kraków 2001, s. 41.

¹³⁰ A. Adamski, *Media w ...*, dz. cyt., s. 267.

¹³¹ D. McQuail, *Teoria komunikowania ...*, dz. cyt., s. 54.

¹³² D. McQuail, *Teoria komunikowania ...*, dz. cyt., s. 54.

¹³³ K. Maciejewska-Mieszkowska, *Telewizja jako aktor polityczny w ocenie społecznej Polaków*, *Media Biznes Kultura*, 2020, nr 2 (9), s.1.

¹³⁴ G. Sartori, *Homo videns. Telewizja i post-myślenie*, przeł. J. Uszyński, Telewizja Polska, Warszawa 2005, s. 38.

¹³⁵ K. Maciejewska-Mieszkowska, *Telewizja ...*, dz. cyt., s. 118.

dezorientacja odbiorców potwierdzających – 80% respondentów uważa, iż przekazy informacyjne powodują, że odbiorcy „nie wiedzą, gdzie leży prawda”, a takie zjawisko i dezorientacja w tym aspekcie może powodować zniechęcenie odbiorców do odbierania i pogłębiania informacji na temat rzeczywistości politycznej¹³⁶.

Nowe Media

Termin *nowe media* w naturalny sposób jest związany z terminem *stare media*. *Stare Media* prezentują cechy analogowości, statyczności i sekwencyjności, *nowe media* są cyfrowe, dynamiczne i interaktywne¹³⁷. Wyrażenie *nowe media* jest używane od lat 60. XX wieku, ale ciągle ewoluuje. S. Livingstone, L. Lievrouw¹³⁸ wskazują na złożoność i łączą informatyczne technologie komunikacyjne wraz z społecznymi kontekstami oraz trzema elementami (*artefakty i urządzenia techniczne; działania, praktyki; układy i organizacje społeczne*). Według McQuaila nowe media cechuje: wzajemne powiązanie, interaktywność, wszechobecność, delokalizacja, otwartość, dostęp do indywidualnych użytkowników (jako nadawców i odbiorców)¹³⁹. Terminem *nowe media* zaczęto posługiwać się około lat 90. XX wieku, głównie w środowisku uczelnianym oraz publicystycznym związanym z komputerowym czasopismem *Wired*¹⁴⁰. W. Pisarek twierdzi, że pojęcie *nowe media*, jak i jego zakres zmienia się w czasie. W latach 80. XX to pojęcie reprezentowało *środki komunikowania masowego o cechach zarówno tradycyjnych (pośredniość, jednoczesność lub krótki przedział czasowy między nadaniem a odbiorem, masowość, rozproszenie terytorialne publiczności, różnorodność zawartości, ogólna dostępność odbioru) jak i nowych mediów (aktywność odbiorcy, interaktywność, w tym także dialogowość)*.¹⁴¹ Prasa, radio, telewizja opierały się na jednokierunkowym przekazie o publicznym i otwartym charakterze w kierunku biernych odbiorców – *nowe media* zmieniają obieg informacji ku modelom interakcyjnym¹⁴².

¹³⁶ K. Maciejewska-Mieszkowska, *Telewizja jako ...*, dz. cyt., s. 128.

¹³⁷ A. Adamski, *Media w...*, dz. cyt., s. 93.

¹³⁸ S. Livingstone, L. Lievrouw, *The Handbook...*, dz. cyt.

¹³⁹ D. McQuail, *Teoria komunikowania...*, dz. cyt., s. 57.

¹⁴⁰ Z. Bauer, *Dziennikarstwo wobec nowych mediów*, Universitas, 2009, s. 129-130.

¹⁴¹ W. Pisarek, *Wstęp do nauki o komunikowaniu*, Wydawnictwa akademickie i profesjonalne, 2008, s. 147-148.

¹⁴² T. Goban-Klas, *Pochwała medioznawstwa czyli renesans McLuhana*, Studia Medioznawcze, 2000, s. 22.

Interaktywność może mieć różne znaczenia¹⁴³. Spiro Kiouisis stworzył *definicję operacyjną* interaktywności zawierającą 4 kryteria: *bliskość społeczną wobec innych, postrzeganą szybkość, aktywizację zmysłów i obecność na odległość*.¹⁴⁴ Według Edward J. Downes, Sally J. McMillan istnieje pięć wymiarów interaktywności¹⁴⁵:

- *Kierunek komunikowania* – zrywanie z jednokierunkowym przekazem na rzecz komunikacji dwukierunkowej;
- *Elastyczność czasowa i role podczas wymiany* – płynność w komunikacji między uczestnikami;
- *Poczucie miejsca w otoczeniu komunikacyjnym*;
- *Poziom kontroli (swoboda komunikacyjna)* – uczestnicy mają większą kontrolę w środowisku komunikacyjnym;
- *Postrzeganie celu* – orientacja na wymianę informacji. Uczestnicy komunikują się w sposób responsywny (reagują).

Nowe media również zmieniły obieg informacji. Jan L. Bordewijk oraz Ben van Kaam stworzyli model ułatwiający badanie zmian, który zawiera wzory obiegu informacji¹⁴⁶:

- *Alokucja* – informacja rozprzestrzenia się od centrum do wielu odbiorców, którzy mają niewielkie szanse na nadanie komunikatu zwrotnego (przykład: koncert). Rodzaj ten jest reprezentowanych przez typowe *stare media*;
- *Rozmowa (konwersacja) i wymiana* – jednostki wchodzi w interakcje wzajemnie ze sobą pomijając pośredników. Strony w konwersacji są równe i może ich być więcej niż dwie;

¹⁴³ S. Kiouisis, *Interactivity: a concept explication*, New Media & Society, 2002.

¹⁴⁴ D. McQuail, *Teoria komunikowania...*, dz. cyt., s. 157.

¹⁴⁵ E. J. Downes, S. J. McMillan, *Defining Interactivity: A Qualitative Identification of Key Dimensions*, New Media and Society, 2000, s. 157-179.

¹⁴⁶ J. L. Bordewijk, B. van Kaam, *Towards a new classification of teleinformation services*, [w:] D. McQuail (red.), *McQuail's Reader in Mass Communication Theory*, London, 1986.

- Konsultacja – jednostka szuka informacji w centralnym zbiorze np. w bibliotece;
- Rejestracja – odwrotność konsultacji, czyli centrum wysyła żądanie otrzymując informację od uczestnika znajdującego się na peryferiach (przykładem mogą być systemy alarmowe).

Media społecznościowe jako część nowych mediów

P. Levinson ukuł w 2007 termin *nowe nowe media* po tym, jak zdał sobie sprawę, że studenci nie chcą rozmawiać o emailu, HTML czy podstawach Internetu (były to pojęcia związane z nowymi mediami w latach 90. XX wieku) tylko o blogach, Facebooku czy YouTube¹⁴⁷. *Nowe media* sprzyjają nawiązywaniu wspólnot oraz komunikowaniu się grup w tych wspólnotach. Wspólnoty te można określić jako społeczność wirtualną utworzoną w Internecie przez jednostki z własnej inicjatywy lub jako odpowiedź na sygnał¹⁴⁸. Społeczność to celowo ustawiona zbiorowość przez osoby o podobnych zainteresowaniach¹⁴⁹ również o nieformalnym charakterze. Społeczności wirtualne mogą aplikować cechy społeczności rzeczywistej (na przykład interakcja pomiędzy osobami, wspólne cele, poczucie przynależności do grupy)¹⁵⁰. Ewolucję social mediów, w których gromadziły się społeczności przedstawia poniższa tabela:

Tabela 2. Ewolucja social mediów

Lata	Opis
1970 - 1980	System <i>PLATO</i> (opracowany na Uniwersytecie Illinois, a następnie sprzedawany przez <i>Control Data Corporation</i>) oferuje wczesne formy mediów społecznościowych, aplikacje, forum i wiadomości – <i>PLATO</i> , <i>TERM-talk</i> przysyłanie wiadomości, <i>Talkomatic</i> czat online, <i>News Report</i> internetowa gazeta i blog oraz listy dostępu umożliwiające właścicielowi pliku ograniczenie dostępu do określonej grupy użytkowników. T. Truscott i J. Ellis tworzą pojęcie <i>Usenetu</i> (ogólnoświatowy system grup dyskusyjnych) w 1979 roku, na bazie pomysłu, który rodzi się w 1980 roku.
1980 -	Systemy operacyjne z graficznym interfejsem użytkownika (Windows, Mac OS) zaczynają pojawiać się i zyskują popularność.

¹⁴⁷ P. Levinson, *Nowe nowe...*, dz. cyt., s.7.

¹⁴⁸ H. Rheingold, *The Virtual Community*, London, Seeker and Warburg, 1994.

¹⁴⁹ T. Lindlof, J. Schatzer, *Media ethnography in virtual space: strategies, limits and possibilities*, Journal of Broadcasting and Electronic Media, 1998, nr 42 (2), s. 170-189.

¹⁵⁰ D. McQuail, *Teoria komunikowania...*, dz. cyt., s. 162.

1990	<i>Bulletin Board Systems</i> (BBS) – komputer zaopatrzony w odpowiednie oprogramowanie, oczekujący na połączenie z nim poprzez łącze modemowe linii sieci telefonicznej; znając numer telefonu do danego BBS-a, można podłączyć się do niego i używać zasobów (na przykład poczty elektronicznej, grup ogłoszeniowo-dyskusyjnych). <i>Internet Relay Chat</i> (IRC) – jedna ze starszych usług sieciowych, która umożliwia rozmowę na tematycznych kanałach komunikacyjnych oraz prywatnie z inną osobą. <i>WELL</i> (<i>Whole Earth 'Lectronic Link</i>) – założona w 1985 roku; jedna z najstarszych działających społeczności internetowych.
2000 - 2010	Rozwijane są różne platformy mediów społecznościowych, takie jak <i>Myspace</i> i <i>Facebook</i>. Blogowanie zaczyna zdobywać popularność. Platformy wiadomości w czasie rzeczywistym, takie jak <i>Windows Live Messenger</i>, również stają się coraz bardziej popularne. Rok 2001 przynosi uruchomienie <i>Wikipedii</i>. W Polsce jednymi z najpopularniejszych komunikatorów staje się <i>Gadu-Gadu</i> i <i>Tlen</i>¹⁵¹, a najpopularniejszą platformą społecznościową kolejno <i>Grono.net</i>, <i>Fotka.pl</i>, <i>Nasza Klasa</i> (później <i>Nk.pl</i>) mająca w swoim szczycie popularności 14 mln użytkowników¹⁵².
2010 - teraz	Następuje rozkwit platform mediów społecznościowych takich jak Facebook, LinkedIn, Myspace, Snapchat, Instagram i Twitter. Serwis Facebook posiada w II połowie 2021 roku 2,8 miliarda użytkowników, serwis Twitter 330 mln użytkowników.¹⁵³ Również TikTok prezentujący krótkie treści video ma ponad 1 miliard aktywnych użytkowników¹⁵⁴. Dzięki postępowi w technologii i rozwojowi rynku smartfonów prawie wszystkie serwisy społecznościowe mają własne aplikacje mobilne.

Źródło: opracowanie własne na podstawie danych z *businessinsider.com.pl*¹⁵⁵, *antyweb.pl*¹⁵⁶ oraz *Wikipedia*¹⁵⁷.

Z uwagi na temat niniejszej pracy, warto przyjrzeć się specyfice medium jakim jest Twitter. Twitter jedną z największych platform społecznościowych, których użytkowanie koncentruje się na mikro-blogowaniu i komentowaniu często bieżących wydarzeń¹⁵⁸.

¹⁵¹ Pierwsze polskie komunikatory- przeżyjmy to jeszcze raz, <https://antyweb.pl/pierwsze-polskie-komunikatory-przezyjmy-to-jeszcze-raz>, 12.08.2021.

¹⁵² Nasza klasa znika z sieci, <https://businessinsider.com.pl/technologie/nasza-klasa-znika-z-sieci-w-szczycie-miala-14-mln-uzytkownikow/hswvc80>, 12.08.2021.

¹⁵³ Użytkownicy Social Media w Polsce i na Świecie, <https://www.whysosocial.pl/uzytkownicy-social-media-w-polsce-i-na-swiecie>, 12.08.2021.

¹⁵⁴ TikTok User Statistics 2022, backlinko.com, 20.11.2022.

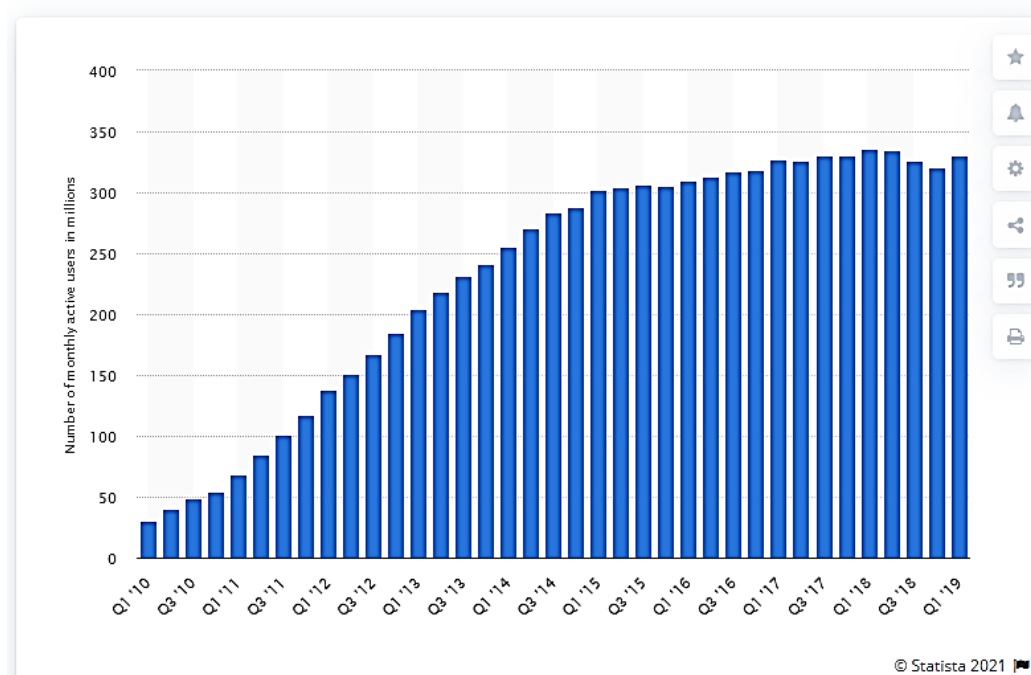
¹⁵⁵ Nasza klasa..., dz. cyt., 12.08.2021.

¹⁵⁶ Pierwsze polskie..., dz. cyt., 12.08.2021.

¹⁵⁷ Timeline of social media, https://en.wikipedia.org/wiki/Timeline_of_social_media, 12.08.2021.

¹⁵⁸ Użytkownicy Social..., dz. cyt., 12.08.2021.

Rysunek 2. Liczba aktywnych użytkowników Twittera miesięcznie na całym świecie od 1 kwartału 2010 do 1 kwartału 2019 – w milionach.



Źródło: www.statista.com, 11.05.2021.

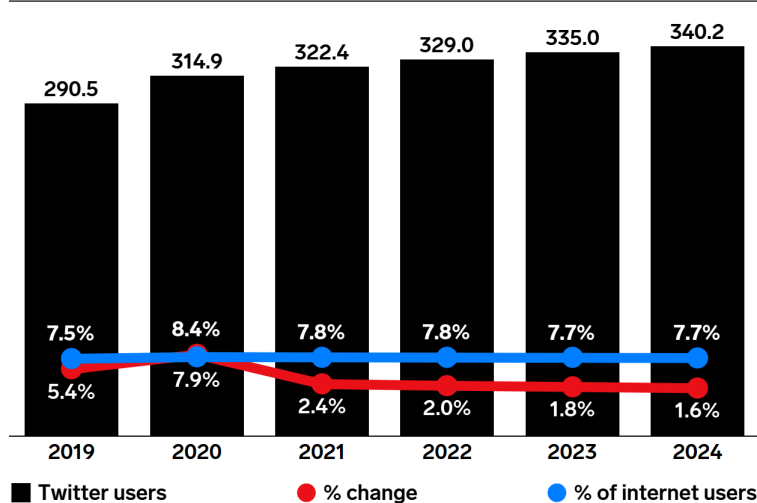
Aktualnie Twitter liczy około 320 mln użytkowników. Wartości pokazują (Rysunek 2.), że liczba użytkowników Twittera rosła do końca 2015, w kolejnych latach dynamika wzrostu jest nieco niższa niż wcześniej.

Rolę Twittera w codziennym życiu pokazuje badanie źródeł, z których ludzie czerpią codzienne informacje. Najwięcej użytkowników Internetu wykorzystuje platformę Facebook (35%), zdecydowanie mniej – YouTube (21%) i WhatsApp (16%). 12% użytkowników używa do tego celu Twittera.¹⁵⁹ Chociaż na przestrzeni lat obserwuje się przyrost użytkowników Twittera, to znacznie większą dynamiką odznacza się Instagram (Rysunek 3. pokazuje dynamikę wzrostu Twittera i Instagrama).

¹⁵⁹ Twitter Statistics, <https://blog.hootsuite.com/twitter-statistics/>, 12.08.2021.

Rysunek 3. Liczba użytkowników Twittera na całym świecie.

Twitter Users Worldwide, 2019-2024
millions, % change, and % of internet users



Note: internet users of any age who access their Twitter account via any device at least once per month

Source: eMarketer, October 2020

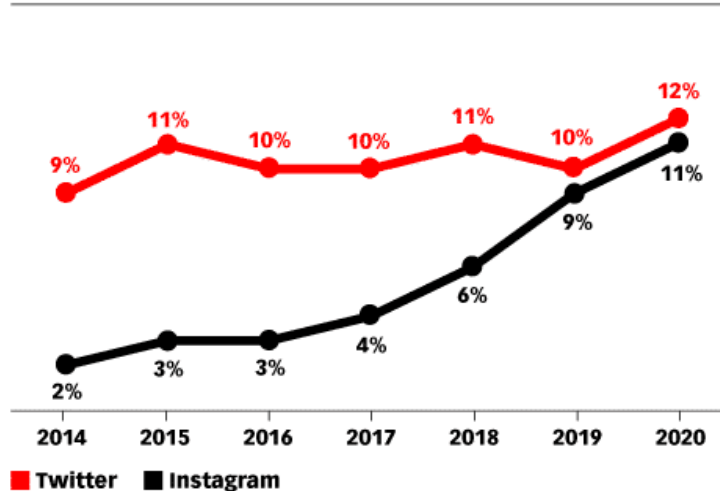
T11270

eMarketer | InsiderIntelligence.com

Źródło: Twitter Statistics, <https://blog.hootsuite.com/twitter-statistics>, 13.08.2021.

Rysunek 4. Najczęstsze źródła dostępu do informacji, Badanie Reuters Institute for the Study of Journalism at the University of Oxford, Digital News Report 2020

Internet Users Worldwide* Who Use Twitter vs. Instagram to Access News, 2014-2020
% of respondents



Note: ages 18+; in the past week; *average across the countries of Australia, Brazil, Denmark, Finland, France, Germany, Italy, Ireland, Japan, Spain, United Kingdom and United States

Source: Reuters Institute for the Study of Journalism at the University of Oxford, "Digital News Report 2020" conducted by YouGov, June 15, 2020

256384

www.eMarketer.com

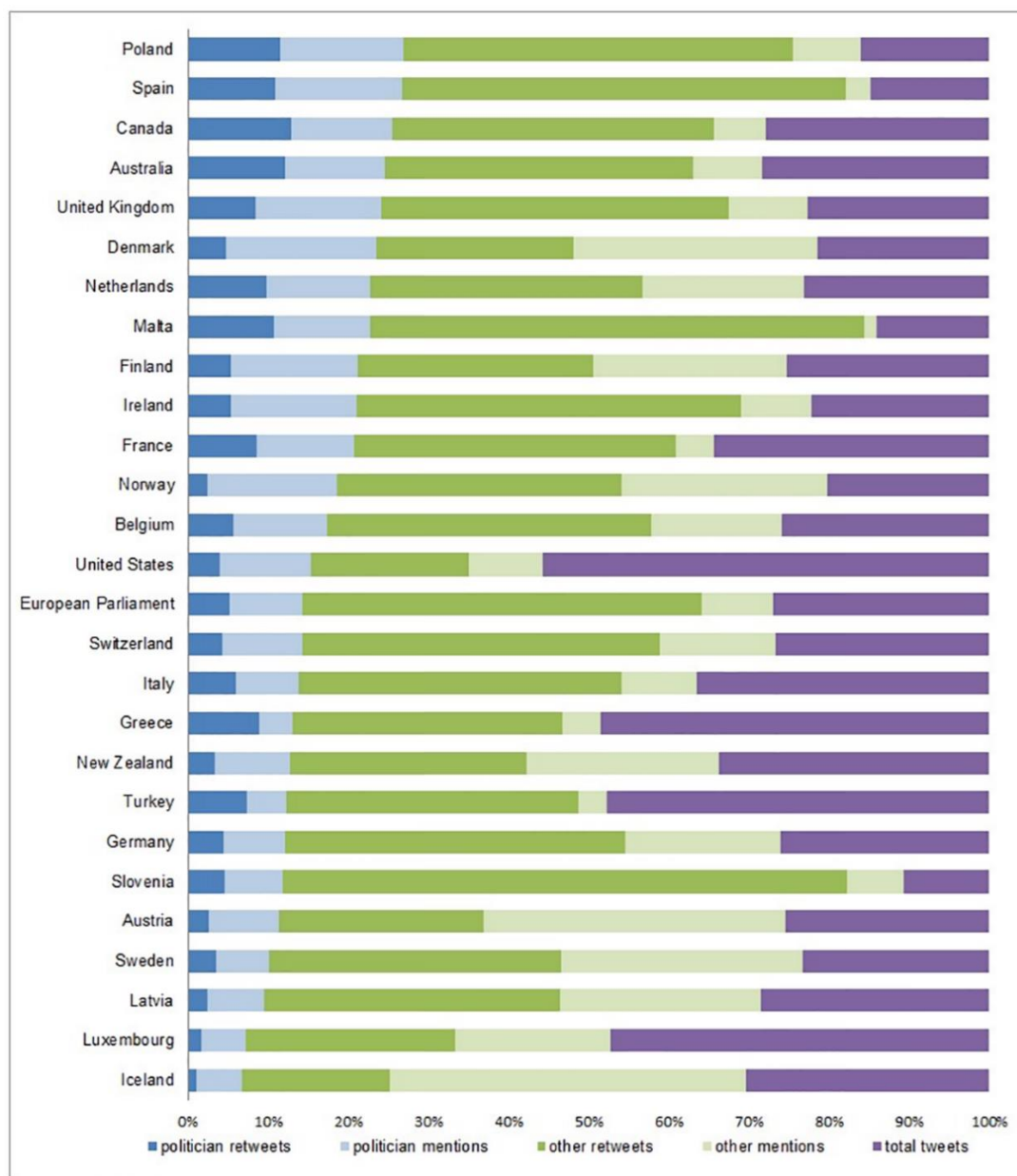
Źródło: Twitter Statistics, <https://blog.hootsuite.com/twitter-statistics>, 12.08.2021.

C. Ducombe w swoim badaniu z 2019 roku wskazała, jak ważne jest komunikowanie za pośrednictwem tej platformy również w polityce¹⁶⁰. Autorka wskazała jak często tematy polityczne były poruszane w *tweetach* i jaką wartość procentową stanowiły. W Polsce występował największy odsetek tweetów politycznych wraz z wzmiankami politycznymi, na drugim miejscu znalazła się Hiszpania, a na trzecim Kanada (Rysunek 5.).

Podsumowując, Twitter zawiera interesującą bazę komunikatów politycznych szczególnie z obszaru Polski. Jest platformą, która notuje stały wzrost liczby użytkowników i stanowi również idealną bazę dla analiz danych tekstowych.

¹⁶⁰ C. Ducombe, *The Politics ...*, dz. cyt.

Rysunek 5. Udział tweetów ze względu na rodzaj w różnych państwach



Źródło: C. Ducombe, *The Politics of Twitter: Emotions and the Power of Social Media*, Oxford Academic, 2019.

1.2. Mediatyzacja

Mediatyzacja (rzeczywistości społecznej) (ang. mediatisation, niem. Mediatisierung) – proces pośrednictwa MEDIÓW w poznawaniu świata; wpływanie przez media na postrzeganie przez człowieka rzeczywistości niedostępnej bezpośredniemu doświadczeniu wraz z konsekwencjami takiego pośrednictwa: kształtowaniem obrazu całej

*rzeczywistości społecznej, a nawet kompleksowych doświadczeń społecznych pod wpływem konstrukcji medialnych.*¹⁶¹

Według S. Hjarvarda pojęcie to było używane w różnych kontekstach do opisu wpływu mediów na różne zjawiska, ale nie włożono odpowiednio dużo pracy, aby zdefiniować termin i zająć się rozwijaniem teoretycznego konceptu¹⁶². Przedstawia on kilka kluczowych punktów klarowania zjawiska i pojęcia mediatyzacji.

Na początkowym etapie mediatyzacja była powiązana z wpływem na polityczną komunikację oraz aspekty z nią związane¹⁶³. O wpływie na politykę pisał K. Asp, który mówił o mediatyzacji życia politycznego jako procesie, w ramach którego „system polityczny” w znacznym stopniu jest pod wpływem dostosowana do wymagań środków masowego przekazu¹⁶⁴. Mediatyzacja jako zbiór narzędzi wykorzystywanych w polityce zaczęła odgrywać kluczową rolę stając się kluczowym aspektem polityki¹⁶⁵.

David L. Altheide i Robert P. Snow w 1979 wskazywali, że logika jest kluczową wartością i ona tworzy zasób, który jest generowany i krąży w społeczeństwie, stworzyli również pojęcie logiki mediów¹⁶⁶. W swoich badaniach wspominają o innych aspektach logiki mediów, poruszając aspekty technologiczne i organizacyjne. Zwracają uwagę na instytucje społeczne, takie jak sport czy religia, ale ich główny nacisk, zarówno w sposób ilościowy i jakościowy, kładziony jest na dużą rolę mediów w komunikacji politycznej.¹⁶⁷

G. Hernes wprowadził określenie „społeczeństwo pokręcone w mediach” (ang. *media-twisted society*). Twierdził, że media miały fundamentalny wpływ na wszystkie instytucje społeczne i ich wzajemne relacje¹⁶⁸. Co prawda nie użył terminu mediatyzacja w swojej koncepcji *społeczeństwa pokręconego w mediach*, ale holistycznie spojrzał na społeczeństwo i pod wieloma względami zgodził się z koncepcją mediatyzacji.¹⁶⁹

¹⁶¹ W. Pisarek, *Słownik ...*, dz. cyt., s. 118.

¹⁶² S. Hjarvard, *The mediatization of culture ...*, dz. cyt., s. 8.

¹⁶³ S. Hjarvard, *The mediatization of culture...*, dz. cyt.

¹⁶⁴ K. Asp, *Måktiga massmedier ...*, dz. cyt.

¹⁶⁵ M. Ampuja, J. Koivisto, E. Välvirronen, *Strong and weak forms of mediatization theory: A critical review*, 2014, s. 112.

¹⁶⁶ S. Hjarvard, *The mediatization of culture ...*, dz. cyt., s. 9.

¹⁶⁷ S. Hjarvard, *The mediatization of Society...* dz. cyt., s. 105-134.

¹⁶⁸ G. Hernes, *Det mediavridde ...*, dz. cyt.

¹⁶⁹ S. Hjarvard, *The Mediatization of Society ...*, dz. cyt., s. 105-134.

Mediatyzacja widoczna jest również w świecie celebrytów, powodując, że rola liderów, ekspertów z różnych dyscyplin i dziedzin zostaje przejęta przez sławne osoby. Obecnie obserwujemy proces celebrytyzacji zwykłych ludzi – zjawiska którego obecny kształt warunkowany jest przez trzy uzupełniające się czynniki: mediatyzację, personalizację i utowarowienie („wszystko jest towarem”) ¹⁷⁰.

G. Mazzoleni i W. Schulz (1999) odnieśli koncepcję mediatyzacji do wpływu mediów na politykę¹⁷¹. Pokazali rosnący wpływ środków masowego przekazu na uprawianie polityki na przykładach wykorzystywania telewizji w brazylijskiej kampanii wyborczej w 1989 r., postaci Silvio Berlusconi w kontekście wykorzystania mediów w drodze do władzy we Włoszech i przypadku Tony'ego Blaira w Anglii. Scharakteryzowali mediatyzację jako *problematiczne współistnienie lub konsekwencje rozwoju nowoczesnych środków masowego przekazu*¹⁷², ale podkreślili też, że instytucje takie jak parlament stały się coraz bardziej zależne od mediów oraz musiały dostosować się do nich.¹⁷³ Mediatyzacja polityki rządzi się podobnymi charakterystykami jak proces celebrytyzacji. Może prowadzić do sytuacji, w której polityk zostaje celebrytą. Mówiąc o mediatyzacji polityki, należy zwrócić uwagę na zmiany zachodzące w zachowaniu i wypowiedziach przez polityków pod wpływem mediów¹⁷⁴.

Koncepcja mediatyzacji została wykorzystana także do badań nad rosnącą rolą marketingu i kulturą konsumpcyjną¹⁷⁵. André Jansson za początkowy punkt przyjął ogólną mediatyzację współczesnej kultury, określając ją jako „proces”, *dzięki któremu pośrednie produkty kultury zyskały na znaczeniu jako kulturowe referencje, aby przyczynić się do rozwoju i utrzymania wspólnot kulturowych*¹⁷⁶. Aktualnie media zajmują kluczową pozycję jako dostawcy produktów i wierzeń kulturowych.

¹⁷⁰ O. Driessens, *The celebrityization of society and culture: Understanding the structural dynamics of celebrity culture*, „International Journal of Cultural Studies”, 2013, nr 6 (16), s. 641.

¹⁷¹ G. Mazzoleni, W. Schulz, *Mediatization of ...*, dz. cyt. s. 247-261.

¹⁷² G. Mazzoleni, W. Schulz, *Mediatization of ...*, dz. cyt., s. 247-261.

¹⁷³ S. Hjarvard, *The Mediatization of Society...*, dz. cyt., s. 105-134.

¹⁷⁴ A. Adamski, *Media as the Intersphere of Human Life: Another View on the Mediatization of Communication Theory*, (red.) D. Petranová, S. Magá, „MEGATRENDS AND MEDIA Media Farm – Totems and Taboo”, 2015, s. 26.

¹⁷⁵ A. Jansson, *The Mediatization of Consumption: Towards an Analytical Framework of Image Culture*, „Journal of Consumer Culture”, 2002, nr 2(1), s. 5-31.

¹⁷⁶ S. Hjarvard, *The Mediatization of Society...*, dz. cyt., s. 105-134.

Pojęcie mediatyzacji było używane w obszarze wpływu mediów na obszar badań. Esa Väliverronen nie traktuje mediatyzacji za „ściśłą” w pojęciu analitycznym, a raczej jako termin niejednoznaczny, odnoszący się do narastającej kultury społecznej środków masowego przekazu i innych form komunikacji¹⁷⁷. Co za tym idzie media odgrywają ważną rolę również w generowaniu i obiegu wiedzy oraz interpretacji nauki i badań. P. Weingart postrzega to jako decydujący element w relacji między mediami i nauką.¹⁷⁸ Mówi o podstawie dla mediatyzacji nauki. Wraz ze wzrostem znaczenia mediów w kształtowaniu opinii publicznej – nauka będzie coraz bardziej zorientowana na media.¹⁷⁹

Wpływ mediatyzacji na religię można odebrać jako stworzenie nowych warunków społecznych, w których sposób definiowania i praktykowania religii ulega przekształceniu¹⁸⁰. Wielkość i intensywność mediatyzacji religii będzie różna się w zależności od położenia geograficznego, obszaru, kraju oraz rozwój infrastruktury medialnej¹⁸¹. Media stały się ważnym, wiodącym źródłem informacji o wydarzeniach religijnych, wiadomościach¹⁸². Podczas wydarzeń rytualnych często można zauważyć interakcję między mediami a kościołem. Niezależnie od tego, czy jest to katastrofa, czy uroczystość narodowa, transmisja ceremonii religijnej odgrywa z jednej strony niewielką, ale z drugiej ważną rolę¹⁸³. Większość relacji medialnych o tego typu wydarzeniach jest prowadzona przez same media przy użyciu ich tradycyjnych formatów i gatunków takich jak wywiady, filmy, wiadomości i komentarze na żywo¹⁸⁴.

John B. Thompson postrzegał rozwój mediów jako integralną część rozwoju współczesnego społeczeństwa¹⁸⁵. Mówił o „zapośredniczeniu kultury współczesnej” (*ang. mediatization of modern culture*). Wynalezienie prasy drukarskiej w połowie XV przyniosło

¹⁷⁷ E. Väliverronen, *From Mediation to Mediatization: The New Politics of Communicating Science and Biotechnology*, [w:] U. Kivikuru, T. Savolainen (red.) *The Politics of Public Issues*. Helsinki, Department of Communication, University of Helsinki, 2001, s. 159.

¹⁷⁸ P. Weingart, *Science and the Media*, „Research Policy”, 1998, nr 27(8), s. 869-879.

¹⁷⁹ S. Hjarvard, *The Mediatization of Society...*, dz. cyt., s. 105-134.

¹⁸⁰ S. Hjarvard, *The Mediatization of Culture and ...* dz. cyt, s. 10.

¹⁸¹ S. Hjarvard, *The Mediatization of Society. A Theory of the Media as Agents of Social and Cultural Change*, *Nordicom Review* 29, 2008, nr 2, nr 29(2), s. 9-11.

¹⁸² K. Niemelä, *Religion in the Nordic media: introduction*. In *Nordic Journal of Religion and Society*, 2013, nr 26(1), s. 1.

¹⁸³ S. Hjarvard, *The mediatization of religion: Theorising religion, media and social change*, „Northern Lights”, 2008, nr 6(1), s. 12.

¹⁸⁴ S. Hjarvard, *The mediatization of religion...*, dz. cyt., s. 12.

¹⁸⁵ J. B. Thompson, *The Media and Modernity: A Social Theory of the Media*, Polity Press, Cambridge 1995; J. B. Thompson, *Ideology and Modern Culture*, Polity Press, Cambridge 1990.

narodziny nowej technologii, która umożliwiła obieg informacji w społeczeństwie. To wydarzenie zinstytucjonalizowało środki masowego przekazu takie, jak książki, czasopi-
sma i gazety, które zaczęły pełnić kluczową w społeczeństwie i umożliwiły interakcję
i komunikację pomijając odległość. Dodatkowo umożliwiło przechowywanie i groma-
dzenie informacji. Co zatem idzie, środki masowego przekazu pomogły przekształcić
społeczeństwo agrarne i feudalne, a mediatyzacja społeczeństwa – tworzyć nowoczesne
instytucje, takie jak państwo, sfera publiczna i nauka. Późniejszy rozwój kolejnych me-
diów, takich jak radio, telewizja i Internet, spowodował dalsze konsekwencje. Komuni-
kacja, która w przeszłości związana była ze spotkaniem osób twarzą w twarz (fizyczne
spotkanie) została zastąpiona komunikacją pośrednią, gdzie relacja między nadawcą
a odbiorcą ulega zmianie w kluczowych aspektach. W przypadku komunikacji masowej
nadawca stara się sprawować kontrolę nad treścią przekazu, mając mały wpływ na to, jak
odbiorca z tego przekazu skorzysta. W przypadku użycia mediów nadawca i odbiorca
wpływają na sposób i treść komunikacji. Thompson zauważył również związek między
mediatyzacją a jej konsekwencjami kulturowymi oraz pojawieniem się dużych global-
nych organizacji medialnych.¹⁸⁶

W. Schulz¹⁸⁷ oraz F. Krotz również użyli pojęcia mediatyzacji do określenia roli me-
diów w szeroko rozumianych zmianach społecznych¹⁸⁸.

W. Schulz zidentyfikował cztery rodzaje procesów, dzięki którym media zmieniają
ludzką komunikację i interakcję. Pierwszy proces to rozszerzanie ludzkich zdolności ko-
munikacyjnych zarówno w czasie, jak i w przestrzeni. Drugi to zastępowanie działań ko-
munikacyjnych, które wcześniej odbywały się fizycznie (na przykład bankowość inter-
netowa). Trzecim procesem jest inicjowanie i łączenie działań – komunikacja fizyczna
łączy się z komunikacją pośrednią, a media przenikają do życia codziennego. W czwar-
tym procesie podmioty z wielu różnych obszarów muszą zmienić swoje zachowanie, aby
dostosować się do mediów.¹⁸⁹

F. Krotz twierdził, że mediatyzacja jako metaprocess jest na równi z indywidualizacją
i globalizacją. Opisywał mediatyzację jako ciągły proces, w którym media zmieniają

¹⁸⁶ S. Hjarvard, *The Mediatization of Society...*, dz. cyt., s. 105-134.

¹⁸⁷ W. Schulz, *Reconstructing Mediatization* ...dz. cyt., s. 87-101.

¹⁸⁸ F. Krotz, *Mediatisierung: Fallstudien zum ...*, dz. cyt.

¹⁸⁹ W. Schulz, *Reconstructing Mediatization...*, dz. cyt, s. 87-101.

relacje i zachowania międzyludzkie, a tym samym zmieniają społeczeństwo i kulturę.¹⁹⁰ Teoria ta pokazuje perspektywę instytucjonalną mediów i ich interakcji z kulturą oraz społeczeństwem. Stosuje zestaw pojęć socjologicznych i określa elementy składające się na „logikę medialną” (*ang. media logic*), aby lepiej przeanalizować wzajemne oddziaływanie mediów i innych sfer społecznych (czyli instytucji). Pojęcie mediatyzacji odnosi się do sytuacji historycznych, w których media uzyskały autonomię jako instytucja społeczna i są wplecione w funkcjonowanie innych instytucji.¹⁹¹

Mediatyzacja według S. Hjarvarda to coś więcej niż tylko etykieta zespołu zjawisk świadczących o zwiększonym wpływie mediów i powinna odnosić się także do innych teorii socjologicznych.¹⁹² Postrzega on mediatyzację społeczeństwa jako proces, w którym społeczeństwo coraz bardziej jest podporządkowane mediom. Proces ten charakteryzuje dwoistość: media zintegrowane są z działalnością innych instytucji społecznych, jednocześnie uzyskując status instytucji społecznych. Powoduje to interakcje społeczne w ramach poszczególnych instytucji, między nimi oraz w całym społeczeństwie, odbywające się za pośrednictwem mediów.

Termin *logika mediów* odnosi się do instytucjonalnego i technologicznego sposobu działania mediów. Stopień podporządkowania się mediom jest zależny od instytucji i sektorów społeczeństwa. Mediatyzacja nie jest taka sama w każdym zakątku świata. Jej rozwój przyspieszył szczególnie w końcówce XX wieku, zwłaszcza w wysoko uprzemysłowionych społeczeństwach, tj. w Europie, USA, Japonii. W miarę postępu globalizacji coraz więcej regionów i kultur będzie podlegać mediatyzacji w różnym stopniu.¹⁹³ Mediatyzacja odnosi się do wielu dziedzin, ale odnosi się również do codziennego życia. Zmediatyzowane mogą zostać także codzienne czynności użytkowników mediów – możemy mieć takie odczucie, że podczas obecności w Sieci przyjaciele naszych przyjaciół mogą stać naszymi przyjaciółmi obserwując ich wspólne dyskusje i rozmowy na forum publicznym lub w mediach społecznościowych¹⁹⁴.

¹⁹⁰ S. Hjarvard, *The Mediatization of Society...*, dz. cyt., s. 105-134.

¹⁹¹ Tamże s. 110.

¹⁹² Tamże s. 113.

¹⁹³ Tamże s. 113.

¹⁹⁴ A. Adamski, *Media as...*, dz. cyt., s. 30.

1.3. Modele komunikacji

Podjęcie systemowe rozwinęło się w nauce o komunikowaniu w latach 50-60. XX wieku i w tym okresie zdefiniowanych zostało kilka modeli komunikacyjnych.

1.3.1. Model aktu perswazyjnego Harolda Lasswella (1948).

H. Lasswell uznawany jest za jednego z ojców nauki o komunikowaniu. W swoim linearnym modelu komunikowania zwanym modelem aktu perswazyjnego zawartego w pracy *The communication of Ideas* odwołuje się do teorii sukcesu propagandowego na podstawie analizy kampanii w USA. Stworzony po II wojnie światowej, reprezentuje uproszczoną koncepcję komunikacji, z jednym nadawcą, jedną wiadomością i jednym lub większą liczbą odbiorców, ale bez interakcji.¹⁹⁵

Koncepcja tego modelu zakłada bierność odbiorcy – jest izolowany od innych wpływów. Model zakłada jednokierunkowość – zakładając osiągnięcie określonego skutku. Model ten jest uznawany jako wzorzec i doczekał się licznych rozszerzeń.¹⁹⁶ Ta prosta formuła została wykorzystana na kilka sposobów, głównie do organizowania i rozpoczynania dyskusji o komunikacji¹⁹⁷.

Lasswell zdefiniował funkcje mediów masowych w postaci trzech punktów: (I) *utrwalenie otoczenia społecznego w którym wyrażany jest system wartości*, (II) *podtrzymywanie relacji członków społeczeństwa i wytwarzanie reakcji na otoczenie*, (III) *przekazywanie dziedzictwa społecznego*.¹⁹⁸

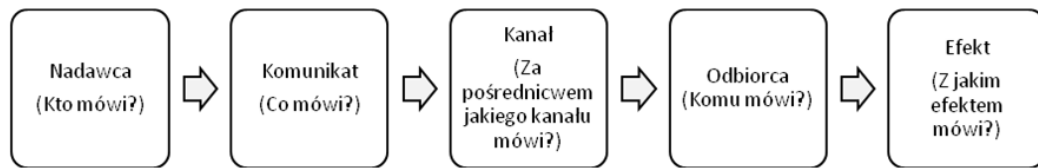
¹⁹⁵ Z. S. Sapienza, N. Iyer, A. S. Veenstra, *Reading Lasswell's Model of Communication*, „Mass Communication and Society”, 2015, nr 18(5), s. 608.

¹⁹⁶ B. Dobek-Ostrowska, *Komunikowanie polityczne...* dz. cyt., s. 31-33.

¹⁹⁷ J. W. Riley, M.W. Riley, *Mass Communication and the Social System*, [w:] R. K. Merton, L. Broom, S. Cotrell (red.), *Sociology Today*, Basic Books, New York 1959.

¹⁹⁸ Lasswell's Model, <https://www.communicationtheory.org/lasswells-model>, 18.05.2022.

Rysunek 6. Model aktu perswazyjnego Harolda Lasswella (1953).



Źródło: Opracowanie własne na podstawie: H. D. Lasswell, *The structure and function of communication in society*, [w:] L. Brynson (red.), *The Communication of Ideas*, Harper & Brothers, New York 1948.

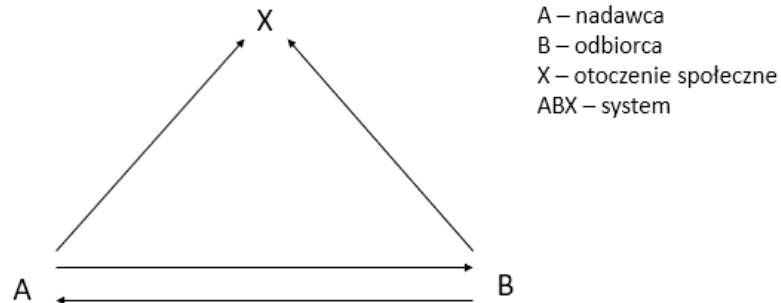
1.3.2. Model systemowy Newcomba (1953).

Model systemowy Theodore Newcomba to pierwszy model, który wprowadził nową kategorię uczestnika komunikowania. A i B działają indywidualnie – są nadawcą i odbiorcą komunikatu. X jest częścią otoczenia społecznego. Głównym celem tej teorii jest wprowadzenie roli komunikacji w relacji społecznej (społeczeństwa) oraz utrzymanie równowagi społecznej w systemie społecznym. Razem tworzą system, wchodząc w interakcje i zależności. Jeden element systemu ma wpływ na inny element systemu¹⁹⁹. Podstawowym założeniem modelu jest dążenie do spójności postaw i związku inicjującego komunikację, gdy warunki na to pozwalają²⁰⁰.

¹⁹⁹ Newcomb's model, <https://www.communicationtheory.org/the-newcomb%E2%80%99s-model/>, 20.05.2022.

²⁰⁰ D. McQuail, S. Windahl, *Communication Models for the Study of Mass Communications*, Routledge, 1993, s. 28.

Rysunek 7. Model systemowy Newcomba (1953).

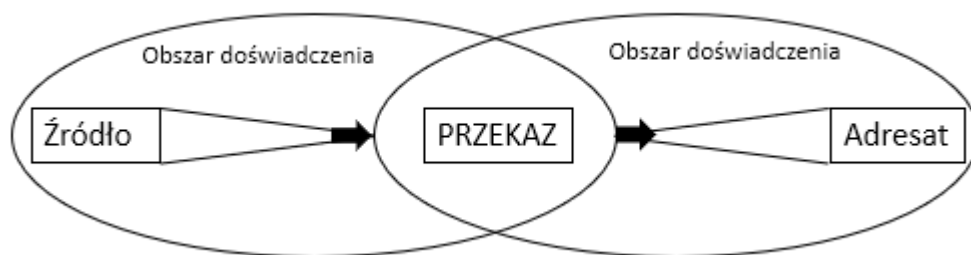


Źródło: Opracowanie własne na podstawie: Newcomb's model, <https://www.communicationtheory.org/the-newcomb%E2%80%99s-model/>, 20.05.2022.

1.3.3. Model wspólnoty doświadczeń Schramma (1954)

Komunikowanie według tego modelu oznacza dzielenie się jednostki doświadczeniami oraz uczestnictwo we wspólnotach z innymi. Źródło może stanowić osoba lub organizacja. Przekaz może przyjmować różne formy pisane (lub formę obrazu). Adresat to jednostka, publiczność masowego przekazu. W całym procesie najważniejszym punktem jest odpowiednia synchronizacja źródła, a wspólnota kodów, wiedzy i postaw definiuje skuteczność komunikowania.²⁰¹

Rysunek 8. Model wspólnoty doświadczeń Schramma (1954).



Źródło: Opracowanie własne na podstawie: W. Schramm, *How Communication Works*, [w:] W. Schramm (red.), *The Process and Effects of Mass Communication*, University of Illinois Press, Urbana 1954.

²⁰¹ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 95-97.

1.3.4. Modele komunikowania Schramma (1954)

Schramm dostosował model wspólnoty do poziomu interpersonalnego i masowego.

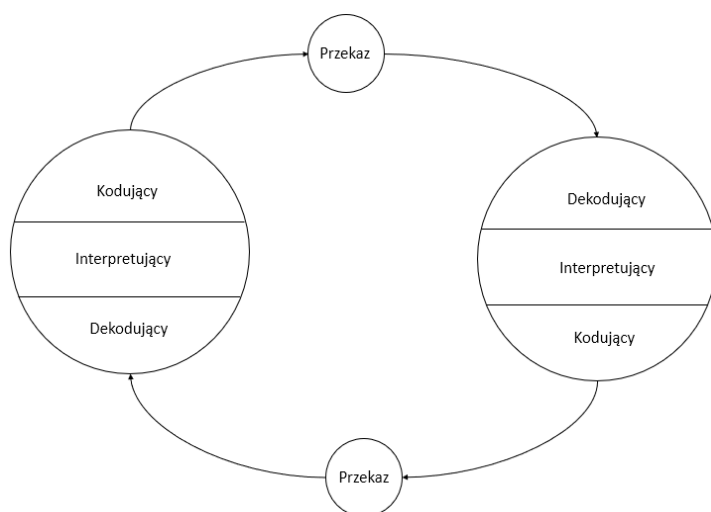
W modelu interpersonalnym każdy z ludzi jest zarówno źródłem i adresatem komunikatu. Człowiek koduje -> wysyła -> odbiera -> dekoduje -> interpretuje -> koduje itd. Pojawienie się tego podejścia oznaczało wyraźne zerwanie z tzw. tradycyjnym liniowym, jednokierunkowym kierunkiem komunikacji.²⁰²

W modelu komunikowania masowego komunikator masowy dekoduje komunikat, interpretuje i koduje. Informacje są bezużyteczne, dopóki nie zostaną starannie ujęte w słowa i przekazane innym. Kodowanie odgrywa bardzo ważną rolę, ponieważ inicjuje proces komunikacji poprzez zamianę myśli w treść. Kiedy informacja dociera do odbiorcy, jego głównym obowiązkiem jest zrozumienie, co mówca zamierza przekazać. Dopóki druga strona nie będzie w stanie zrozumieć lub rozszyfrować informacji, którą chce przekazać nadawca, wiadomość staje się bezużyteczna.²⁰³

²⁰² D. McQuail, S. Windahl, *Communication Models ...*, dz. cyt., s. 20.

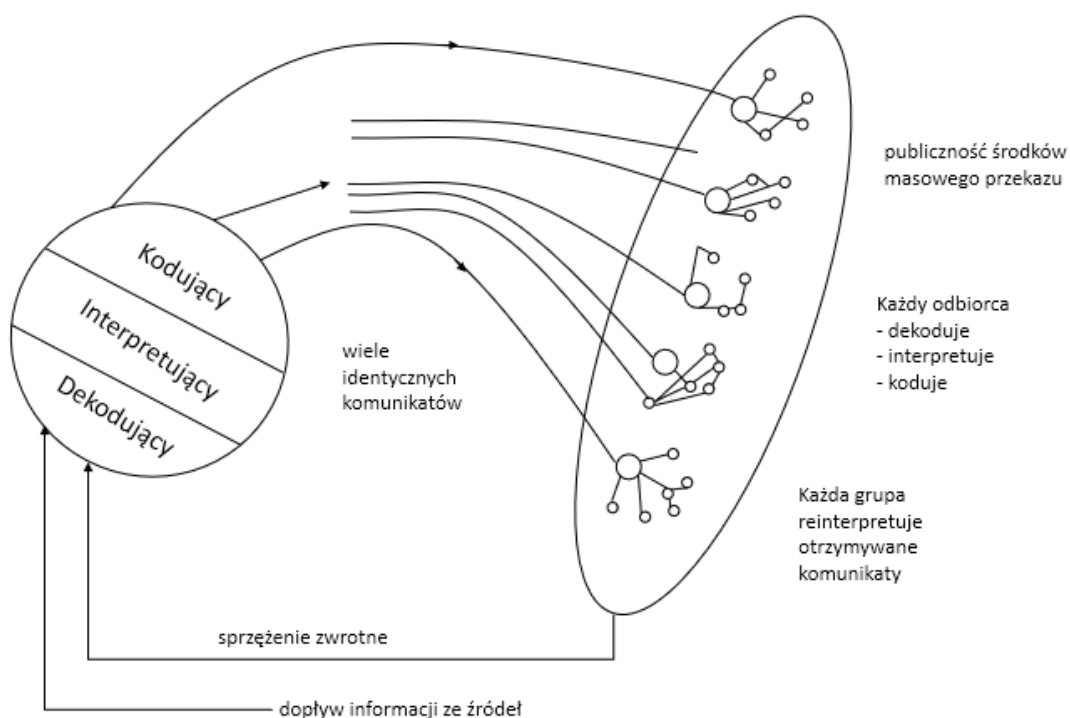
²⁰³ Schramm Model, <https://www.managementstudyguide.com/schramm-model-of-communication.htm>, 21.05.2022.

Rysunek 9. Model komunikowania interpersonalnego Schramma.



Źródło: Opracowanie własne na podstawie: W. Schramm, *How Communication Works*, [w:] W. Schramm (red.), *The Process and Effects of Mass Communication*, University of Illinois Press, Urbana 1954.

Rysunek 10. Model komunikowania masowego Schramma.



Źródło: Opracowanie własne na podstawie: W. Schramm, *How Communication Works*, [w:] W. Schramm (red.), *The Process and Effects of Mass Communication*, University of Illinois Press, Urbana 1954.

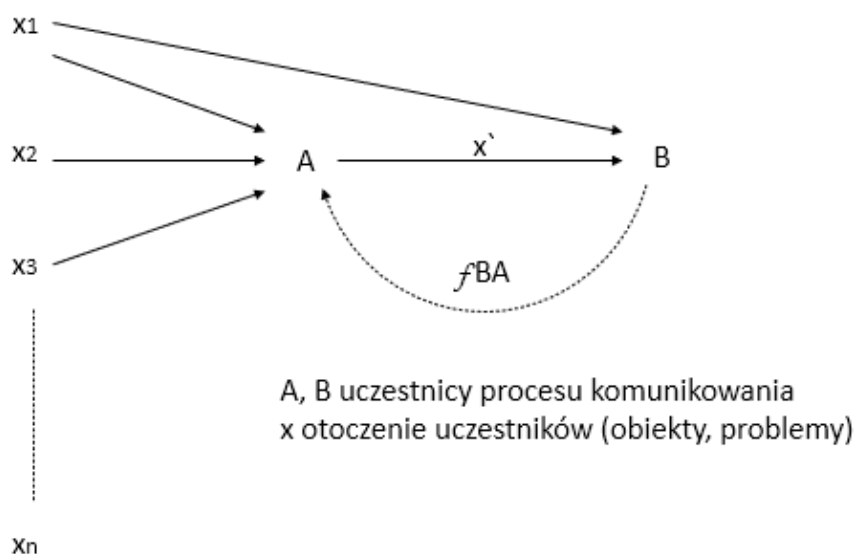
1.3.5. Model selekcji Bruce Westleya i Malcolma McLeana (1957)

B. Westley oraz M. McLean na podstawie modelu Newcomba stworzyli dwa modele procesu komunikowania, w dwóch podejściach: interpersonalnym i masowym.

W podejściu procesu komunikowania interpersonalnego obiekty w otoczeniu oznaczono ($x_1, x_2 \dots x_n$) zaś ich odpowiedniki (x', x'', x''' itd.). Obiekty (tematy, przedmioty itd.) podlegają procesowi selekcji przez nadawcę (A) i są przekazywane w formie (x') do odbiorcy (B).²⁰⁴

Model masowy Westleya McLeana powiązany jest z modelem Newcomba, który opiera się na fakcie, że komunikacja jest generowana przez brak równowagi systemu społecznego: model Westley Macleana ma czwarty element oprócz modelu Newcomba – C pełni rolę kanału, ale działając jednocześnie jako komunikujący *agent opinii publicznej*.²⁰⁵

Rysunek 11. Model komunikowania interpersonalnego Westleya i McLeana.



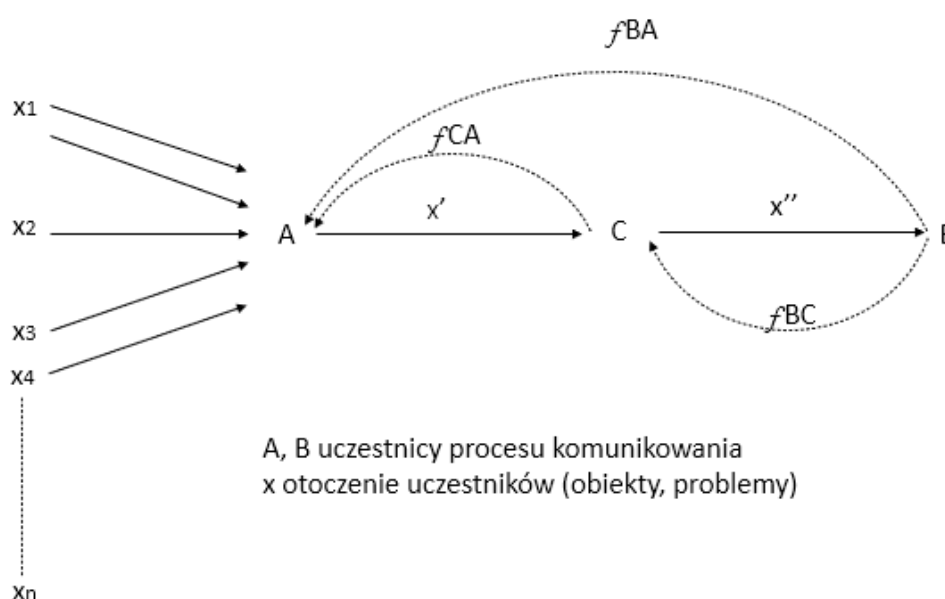
Źródło: Opracowanie własne na podstawie: B.H. Westley, M. MacLean, *A Conceptual Model for Mass Communication Research*, "Journalism Communication", 1957, nr 34, s. 31-38.

²⁰⁴ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 98-99.

²⁰⁵ M. Bucur, A. Ban, *Study on the Westley Maclean model and the co-orientation model and their exemplification in an industrial enterprise*, „Acta Marisiensis Seria Technologica”, 2019, nr 16(2), s. 26-27.

W podejściu procesu komunikowania masowego następują różnice w stosunku do komunikowania interpersonalnego: (I) sprzężenie zwrotne w komunikowaniu masowych prawie nie występuje, (II) liczba nadawców w komunikowaniu masowym jest znacząco większa²⁰⁶. Pośrednik (C) pełni rolę masowego komunikatora – decyduje on o treści komunikatu. Pełni on rolę *gate-keepera* (kontroler przepływu informacji). A stanowi źródło informacji, a B to członek społeczeństwa²⁰⁷.

Rysunek 12. Model komunikowania masowego Westleya i McLeana.



Źródło: Opracowanie własne na podstawie: B.H. Westley, M. MacLean, *A Conceptual Model for Mass Communication Research*, "Journalism Communication", nr 34, s. 31-38.

1.3.6. Model socjologiczny Rileyów (1959)

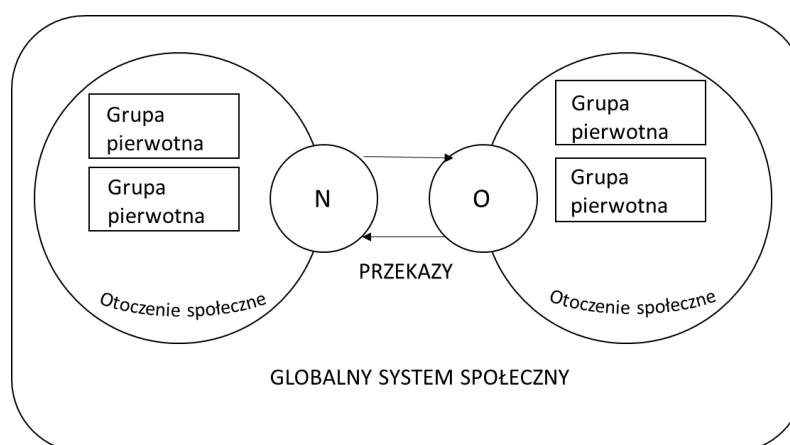
Model ten opiera się na tradycyjnych elementach: Nadawca (N), Przekaz (P), Odbiorca (O), uwytatniając aspekt społeczny i pokazując, jak ważną rolę pełni komunikowanie masowe jako proces społeczny. Grupy pierwotne to na przykład rodzina, zespół w pracy. Każda jednostka pozostaje w obrębie wpływu grupy²⁰⁸.

²⁰⁶ D. McQuail, S. Windahl, *Communication Models...*, dz. cyt., s. 38.

²⁰⁷ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 98-99.

²⁰⁸ Riley & Riley Model of Communication, <https://www.communicationtheory.org/riley-riley-model-of-communication>, 20.05.2022.

Rysunek 13. Model socjologiczny Rileyów (1959).



Źródło: Opracowanie własne na podstawie: J. W. Riley, M.W. Riley, *Mass Communication and the Social System*, [w:] R.K. Merton, L. Broom, S.Cotrell (red.), *Sociology Today*, Basic Books, New York 1959.

1.3.7. Model systemowy DeFleura (1966)

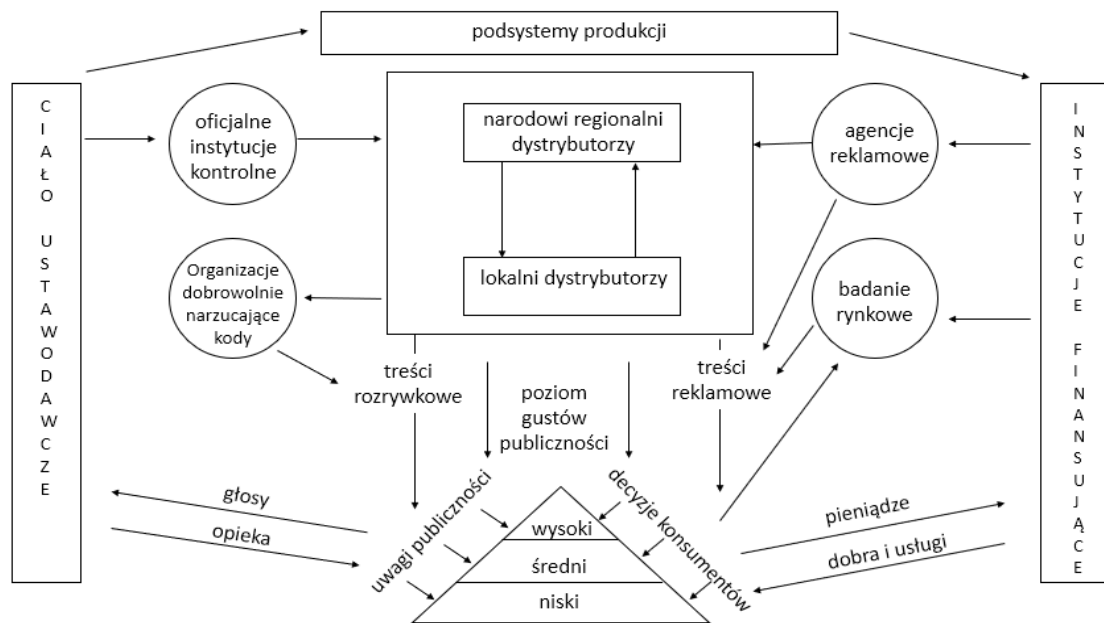
Jest to model, który stara się uwzględnić związki pomiędzy systemami politycznymi, medialnymi i ekonomicznymi. Ujmuje dodatkowo instytucje, które kontrolują i regulują media. W modelu tym sfera polityki ma swoją reprezentację przez agencje kontroli, powoływane przez ciało ustawodawcze (rząd, parlament, Senat). Sfera gospodarki ma reprezentację przez instytucje finansujące, w tym instytucje finansujące media. Podsystem produkcji obejmuje tworzenie przekazu medialnego i ogłaszanie publiczności. Publiczność występuje jako odbiorca środków komunikowania – obywatele, wyborcy oraz konsumenci.²⁰⁹

Innym ważnym modelem DeFleura jest model w którym jednym z ważnych aspektów w poprawie modelu Shannona i Weavera przez modelu komunikacji DeFleura (1970) jest dwukierunkowy proces komunikacji, który jest rekomendowany przez autora. Ten model jest pierwszym, który wprowadza zarówno dwukierunkową informację zwrotną, jak i grupę docelową w proces komunikacji.²¹⁰

²⁰⁹ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 93-103.

²¹⁰ M. DeFleur, *Theories of Mass Communication*, David McKay, New York 1970 ; De Fleur Model of Communication, <https://www.communicationtheory.org/de-fleur-model-of-communication>, 22.05.2022.

Rysunek 14. Model systemowy DeFleura (1966).



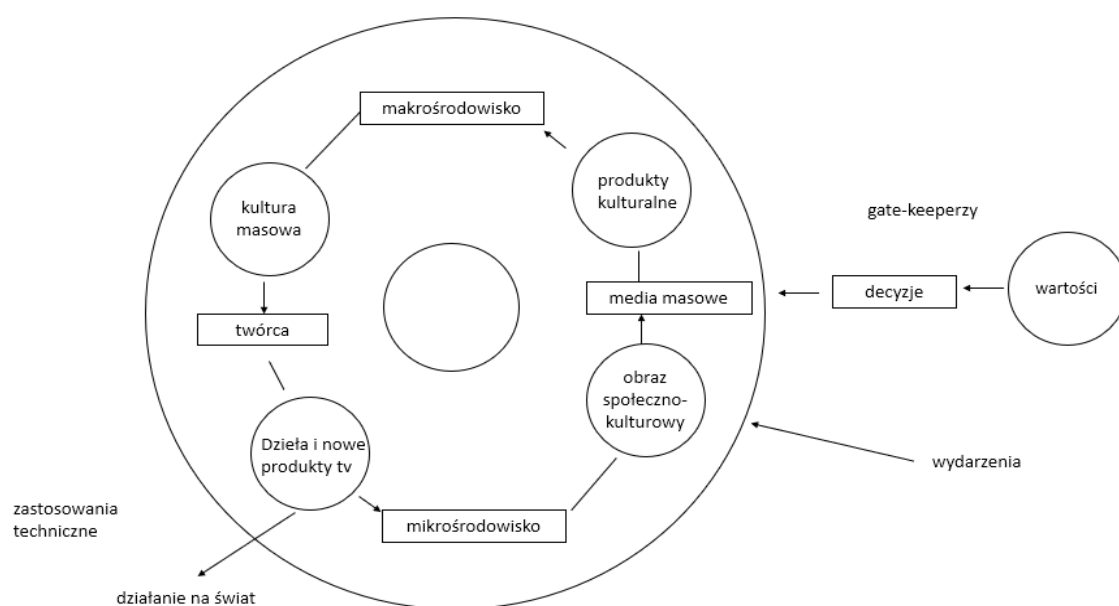
Źródło: Opracowanie własne na podstawie: M. DeFleur, *Theories of Mass Communication*, Longman, New York 1966; B. Dobek-Ostrowska, *Komunikowanie polityczne i publiczne*, PWN, Warszawa 2006, s. 93-103.

1.3.8. Model komunikowania Molesa (1967)

Model ujmuje środki przekazu przekształcane na tzw. kulturemy (*ang. culturemes*) – krążące w społeczeństwie nowe idee scalane i rozpowszechniane przez media. Moles ujmuje w swoim modelu kilka etapów – jest to obieg zamknięty. Początek jest związany z obecnością twórcy i dzięki jego inwencji powstają nowe idee (kulturemy). Następnie twórca przekazuje kulturemy do mikrośrodowiska. Kolejny wytworzony obraz społeczno-kulturowy rozpowszechniany jest przez masowe media, a następnie kulturemy te jako produkty trafiają do makrośrodowiska tworząc kulturę masową.²¹¹

²¹¹ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 93-103.

Rysunek 15. Model komunikowania Molesa (1967).



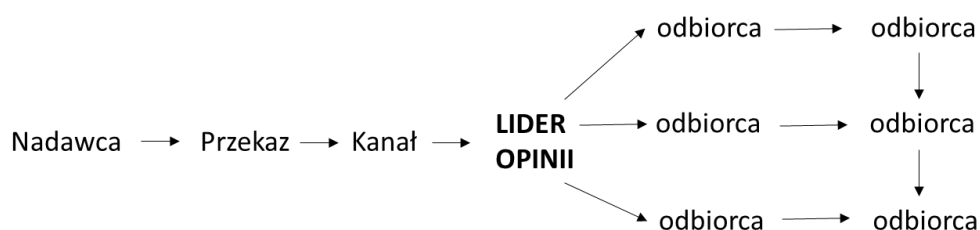
Źródło: Opracowanie własne na podstawie: A.A. Moles, *Sociodynamique de la culture*, Paris, La Haye: Mouton &C, 1967.

1.3.9. Model Lazarsfelda i Katza (1995)

Z punktu widzenia niniejszej pracy warto zwrócić na model Lazarsfelda i Katza. Paul Lazarsfeld wprowadził podejście sondażowe – koncentrował się na zastosowaniu metod statystycznych, ilościowych oraz jakościowych. W 1940 roku badacze prowadzili badania nad zależnością pomiędzy decyzjami wyborczymi, a wykształceniem, wiekiem, pozycją zawodową. W 1948 badacze bazując na wynikach badań ankietowych wskazali pozytywne korelacje pomiędzy opiniami obywateli i że efekt komunikowania jest powiązany z wzmocnieniem istniejących opinii. W swoich pracach zajmowali się pytaniami: w jaki sposób oddziaływać na postawy, aby je zmienić, jaki jest wpływ mediów na jednostki oraz w jaki sposób rozprzestrzeniają się komunikaty. Badacze stworzyli model dwuetapowego przepływu komunikowania, mający następujące charakterystyki: (I) proces komunikowania zależy od struktury powiązań pomiędzy odbiorcami, (II) ważną rolę pełnią interpersonalne kontakty odbiorców z mediów do społeczeństwa, (III) komunikowanie to złożony proces odbywający się dwuetapowo od mediów do liderów opinii i od liderów tych opinii w kierunku publiczności – rola tych liderów opinii jest kluczowa, (IV)

pośredni i opóźniony wpływ środków masowego przekazu na odbiorców.²¹² Według autorów *idee często wypływają z radia i podążają do liderów opinii i od nich do mniej aktywnych sekcje ludności*²¹³.

Rysunek 16. Model Lazarsfelda i Katza (1955).



Źródło: Opracowanie własne na podstawie: E. Katz, P. F. Lazarsfeld, *Personal Influence*, Free Press, Glencoe 1955.

1.4. System społeczny, polityczny i medialny jako środowisko dla systemu komunikowania politycznego

Komunikacja masowa może być uznana jako zespół zabiegów, poprzez które grupy specjalistów wykorzystują media do rozpowszechniania treści informacyjnej lub symbolicznej. Charakteryzuje się [ona] wykorzystaniem różnych technik (prasy, radia, kina, telewizji, itd.), zasięgiem obszaru transmisji (regionalnej, krajowej czy światowej), różnorodnością publiczności.²¹⁴

Termin system komunikowania ma dwa znaczenia. Pierwsze – system posiada charakterystykę instytucjonalności – instytucje takie, jak organy, jednostki i grupy są powiązane między sobą i termin ten to uwypatnia. Drugie znaczenie – system posiada charakterystykę funkcjonalności, postrzegany jest jako proces, uwypatnia działania i pokazuje interakcje. System komunikowania jest złożony, wielopłaszczyznowy, z nadawcami zespołowymi (np. formalnie ustrukturyzowane organizacje). Zamiast indywidualnego

²¹² B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 34-36.

²¹³ D. McQuail, S. Windahl, *Communication Models ...*, dz. cyt., s. 62.

²¹⁴ Ch. Baylon, X. Mignot, *Komunikacja*, Kraków 2008, s. 177.

odbiorcy mamy bardziej sformalizowaną grupę ludzi (zwykle pozbawionych interpersonalnego typu styczności).²¹⁵

O typie systemu według B. Dobek-Ostrowskiej decydują²¹⁶:

- uczestnicy systemu (struktura, skład, cechy);
- kontrola treści komunikatów;
- źródła dojścia informacji do systemu – zarówno do nadawców jak i odbiorców;
- typ kontaktów i rodzaj styczności pomiędzy członkami w systemie;
- zadania przeznaczone do wykonania przez członków systemu;
- wzorce, reguły, czy normy kierujące członkami systemu;
- forma sprzężenia zwrotnego.

Wraz z rozwojem mediów cyfrowych zmienia się charakter tej komunikacji. Media cyfrowe łączą funkcje komunikowania masowego i interpersonalnego oraz następuje prywatyzowanie sfery publicznej²¹⁷.

Największym systemem jest system komunikowania społecznego określany mianem ogółu społeczeństwa i relacji komunikacyjnych. Jednostka, która należy do tego systemu może również należeć do innych formalnych lub nieformalnych grup, które także mogą się między sobą komunikować. Człowiek będący w pracy jest członkiem systemu w swojej organizacji, jako obywatel odbiera komunikaty publiczne, jako wyborca odbiera komunikaty w obrębie systemu politycznego, a jako widz komunikuje się za pomocą systemu komunikowania masowego.²¹⁸ Komunikowanie może być traktowane jako składnik procesu społecznego: *jako akt komunikatywny jest środkiem przez który wyrażane są normy grupowe, sprawowana jest kontrola społeczna,*

²¹⁵ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 103.

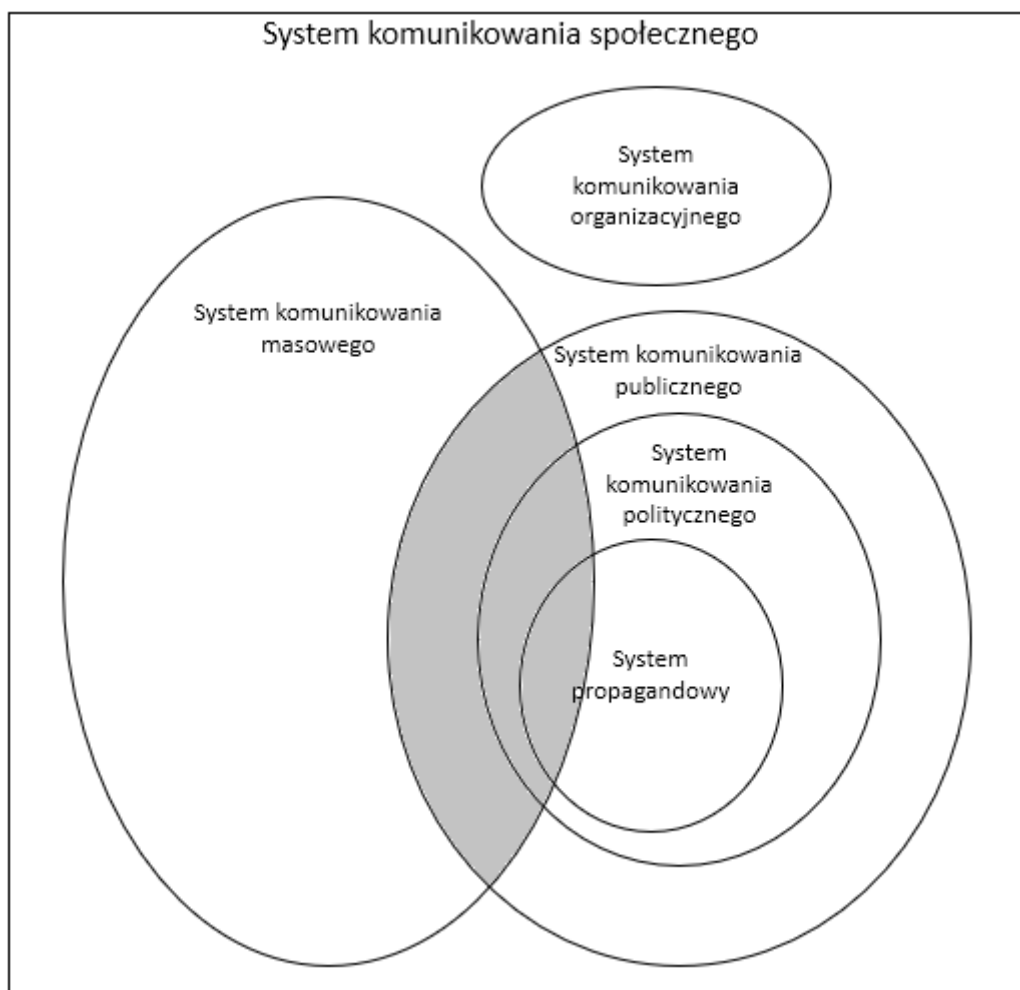
²¹⁶ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 104.

²¹⁷ B. Kudra, *O komunikacji społecznej*, [w:] *Komunikowanie publiczne. Zagadnienia wybrane*, B. Kudra, E. Olejniczak (red.), Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2014, s. 21.

²¹⁸ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 105.

przydzielane są role, osiągnięta koordynacja wysiłków, ujawniane są oczekiwania i przenoszony jest proces społeczny²¹⁹.

Rysunek 17. Relacje systemów komunikowania w systemie społecznym.



Źródło: Opracowanie własne na podstawie: B. Dobek-Ostrowska, *Komunikowanie polityczne i publiczne*, PWN, Warszawa 2006, s. 104-105.

System komunikowania organizacyjnego swoją rozpiętością obejmuje procesy komunikowania wewnątrz struktur organizacji. Przykładem takich systemów mogą być przedsiębiorstwa, związki zawodowe, partie polityczne, urzędy. W zależności od typów przekazu zakres oficjalności komunikacji, np. internetowej, elektronicznej, jest znacząco węższy od komunikacji rzeczywistej (realnej)²²⁰. W systemie tym mamy dwa poziomy komunikacji wewnętrznej: (I) Komunikowanie interpersonalne – osobiste kontakty

²¹⁹ M. Defleur, *Theories of ...* dz. cyt., s. 90.

²²⁰ J. Grzenia, *Komunikacja językowa...*, dz. cyt., s. 64.

członków grupy, przekazywanie informacji następuje od menedżera w dół struktury, (II) Komunikowanie grupowe – jest obligatoryjnym elementem systemu komunikowania w organizacji formalnej, nieformalnej, małej, ogromnej w której wszyscy uczestnicy nie muszą znać się osobiście. Wyróżniamy 4 kierunki komunikowania się: kierunek w dół (przesyłanie poleceń od jednostki na szczycie w głąb struktury), kierunek w górę (*feedback* od podwładnych w kierunku menedżera), kierunek poziomy (komunikowanie uczestników będących na tym samym poziomie), kierunek ukośny (komunikowanie się jednostek różniących się pozycją w strukturze).²²¹

System komunikowania masowego zawiera wszystkie inne podgrupy komunikowania masowego, zasięgiem obejmując największą liczbę odbiorców komunikatów. Każdy z poziomów ma więcej przypadków stwarzających odrębne problemy badawcze.²²²

Rysunek 18. Piramida komunikacji.



Źródło: Opracowanie własne na podstawie: D. McQuail, *Teoria komunikowania masowego*, Wydawnictwo Naukowe PWN, Warszawa 2008, s. 36.

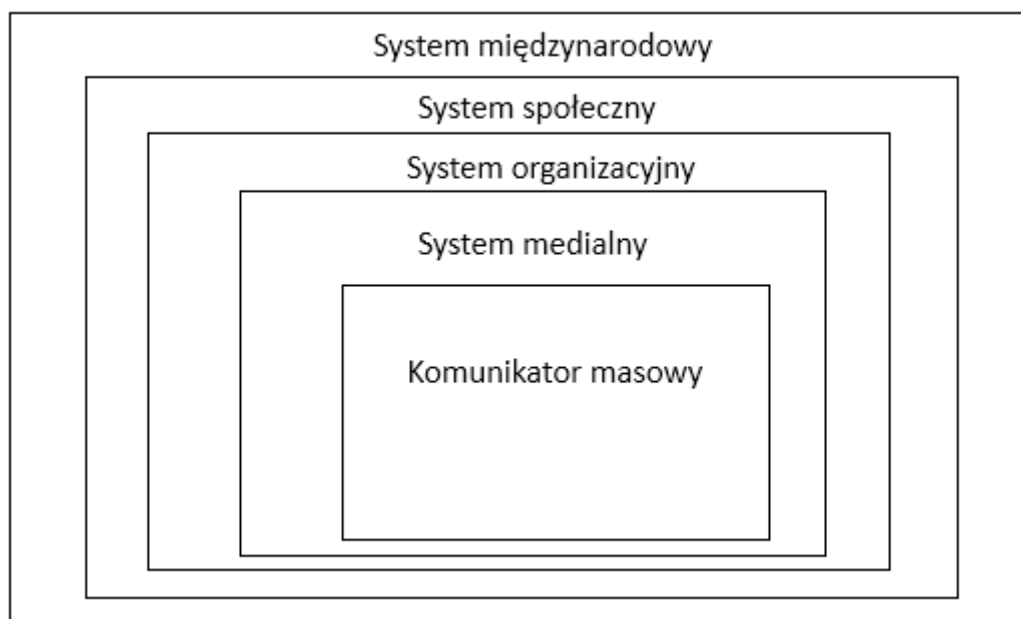
Profesjonalni nadawcy medialni, publiczność, urządzenia techniczne, kanały dystrybuujące przekazy medialne oraz inne są elementami systemu komunikowania masowego. Szeroka publiczność jest stymulatorem publicznego, społecznego rodzaju

²²¹ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 103-120.

²²² D. McQuail, *Teoria komunikowania...*, dz. cyt., s. 36.

komunikowania. System komunikowania masowego nie jest odizolowany, lecz wchodzi w interakcje z innymi systemami: systemem politycznym oraz systemem ekonomicznym. Parlament, rząd czy prezydent mogą powoływać instytucje kontrolujące media (w Polsce Krajowa Rada Radiofonii i Telewizji). System ekonomiczny również poprzez agencje reklamowe czy informacyjne, właściciele mediów może wywierać wpływ na elementy systemu komunikowania masowego, jakimi są m.in. media. W systemie tym osoby pracujące w środkach masowego przekazu związane są z komunikatorem masowym. Komunikator masowy jako nadawca medialny posiada cztery kategorie: (I) właściciele, którzy zarządzają mediami, (II) twórcy, (III) dziennikarze, (IV) technicy. Komunikator masowy wytwarza komunikaty – jest w samym centrum innych systemów.²²³ Mimo złożoności społeczeństwa sieciowego każdy z poziomów zawiera podobne problemy związane z teorią badań nad komunikacją: źródła i odbiorcy, funkcje i cele komunikacji, kanały komunikacji, treść, skutek komunikacji (idee, działania)²²⁴.

Rysunek 19. Organizacja mediów masowych.



Źródło: Opracowanie własne na podstawie: D. McQuail, Teoria komunikowania masowego, Wydawnictwo Naukowe PWN, Warszawa 2008, s. 283.

²²³ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 103-120.

²²⁴ D. McQuail, *Teoria komunikowania...*, dz. cyt., s. 38.

System medialny ulega zmianom wraz z rozwojem technologicznym. *System medialny jest to zespół mediów masowych działających w społeczeństwach narodowych, w różnych sektorach, jak prasa, radio, telewizja, nowe media.*²²⁵ Stanowi jądro systemu komunikowania masowego – do systemu tego należą: system prasowy (redakcje), system radiowy (stacje radiowe), system telewizyjny (kanały i sieci telewizyjne), system nowych mediów (portale internetowe), system telekomunikacyjny (operatorzy sieci satelitarnych, telefonicznych i kablowych)²²⁶. Relacje między nadawcami a otoczeniem są interaktywne: organizacje medialne utrzymują autonomiczne granice i wewnątrz nich działają²²⁷.

Rynek pierwotny systemu medialnego to przede wszystkim publiczność mediów (czytelnicy, słuchacze, oglądający telewidzowie, korzystający z nowych mediów) oraz rynek reklamy (agencje reklamowe, zleceniodawcy). Rynek wtórny to dostawcy produktów medialnych, dystrybutorzy, dostawcy informacji (agencje prasowe), dostawcy technologii oprogramowania, sprzętu oraz instytucje władzy, jak rząd i instytucje kontrolujące. Naturę mediów masowych określają 3 następujące cechy: media są przemysłem – utrzymują się na rynku, pełnią usługi publiczne – spełniają funkcję informacyjną i edukacyjną, są aktorem politycznym – są elementem systemu politycznego (szczególnie w okresie wyborczym).²²⁸

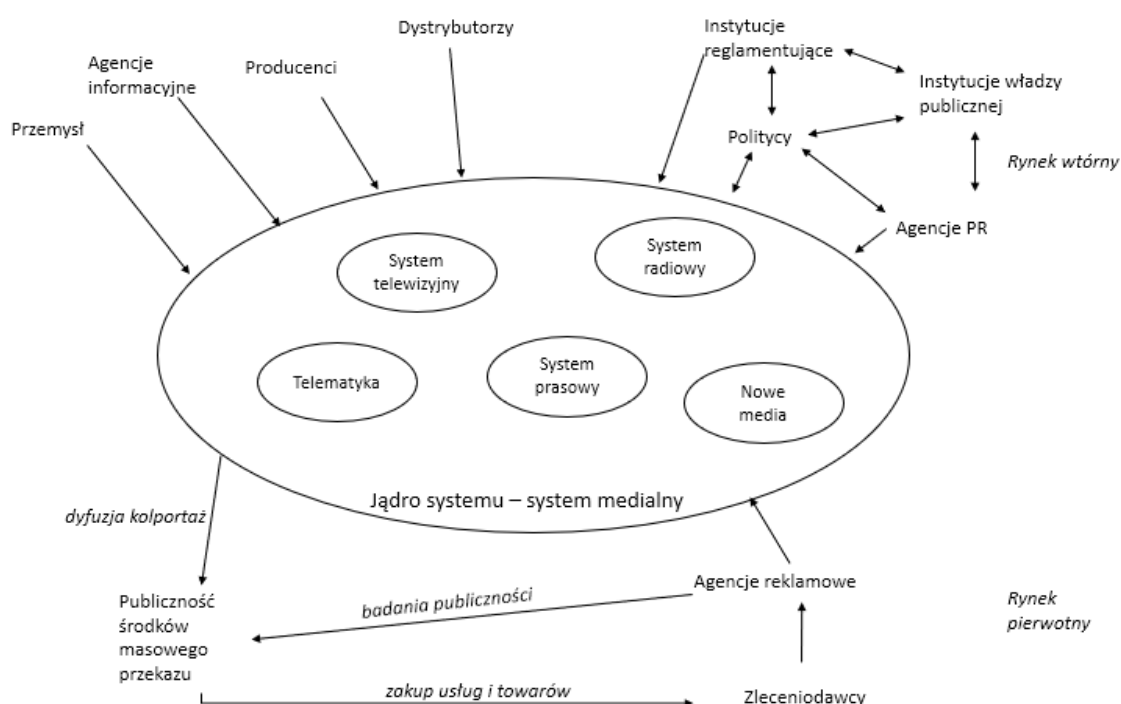
²²⁵ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 114.

²²⁶ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 103-120.

²²⁷ D. McQuail, *Teoria komunikowania...*, dz. cyt., s. 282.

²²⁸ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 103-120.

Rysunek 20. Struktura systemu komunikowania masowego.



Źródło: Opracowanie własne na podstawie: B. Dobek-Ostrowska, *Komunikowanie polityczne i publiczne*, PWN, Warszawa 2006, s. 116.

Komunikowanie systemu medialnego z systemem politycznym jest ściśle związane – informacja i komunikowanie są znaczące w wielu koncepcjach systemu politycznego. System polityczny korzysta z legitymizacji działania mediów masowych w postaci rozpowszechnianych treści, decyzji politycznych oraz potrzeb i oczekiwań społeczeństwa. Media masowe są traktowane jako kanały przez które informacja dociera do elementów systemu politycznego.²²⁹

D. McQuail wskazuje, że polityka jest jasnym przykładem adaptowania instytucji społecznej do rozwoju środków masowego przekazu, stąd media stały się głównym źródłem informacji oraz opinii politycznych. Powoduje to kilka wyzwań i aspektów dla polityki dla której media stanowią coraz ważniejszą pozycję: mniej czasu na uczestniczenie w polityce kosztem oglądania telewizji; marketing polityczny i spadek zaufania

²²⁹ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 119.

polityków w społeczeństwie; negatywne, coraz droższe i zbiurokratyzowane kampanie polityczne.²³⁰

1.5. Model systemu komunikowania politycznego i jego uczestnicy

Zdaniem B. Dobek-Ostrowskiej, *System komunikowania politycznego to efekt interakcji systemów: politycznego, medialnego i społecznego. Jest to podzbiór zawierający się w tych systemach, a nie zbiór równoważny w stosunku do nich. Uczestnicy systemu wchodzą w interakcje, podejmują konkretne działania, obejmujące przygotowanie, wysyłanie i odbieranie komunikatów o polityce, projektowanie i realizację strategii mających na celu zaspokojenie własnych politycznych interesów i potrzeb.*²³¹

System komunikowania politycznego tworzą trzy inne systemy:

- (I) system polityczny, który tworzą partie polityczne i inne organizacje, władza publiczna, prezydent, rząd itd.;
- (II) system medialny tworzony przede wszystkim przez organizacje medialne wraz z personelem mediów, stacje telewizyjne i radiowe, portale internetowe, firmy fonograficzne, agencje reklamowe;
- (III) system społeczny czyli całe społeczeństwo jako odbiorcy przekazów medialnych, wyborcy, obywatele, organizacje, które działają w sferze publicznej.

W. Schulz łączy wymianę komunikatów politycznych z aktorami, a mianowicie *wszystkimi grupami, organizacjami i jednostkami, które uczestniczą w procesie zbiorowo wiążącego podejmowania decyzji w sprawie dystrybucji ograniczonych zasobów w społeczeństwie*²³². W najogólniejszym znaczeniu komunikacja polityczna obejmuje interaktywne procesy informacji oraz formalne i nieformalne (podejmowanie decyzji za kulisami i negocjacje) sposoby przepływu wiadomości²³³.

²³⁰ D. McQuail, *Teoria komunikowania...*, dz. cyt., s. 514.

²³¹ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 119.

²³² W. Schulz, *Political communication*, [w:] W. Donsbach (red.), *The international encyclopedia of communication* (pp. 3671 -3682), MA: Blackwell, Cambridge 2008, s. 367.

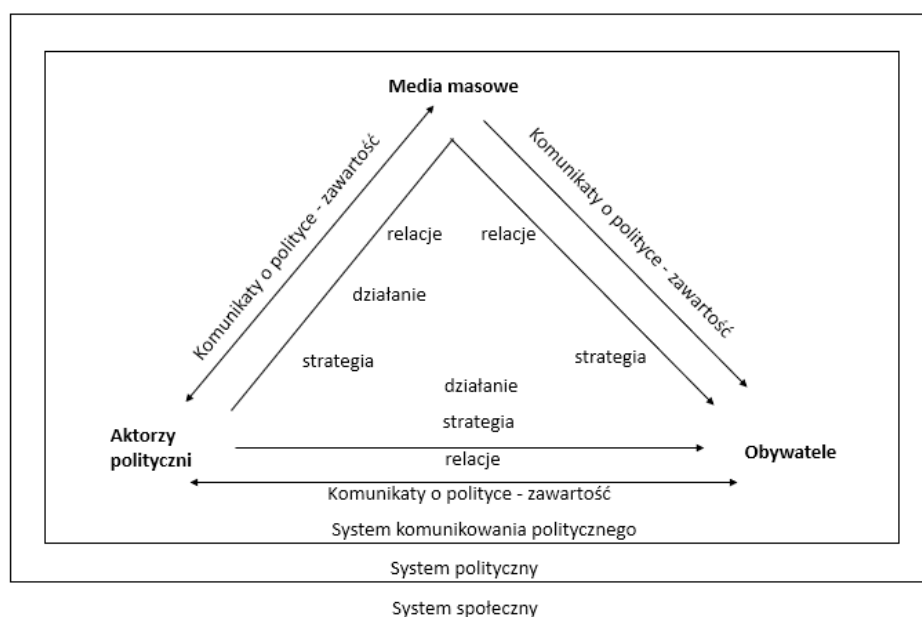
²³³ B. Pfetsch, *Comparing Political Communication*, ACADEMIA, 2012, s. 25-26.

Negocjacje, badania nad komunikacją są zainteresowane przede wszystkim publicznym sposobem komunikacji politycznej, który jest nieuchronnie związany z mediami masowymi.

W systemie politycznym wytwarzane są relacje pomiędzy uczestnikami: aktorzy polityczni a media, media a obywatele, aktorzy polityczni a obywatele. Relacje te mają znaczący wpływ na całkowity efekt komunikacji, mimo że wszystkie te elementy są jednakowo równoważne to spośród trzech elementów najliczniejsza grupa obywateli są najbardziej zaniedbywani.²³⁴

System komunikowania politycznego jest wrażliwy na procesy zachodzące we wszystkich systemach (politycznym, społecznym, medialnym). Głównym zadaniem systemu komunikowania politycznego jest wymiana informacji w polityce. W procesie komunikowania politycznego widoczne jest także działanie komunikacyjne i polityczne pomiędzy uczestnikami systemu, jak również sprzężenie zwrotne wytwarzające określone relacje między aktorami politycznymi, mediami masowymi a obywatelami.²³⁵

Rysunek 21. Model systemu komunikowania politycznego.



Źródło: Opracowanie własne na podstawie: B. Dobek-Ostrowska, *Media masowe i aktorzy polityczni w świetle studiów nad komunikowaniem politycznym*, Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław 2004, s. 61.

²³⁴ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 120.

²³⁵ Tamże.

Z uwagi na silne powiązanie z systemem publicznym i dyfuzję pojęć w tych dwóch systemach warto zwrócić uwagę na określenie ram i różnic. B. Dobek-Ostrowska pisze, że *Fundamentalnym elementem systemu [publicznego] są instytucje o charakterze publicznym, których działania dotyczą ogółu ludzi i są dostępne dla wszystkich jednostek. Cele instytucji, sposób powoływania i kompetencje są określone przez istniejące normy prawne. Instytucje te spełniają funkcję regulacyjną, zabezpieczającą i służebną zarówno wobec pojedynczego obywatela jak i całego społeczeństwa.*²³⁶

Komunikowanie publiczne dotyczy przede wszystkim komunikowania formalnego, którego celem jest przekazywanie informacji oraz wymiana informacji w formie publicznej, podtrzymująca przy tym więzi społeczne. System komunikowania publicznego swoim zasięgiem obejmuje organizacje społeczne, instytucje władzy państwa i samorządów, publiczności i szeregu powiązań między nimi.²³⁷

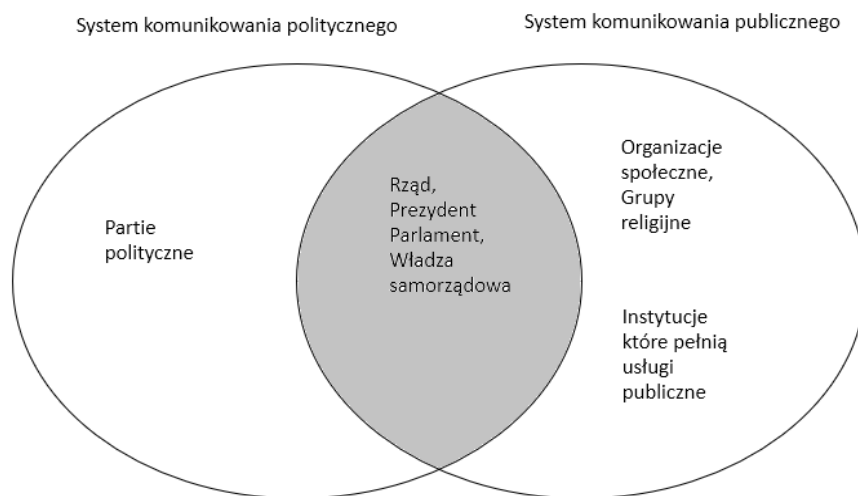
Komunikowanie publiczne ma cechy bardziej sformalizowane (w porównaniu do komunikowania politycznego), dotyczy sfery publicznej i ograniczone jest do terytorium regulowanego przez prawo i przepisy. Związane jest z prawem, wprowadzaniem procedur lub podejmowaniem decyzji publicznych. Zawiera m.in. obowiązek informowania społeczeństwa o sprawach ważnych poprzez informowanie wiadomości w serwisach informacyjnych do tego przeznaczonych, kampaniach reklamowych, zawiera także obowiązek prowadzenia dialogu (słuchania problemów obywateli) między nadawcami publicznymi i ich odbiorcami. Komunikowanie publiczne zawiera również zagwarantowanie obywatelom swobodnego dostępu do dokumentów i akt stworzonych przez administrację publiczną. Podsumowując, komunikowanie publiczne zbliża się w niektórych kwestiach do komunikowania politycznego czy organizacyjnego. Procesy te czasami zachodzą na siebie swymi zakresami.²³⁸

²³⁶ P. Zémor, *La communication*, Paris: PUF, 1995, s. 5.

²³⁷ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 123.

²³⁸ B. Kudra, *O komunikacji...*, dz. cyt., s. 20-22.

Rysunek 22. Instytucje publiczne jako aktor komunikowania publicznego i politycznego.



Źródło: Opracowanie własne na podstawie: B. Dobek-Ostrowska, *Komunikowanie polityczne i publiczne*, PWN, Warszawa 2006, s. 122.

Zasługującym na uwagę elementem systemu są instytucje władzy publicznej mające dwojaką przynależność – zarówno komunikowania publicznego, jak i systemu komunikowania politycznego. Do instytucji pełniących usługi publiczne zaliczamy m.in.: straż pożarną, policję, instytucje sanitarne, urzędy celne, urzędy skarbowe itd. Instytucje pełniące usługi publiczne nie powinny być elementami procesu komunikowania politycznego. Różnicą między komunikowaniem politycznym a publicznym jest to, że rolą komunikowania politycznego jest formułowanie ofert politycznych w celu przekonania wyborców, zaś komunikowania publicznego przekazywanie treści pozwalających obywatelom samodzielnie oceniać przesyłane komunikaty.²³⁹

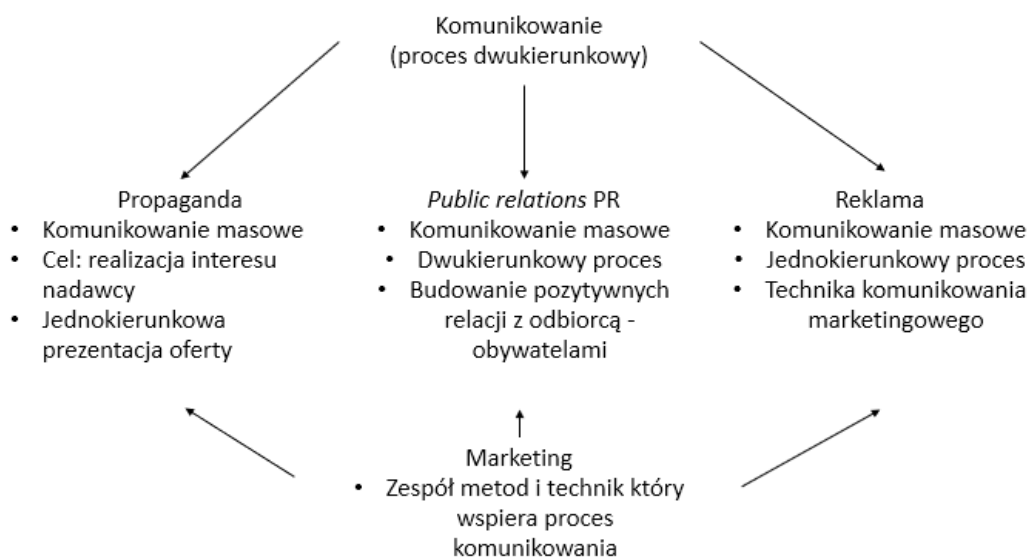
Komunikowanie polityczne to proces wielowektorowy; ma w sobie elementy propagandy, reklamy, elementów public relations i mechanizmy tzw. marketingu politycznego. W systemach demokratycznych i otwartych media ułatwiają przepływ komunikatów dwukierunkowo, lecz bardzo często jest to komunikowanie jednokierunkowe. Kanałem transmisyjnym dla przepływu informacji są media. Zazwyczaj komunikowanie polityczne przybiera formę wystąpień publicznych jak konferencje prasowe przemówienia, czy wywiady.²⁴⁰

²³⁹ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 122-124.

²⁴⁰ B. Kudra, *O komunikacji...*, dz. cyt., s. 21.

Polityka jest to proces, w którym stanowi się prawo, podejmuje decyzje, a komunikowanie polityczne przedstawia obywatelom ten proces. Komunikowanie polityczne ma dwie perspektywy: perspektywę teoretyczną (jako dziedzinę badań naukowych) oraz praktyczną (traktowane jako propaganda polityczna i marketing polityczny). Komunikowanie polityczne ma swoje dziedziny pokrewne.²⁴¹

Rysunek 23. Komunikowanie polityczne i dziedziny pokrewne.



Źródło: Opracowanie własne na podstawie: B. Dobek-Ostrowska, *Komunikowanie polityczne i publiczne*, PWN, Warszawa 2006, s. 132.

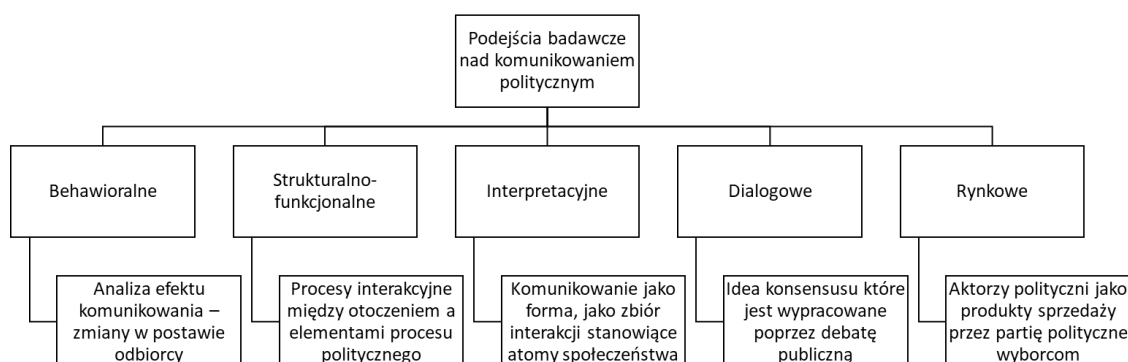
Heinz Eulau, Samuel J. Eldersveld, M. Janowitz stworzyli pierwszą definicję komunikowania politycznego w 1956 roku, w której przedstawili je jako pośredniczenie w relacjach między wyborcami a instytucjami politycznymi²⁴². Pierwsze próby wyklarowania pojęcia komunikowania politycznego ograniczały się do jednokierunkowego przekazu komunikatów politycznych od rządzących. Istnieje pięć różnych podejść do

²⁴¹ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 132-135.

²⁴² H. Eulau, S. J. Eldersveld, M. Janowitz, *Political Behavior: A Reader in Theory and Research*, „American Journal of Sociology”, nr 63(1), 1956.

komunikowania politycznego: behawioralne, strukturalno-funkcjonalne, interpretacyjne, dialogowe, rynkowe.²⁴³

Rysunek 24. Podejścia badawcze nad komunikowaniem politycznym.



Źródło: Opracowanie własne na podstawie: B. Dobek-Ostrowska, *Komunikowanie polityczne i publiczne*, PWN, Warszawa 2006, s. 135-142.

W latach 90. XX wieku powstało kilka definicji procesu komunikowania politycznego:

- Socjologiczne ujęcie Rusha (1992) nawiązuje do podejścia strukturalno-funkcjonalnego. Komunikowanie polityczne jest to transmisja polityczna i krążenie informacji pomiędzy systemem społecznym i politycznym²⁴⁴;
- L.W. Pye (1993) definiuje komunikację polityczną jako *przepływ wiadomości i informacji, który nadaje strukturę i znaczenie procesowi politycznemu*²⁴⁵.

²⁴³ B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 132-134.

²⁴⁴ M. Rush, *Politics and Society*, An Introduction to Political Sociology, Prentice Hall, London 1992.

²⁴⁵ L.W. Pye, *Political communication*, [w:] V. Bogdanor (red.), *The Blackwell encyclopedia of political science* (2nd edition, pp. 442-445), "Cambridge, MA: Blackwell", 2018, nr 68(2) s. 442.

- Medioznawcza perspektywa Negrine'a (1994) wskazuje na różnice w przekazie medialnym, różny stopień wpływu aktorów na zawartość mediów, zróżnicowany charakter odbiorców, proces tworzenia informacji²⁴⁶;
- Definicja systemowa Blumlera i Gurevitcha (1995) nawiązuje do teorii systemów, system składa się z elementów takich, jak polityczne instytucje, instytucje medialne, publiczność, aspekty kultury politycznej²⁴⁷;
- Koncepcja McNaira (1995) skupia się na inercyjności i celowości komunikowania o polityce²⁴⁸;
- Wizja Dentona i Woodwarda (1990, 1998) ujmuje komunikowanie polityczne jako debatę publiczną na tematy finansów publicznych, sankcji, kontroli i społecznego (implikacji społecznych)²⁴⁹;
- Typ strategicznej kultury komunikacji politycznej (1998) – według tej koncepcji proces komunikowania politycznego jest zdefiniowany przez grupy polityczne i udziałowców (*ang. stakeholders*) oraz powiązań między nimi – tworząc swojego rodzaju grę komunikowania politycznego²⁵⁰.
- Złoty trójkąt komunikowania Perloffa (1998) wskazuje proces wymiany oraz interpretację komunikatów pomiędzy liderami, obywatelami i mediami masowymi – układ ten autor nazywa złotym trójkątem komunikowania politycznego²⁵¹.

Należy zaznaczyć, że pojęcia wypracowane powinno traktować się komplementarnie a komunikowanie polityczne jest silnie związane ze sferą publiczną oraz demokracją.²⁵²

²⁴⁶ R. Negrine, *Politics and the Mass Media in Britain*, Routledge, London 1994; R. Negrine, *The Communication of Politics*, Sage, London 1996.

²⁴⁷ M. Gurevitch, J. Blumer, *Political Communication Systems and Democratic Values*, [w:] J. Lichtenberg (red.), *Democracy and the Mass Media*, Cambridge University Press, Cambridge 1990, s. 269-287.

²⁴⁸ B. McNair, *Introduction to Political Communication*, Routledge, London 1998.

²⁴⁹ R. Denton, G. Woodward, *Political Communication in America*, Longman, New York 1998.

²⁵⁰ J. B. Manheim, *The news shapers: Strategic communication as a third force in news making*, [w:] D. Graber, D. McQuail, & P. Norris (red.), *The politics of news: The news of politics* (pp. 94-109), Congressional Quarterly Press, Washington 1998.

²⁵¹ R. M. Perloff, *Political Communication, Politics, Press and Public in America*, LEA, London 1997.

²⁵² B. Dobek-Ostrowska, *Komunikowanie polityczne...*, dz. cyt., s. 142-160.

1.6. Język

Język jest narzędziem, które umożliwia komunikowanie się i wyrażanie własnych myśli. Nie jest to tylko zbiór słów, ale musi mieć pewne cechy. Biorąc pod uwagę, że używany jest w życiu codziennym w sposób bezrefleksyjny, często brakuje refleksji, jak ważną rolę pełni. Aby opisać czym jest język posłużono się definicją R. Grzegorzczukowej: *system znaków konwencjonalnych, fonicznych, służący do porozumiewania się o wszystkim (uniwersalny), charakteryzujący się tzw. dwuklasowością, tzn. udziałem gramatyki, umożliwiającą nieograniczone tworzenie nowych konstrukcji*²⁵³.

M. Bugajski dzieli komunikację na dominującą (jest to rodzaj komunikacji oficjalnej, powszechniejszej przez środki masowego przekazu; jednokierunkowym, zbiorowym i pośrednim kontakcie z uprzywilejowaną pozycją nadawcy komunikatu i wyszczególnioną perswazyjną funkcją języka) i komunikację alternatywną (komunikację wspólnot komunikatywnych, uczestniczącą, z aktywnymi, upodmiotowionymi twórczymi uczestnikami) oraz nieoficjalną i oficjalną.²⁵⁴

Komunikacja jako system nadawczo-odbiorczy, w którym nadawca ma za zadanie przekazać pewną informację odbiorcy musi obrać język jako sposób kodowania tych informacji. Zarówno nadawca wiadomości, jak i odbiorca powinni posługiwać się tym samym kodem, aby informacja mogła zostać przekazana w odpowiedni sposób. Nadawca wybiera słownictwo (wyrazy) i struktury gramatyczne, które powinny być znane również odbiorcy. W przypadku rozmowy twarzą w twarz lub poprzez telefon bądź inne urządzenie przesyłające dźwięk uruchamiany jest narząd mowy i w postaci fal akustycznych (w przypadku urządzeń dodatkowo przetwarzających sygnały) trafia do ucha odbiorcy przekazu. Odbiorca odczytuje wiadomość przy pomocy kodu i znanych struktur gramatycznych. Jest to sytuacja idealna, w praktyce bowiem mogą wystąpić liczne przypadki zakłóceń: w przypadku wykorzystania narządu mowy na przykład niedosłyszenie, w przypadku tekstu pisanego – niezrozumienie czyli niepełna zgodność kodu (odbiorca może mieć inne skojarzenia z wybranymi słowami jak nadawca, inne emocje kojarzą się

²⁵³ R. Grzegorzczukowa, *Wstęp do ...* dz. cyt., s. 13.

²⁵⁴ M. Bugajski, *Język w ...* dz. cyt., 2006.

z danym słownictwem). Takie sytuacje mogą powodować nieporozumienia i może zaistnieć potrzeba dalszych wyjaśnień, by nadać szerszy kontekst wypowiedzianym słowom.²⁵⁵

Tadeusz Milewski podaje definicję znaku jako *element rzeczywistości, który jest istotny nie ze względu na swoje własne cechy, ale ze względu na relację do innego elementu rzeczywistości, do którego się odnosi*²⁵⁶. Znaki w języku mają dwie charakterystyki: forma znaku (jest to zjawisko obserwowane zmysłami) i znaczenie (treść). Dzięki temu między innymi możemy wnioskować i szukać powiązań.

Według R. Grzegorzczukowej możemy wyróżnić siedem podstawowych cech języka²⁵⁷:

- *Konwencjonalność znaków*. Znaki konwencjonalne to takie, które opierają się o pewien wzorzec (konwencję) znany nadawcy i odbiorcy (tej znajomości muszą się nauczyć). Przykładem konwencji znakowej może być system znaków drogowych. W przypadku języka naturalnego nie jest możliwe wytworzenie takiego systemu, ponieważ język ciągle się zmienia, dlatego też autorka mówi o *uzusie semiotycznym*, który warunkuje użycie znaku;
- *Foniczność znaków językowych*. We współczesności odbieranie języka w postaci pisanej jest niemal równie powszechne, co odbierane foniczne. Ludzie, którzy mówią, dysponują dwiema formami: foniczną (ciąg dźwięków) i graficzną (ciąg liter);
- *Dwustopniowość znaków językowych*. Znaki językowe (które mają jakieś znaczenie) są zbudowane z mniejszych nie mających znaczenia elementów (diakrytów) – przykładowo diakryty *r*, *k*, *a* w języku polskim tworzy znaki (ciągi znaczące) jak *rak*, *kra*, *arka*;
- *Dwuklasowość języka*. Język posiada dwie klasy: *Słownik* (liczny, w danym momencie również stanowi zamkniętą klasę) oraz *gramatyki* (reguły, które pozwalają na tworzenie nowych wypowiedzi);

²⁵⁵ R. Grzegorzczukowa, *Wstęp do...*, dz. cyt., s.12-21.

²⁵⁶ T. Milewski, *Językoznawstwo*, Warszawa 1965.

²⁵⁷ R. Grzegorzczukowa, *Wstęp do...*, dz. cyt., s. 22-33.

- *Uniwersalność systemu językowego.* Językiem naturalnym możemy przekazywać informację o wszystkim. Cecha ta jest powiązana z cechą dwuklasowości języka;
- *Abstrakcyjność znaków językowych.* Umożliwia operowanie na pojęciach abstrakcyjnych – mówienie o zjawiskach przyszłych, przeszłych, odległych, a także tylko pomyślanych;
- *Polisemiczność znaków.* Umiejętność posługiwania się znakami do opisu zjawisk, które nie mają swojej nazwy, ale ze względu na relację i podobieństwo mogą zostać określone w metaforyczny sposób, na przykład: *dochodzić do domu* – ruch fizyczny, *dochodzić do rozwiązania jakiejś zagadki* – ruch mentalny.²⁵⁸

Język spełnia dwie funkcje jako zjawisko społeczne: (1) funkcja socjalizująca – jednocząca członków danej społeczności (narodu), (2) funkcja kulturotwórcza polegająca na przechowywaniu i gromadzeniu wiedzy²⁵⁹. Należy pamiętać że zakres oficjalności komunikacji drogą elektroniczną poprzez wykorzystanie Internetu jest znacznie węższy niż w komunikacji rzeczywistej²⁶⁰.

Język również jest powiązany ze sposobem myślenia. Papież Benedykt XVI ujmuje tę korelację w następujący sposób: *myśl i relacja powstają zawsze w sferze języka, pojmowanego oczywiście w szerokim sensie, nie tylko werbalnym. Język nie jest zwykłą otoczką koncepcji, zamienną i prowizoryczną, ale żywym kontekstem, w którym myśli, niepokoje i projekty ludzi rodzą się w świadomości i przekształcają się w gesty, symbole i słowa. Tak więc człowiek nie tylko «używa» języka, ale w pewnym sensie «mieszka» w nim²⁶¹. Benedykt XVI twierdzi również, że nowe języki powstające w komunikacji cyfrowej rozwijają umiejętności bardziej w sposób emocjonalny i intuicyjny niż analityczny²⁶². Innym ujęciem języka może być forma dialogu personalnego – rozmowa*

²⁵⁸ R. Grzegorczykowa, *Wstęp do...*, dz. cyt., s. 22-33.

²⁵⁹ R. Grzegorczykowa, *Wstęp do...*, dz. cyt., s. 55.

²⁶⁰ J. Grzenia, *Komunikacja językowa...*, dz. cyt., s. 64.

²⁶¹ Papież Benedykt XVII, *Przemówienie z okazji zgromadzenia plenarnego Papieskiej Rady ds. Środków Społecznego Przekazu*, 28.02.2011, https://opoka.org.pl/biblioteka/W/WP/benedykt_xvi/przemowienia/massmedia_28022011.html, 01.11.2022.

²⁶² Papież Benedykt XVII, *Przemówienie z okazji ...*, dz. cyt.

prowadzona między ludźmi²⁶³. Język cechuje rozwój, który jest sprzężony z rozwojem człowieka²⁶⁴. Stąd język sam w sobie powinien być pielęgnowany i rozwijany w różnych formach aktywności myślowej²⁶⁵.

Język daje silne przesłanki ku temu, że jest on nieodłączną częścią ludzkich myśli. To co mówimy, w jaki sposób mówimy, jakiego słownictwa używamy to powstaje w głowie i wychodzi na zewnątrz. Człowiek jako istota społeczna należy do pewnej grupy ludzi, którzy przejawiają podobne cechy m. in. języka, dzięki któremu mogą się komunikować. System językowy jest niezbędnym narzędziem służącym do komunikacji między ludźmi oraz jedynym narzędziem do interpretacji otaczającej nas rzeczywistości²⁶⁶. Nazywanie przez nas przedmiotów i kategoryzowanie świata to sposób *porządkowania chaosu*, który jest wokół nas oraz *redukowanie złożoności rzeczywistości*²⁶⁷.

Jest to zjawisko ściśle powiązane z czymś co badacze ujmują w pojęciu *językowego obrazu świata*. Encyklopedia PWN definiuje to jako: *obraz rzeczywistości utrwalony w języku i niejako narzucany lub sugerowany użytkownikom, podlegający weryfikacji na drodze obserwacji kontrolnych i kryt. refleksji; nie jest odbiciem obrazu świata, lecz jego interpretacją, a po części kreacją (byty fikcyjne)*²⁶⁸. Próbą definicji tego pojęcia w swojej pracy zajmuje się M. Kowalski²⁶⁹. Autor wskazuje różne ujęcia tego zjawiska m. in. zaczynając od definicji J. Bartmińskiego i R. Tokarskiego jako *Pewien zespół sądów mniej lub bardziej utrwalonych w języku, zawartych w znaczeniach wyrazów lub przez te znaczenia implikowanych, który orzeka o cechach i sposobach istnienia obiektów świata pozajęzykowego*²⁷⁰.

Autor twierdzi, że definicję można traktować jako *reprezentatywną dla ontologicznego realizmu*²⁷¹. Obiekty pozajęzykowe mogą mieć różne typy: obiekty o strukturze

²⁶³ Cz. S. Bartnik, *Personalizm*, Oficyna wydawnicza „Czas”, Lublin 1995, s. 360.

²⁶⁴ F. Courth, *Bóg trójjedyniej miłości*, tł. M. Kowalczyk, adaptacja L. Bartel, F. Mickiewicz, (oryg. Der Gott der dreifaltigen Liebe), Pallotinum, Poznań 1997, s. 255.

²⁶⁵ P. Liszka, *Język jako instrument i przedmiot teologii*, Papieski Wydział Teologiczny we Wrocławiu, 2005, s. 341.

²⁶⁶ M. Kopińska, *Język jako ...*, dz. cyt., s. 55.

²⁶⁷ J. Maćkiewicz, *Kategoryzacja a ...*, dz. cyt., s. 53.

²⁶⁸ Językowy obraz świata, <https://encyklopedia.pwn.pl/haslo/jezykowy-obraz-swiata;3917950.html>, 11.11.22.

²⁶⁹ M. Kowalski, *Językowy obraz świata i jego istnienie*, „The Peculiarity of Man”, 2018, nr 22.

²⁷⁰ J. Bartmiński, R. Tokarski, *Językowy obraz...*, dz. cyt., s. 72.

²⁷¹ M. Kowalski, *Językowy obraz...*, dz. cyt., s. 142.

biologicznej (np.: kot, dąb), obiekty, które warunkowane są kulturowo (np.: wdowa, kawaler), subiektywne (np.: śmierć), abstrakcyjne (np.: procesy psychiczne), obiekty które nie występują w realnym świecie (np.: kapitalizm, komunizm)²⁷². Początki *problematyki językowego obrazu świata* sięgają już czasów Marcina Lutra (XVI wiek), który w jednej ze swych prac skupił się tym, że każdy z języków ma własne specyficzne cechy, które w procesie pojmowania świata są powodem, że te same rzeczy czy zjawiska są odmienne w różnych społecznościach, są różnie nazywane czy rozumiane – stąd świat jest pojmowany inaczej dla osobnych społeczności²⁷³. O języku nie możemy też mówić jako o prostej relacji język – świat ze względu na to, że języki stanowią dla oglądu świata *mentalne filtry*²⁷⁴. Piętno społecznego pochodzenia języka uwidacznia się w tym, że nie może on być kreowany w sposób indywidualny przez jednostkę, ale jest uzależniony od użycia zbiorowego – jako wytwór społeczności przekazywany z pokolenia na pokolenie wraz z światopoglądem i stereotypami – na przykład konkretny naród²⁷⁵. R. Grzegorczykowa wskazuje na różnicę pomiędzy *językowym obrazem świata* w językowym, a tekstowym aspekcie szczególnie w tekstach literackich: *Twórczość literacka, a w szczególności twórczość poetycka, przekazuje swoiste widzenie świata, właściwe poecie, odmienne od widzenia potocznego. Dokonuje się to oczywiście przez wykorzystanie systemu językowego, wykorzystanie polegające na przekraczaniu standardowych znaczeń i łączliwości, odbijających owo przeciętne widzenie świata*²⁷⁶. Autorka później zmienia zdanie w tej kwestii: *nie ma ostrych granic między konotacjami ustabilizowanymi a pojawiającymi się indywidualnie, tekstowo. Poezja wydobywa na jaw to, co jako potencja semantyczna dane jest systemowo*²⁷⁷. Pojęciem *tekstowego obrazu świata* posługuje się również J. Maćkiewicz (powołując się na pracę T. Dobrzyńskiej): rozróżnienie metafor języka (powtarzalnych i zleksykalizowanych) oraz poetyckich (tekstowych, kreacyjnych, indywidualnych),

²⁷² J. Maćkiewicz, *Co to jest językowy obraz świata*, Rozprawy i Analizy, Lublin 1999, s. 13.

²⁷³ A. Mańczyk, *Wspólnota językowa i jej obraz świata*, Krytyczne uwagi do teorii językowej Leo Weisgerbera, WSP, Zielona Góra 1982.

²⁷⁴ M. Kowalski, *Językowy obraz...*, dz. cyt., s. 142.

²⁷⁵ M. Kopińska, *Język jako ...*, dz. cyt., s. 55.

²⁷⁶ R. Grzegorczykowa, *Pojęcie językowego obrazu świata*, [w:] *Językowy obraz świata*, (red.) J. Bartmiński, Lublin 1990, s. 47.

²⁷⁷ R. Grzegorczykowa, *Wprowadzenie do semantyki językoznawczej*, Warszawa 2001, s. 164.

badaczka twierdzi, że kryterium wydobywania danych z tekstu musi być powtarzalność (konwencjonalność użycia)²⁷⁸.

Pojęcie *językowy obraz świata* (JOS) nie jest rozumiane w sposób jednoznaczny, mimo zaliczenia do podstawowych pojęć gramatyki treści i semantyki²⁷⁹. Termin ten pojawił się w okresie francuskiego i włoskiego Oświecenia, lecz dopiero na przełomie XVIII i XIX w. zagadnienie JOS omówili w aspekcie filozoficznym Johan Georg Hamann i Johann Gottfried Herder, a opracowania teoretyczne zjawiska opracował W. von Humboldt, który stwierdził, iż język jest nie tylko środkiem komunikowania się, lecz także *wyrazem ducha oraz oglądu świata przez użytkowników języka*²⁸⁰. J. G. Herder twierdził, że *każdy naród ma własny rezerwuar myśli, które stały się znakami*²⁸¹. Wyrażenia językowe mają różne pochodzenie. Między innymi kryterium obserwacyjne, które zastosował człowiek w sytuacji pierwotnej. Noah J. Jacobs w książce *Dzień nadawania imion w Raju* (ang. *Naming-day in Eden*) wskazuje na kilka przykładów: miejsce pochodzenia (np. pekińczyk), wielkość (np.: wieloryb), kolor (np.: szarak), wydawany dźwięk (np.: kwiczoł)²⁸². Niekiedy, tworzony opis z zjawisk za pomocą języka nie jest zgodny z stanem faktycznym nauki, a jedynie opartym na tradycji na przykład: *słońce wschodzi i zachodzi*²⁸³.

J. G. Herder i J. G. Humboldt twierdzili, że bez słów nie byłoby myśli, a zatem język jest niezbędny, aby otaczającej nas rzeczywistości nadać *kontury* oraz *kształt* – aby móc nadać nazwę jej i jej częściom – struktury myślenia są spójne ze strukturami języka. Człowiek myśli za pomocą języka i uczy się myśleć, stąd bez języka rozumowanie nie byłoby możliwe.²⁸⁴ Według Leo Weisgerbera doświadczanie świata zaczyna się na języku i kończy na nim, dlatego, że wszelkie poznanie jest dokonywane przez język i dzięki

²⁷⁸ J. Maćkiewicz, *Wyspa – językowy obraz wycinka rzeczywistości*, [w:] *Językowy obraz świata*, (red.) J. Bartmiński, Lublin 1990, s. 208.

²⁷⁹ R. Jedliński, *Językowy obraz świata wartości w wypowiedziach uczniów kończących szkołę podstawową*, Wydawnictwo Naukowe Akademii Pedagogicznej, Kraków 2000, s. 19.

²⁸⁰ R. Jedliński, *Językowy obraz...*, dz. cyt., s. 19.

²⁸¹ M. Kopińska, *Język jako ...*, dz. cyt., s. 57.

²⁸² J. Bartmiński, *Językowe podstawy obrazu świata*, Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, 2006, s. 6.

²⁸³ J. Anusiewicz, A. Dąbrowska, M. Fleischer, *Językowy obraz świata i kultura. Projekt koncepcji badawczej*, Język a kultura (tom: 13), Acta Universitatis Wratislaviensis (tom: 2218), Warszawa 2000, s. 30.

²⁸⁴ M. Kopińska, *Język jako ...*, dz. cyt., s. 56.

językowi²⁸⁵. Jednym z najważniejszych elementów jest przekazywanie wszystkich kwestii językowych kolejnemu pokoleniu: *istota językowego obrazu świata polega na tym – żeby zbadać i określić, jakie treści poznawcze, jaka wiedza i doświadczenie oraz wartościowanie ujmowanej poznawczo rzeczywistości jest zmagazynowane w języku, przezeń przenoszone i przekazywane przyszłym pokoleniom*²⁸⁶. Każde pokolenie dokłada coś swojego, ale ogółem jest oparte na tradycji i nie jest w stanie się od tej tradycji odciąć w sposób stuprocentowy, między innymi dlatego, że posługuje się tym samym językiem narodowym – nośnikiem oraz głównym elementem tradycji²⁸⁷.

Użycie języka to zwykle wytworzenie wypowiedzi, tekstu mogącego być całkowicie nowym tworem. Termin *tekst* może być rozumiany dwojako – jako wypowiedź pisemna lub ustna²⁸⁸. Tekst to *skończony i uporządkowany ciąg elementów językowych, mogących spełniać łącznie funkcję komunikatywną, a więc stanowiących jeden globalny znak*²⁸⁹. Generalnie tekst powinien posiadać spójność semantyczną (informacyjność i sensowość) będącą konsekwencją zamiaru komunikacyjnego.²⁹⁰

Media mogą w znaczącym stopniu wpływać na język Polaków, a polszczyzna, język ludzi wypowiadających w mediach np. telewizji stanowi dla wielu wzorzec tego, jak mówić poprawnie²⁹¹. W języku widzimy proces semantycznego przesunięcia; z biegiem czasu słowo lub wyrażenie może nabierać znaczenia innego od pierwotnego, skutkiem czego może to powodować nieporozumienia. Wilga Herman twierdzi, że neosemantyzmy często mogą powstawać pod wpływem ekspresji w obszarze potocznej polszczyzny, chęci wyrażania emocji, a użyte są skojarzenia związane z *kształtem, funkcją czy uczuciami towarzyszącymi danemu zjawisku*²⁹².

²⁸⁵ A. Mańczyk, *Wspólnota językowa...*, dz. cyt.

²⁸⁶ J. Anusiewicz, *Problematyka językowego obrazu świata w poglądach niektórych językoznawców i filozofów niemieckich XX wieku*, [w:] *Językowy obraz świata*, (red.) J. Bartmiński, Lublin 1990, s. 281-282.

²⁸⁷ G. Żuk, *Językowy obraz świata w polskiej lingwistyce przełomu wieków*, [w:] *Przeobrażenia w języku i komunikacji medialnej na przełomie XX i XXI wieku*, red. M. Karwatowska, A. Siwiec, Chełm 2010, s. 243.

²⁸⁸ R. Grzegorzczkowska, *Wstęp do...*, dz. cyt., s. 37.

²⁸⁹ T. Dobrzyńska, *Tekst. Próba syntezy* [w:] *Współczesny język polski*, (red.) J. Bartmański, Lublin 2001, s. 297.

²⁹⁰ R. Grzegorzczkowska, *Wstęp do...*, dz. cyt., s. 40.

²⁹¹ R. Pawelec, M. Trysińska, *Najnowsze słownictwo a współczesne media elektroniczne*, „Poznańskie Studia Polonistyczne Seria Językoznawcza”, Warszawa 2008, nr 17(37), s. 174.

²⁹² W. Herman, *Przesunięcie semantyczne w nowej leksyce radia i telewizji*, Semper, 2008, s. 23.

R. Grzegorzczkowska przedstawia 12 typów funkcji wypowiedzi ze względu na cel nadawcy:

- Informacyjna;
 - Informacyjna, opisowa – stanowią opis zdarzeń;
 - Informacyjna oceniająco-postulatywna – mające charakter o przekonaniu nadawcy komunikatu (np. *Jan nie powinien tam chodzić*);
- Pozainformacyjne;
 - Sprawcza – celem jest stwarzanie stanów rzeczy;
 - na mocy wiary – w tej grupie wydziela się podgrupy: magia ludowa (np. klątwy), akty sakralne (np. sakramenty, rozgrzeszenie);
 - na mocy konwencji społecznej – akty performatywne (akty prawne, wyroki, ustawy);
 - Ekspresywna – nie są skierowane na odbiorcę, stanowią ekspresję nadawcy i jego emocji czy sądów (np. *Jan wrócił!*, *Oby jutro było ładnie!*);
 - Nakłaniające: angażujące odbiorcę;
 - w celu wywołania działań;
 - pytania – w charakterze pytań;
 - dyrektywy, rozkazy – w charakterze próśb, żądań;
 - w celu na wpływania na stan mentalny odbiorcy;
 - na świadomość: perswazja – działające na świadomość;
 - bez świadomości odbiorcy: manipulacja – cele propagandowe, reklamy;

- Fatyczna – mają na celu nawiązanie, podtrzymanie lub kończenie kontaktu z odbiorcą;
- Kreatywna – teksty poetyckie, tworzenie i używanie jako formy myślenia, używanie języka jako sposobu kreowania wizji świata;
- Misteryjna – postawa kontemplacyjna wobec świata, uczestniczenie w *sacrum*²⁹³.

Nie tylko odbiorcy komunikatów politycznych, mogą różnie odbierać przekaz, ale także dziennikarstwo polityczne cechuje pewna prawidłowość, otóż Barbara Bogolebska opisuje, że w obrębie tej grupy *agonistyczność dziennikarstwa politycznego widoczna jest szczególnie w okresie przedwyborczym – ścierania się racji partii i ugrupowań dążących do władzy*²⁹⁴. Rodzaje publicystyki, np. publicystyka polityczna lub społeczna, zakłada zdolność przekazywania wyników analiz, raportów i wniosków odbiorcom przekazu – mniej lub bardziej przygotowanym do tego laikom: granice pomiędzy rodzajami dyskursu specjalistycznego i komunikacji nie muszą być jednak ostre²⁹⁵. Aby znacznie poprawić przekaz publikacji internetowych (nie tylko o charakterze politycznym) m. in. poradników istnieje szereg technik optymalizacyjnych m. in. pisanie w formie zrozumiałej jasnej czytelnej, pisanie w sposób życzliwy dla czytającego, jeśli to możliwe stosowanie przykładów praktycznych i unikanie specjalistycznego nazewnictwa, ogólne pisanie w myśl zasady *Napisz tak, żeby zrozumiała Twoja sąsiadka, a Twój kolega – profesor na pewno się nie obraził*²⁹⁶.

Według J. Grzeni Sieć w porównaniu z innymi mediami, jak radio, telewizja czy prasa daje użytkownikom bardzo dużą aktywność²⁹⁷. Stąd też wśród internetowych wypowiedzi często następuje modyfikacja pisowni m. in. zwielokrotnienie znaków interpunkcyjnych lub liter, stosowanie wersalików, dużych liter oznaczających krzyk, niestosowanie dużych liter tam gdzie wymaga tego konwencja ortograficzna, pisemne oznaczenie

²⁹³ R. Grzegorzczkowska, *Wstęp do...*, dz. cyt., s. 56-61.

²⁹⁴ B. Bogolebska, *Retoryka w odmianach dziennikarstwa*, „FOLIA LITTERARIA POLONICA”, 2016, nr 1(31), s. 55.

²⁹⁵ T. Konik, *Przypadki retorycznie istotnych zmian redakcyjnych w tłumaczeniach publicystyki prasowej o tematyce gospodarczej*, „Retoryka w komunikacji specjalistycznej”, (red.) M. Załęska, Polskie Towarzystwo Retoryczne, Warszawa 2015, s. 191-198.

²⁹⁶ M. Worsowicz, *Jak pisać ...* dz.cyt., s. 281-298.

²⁹⁷ J. Grzenia, *Komunikacja Językowa ...*, dz. cyt., s. 21.

reakcji niewerbalnych (np. śmiech jako: hahaha), opisy akcji (np. *wychodzę*), stosowanie emotikon (połączenia znaków przypominających m.in. twarz, mimikę), niestandardowe spacje lub kombinacje tekstu i grafiki, stosowanie akronimów (skrót typu *ASAP* – *ang. As Soon As Possible* – oznaczające zwrot najszybciej jak to możliwe)²⁹⁸. Wypowiedzi pisane w mediach społecznościowych często mogą mieć powyższe charakterystyki dodatkowo m.in. różny poziom znajomości ortografii, błędy stylistyczne czy nieumiejętne korzystanie ze znaków interpunkcyjnych, co stanowi informację do oczyszczenia ich przed rozpoczęciem analiz.

Inną kwestią jest zjawisko wartościowania w języku. J. Puzynina w pracy *Problemy wartościowania w języku i w tekście* autorka stawia z pewnością, że wartościowanie jest dokonywane w tworzonych tekstach; podmiotami tego wartościowania są ich autorzy, a przedmiotem jest otaczająca rzeczywistość w tym człowiek, który wartościuje siebie oraz własne możliwości i poczynania. Oprócz wartościowania podmiotu mamy również wartościowania bohaterów literackich oraz wartościowania relacjonowane mowie zależnej i niezależnej. Istnieją często występujące przypadki wartościowań, które są ukryte w presupozycjach (w presupozycjach wniosek wynika jednakowo ze zdania, jak i z jego negacji: np. ze zdania „Aktualny król Wielkiej Brytanii jest łysy”, jak i zdania „Aktualny król Wielkiej Brytanii nie jest łysy” wynika zdanie „Wielka Brytania ma obecnie króla”). Wartościowaniu mogą służyć różne jednostki struktur językowych m. in.: wyrazy (piękny, brzydki), związki frazeologiczne (coś jest do kitu), konstrukcje słowotwórcze (artykulisko). Wartościowania mają różne funkcje, ale przede wszystkim są elementem identyfikacji grupowej.²⁹⁹

Zjawisko występowania wartościowania istnieje w różnych dziedzinach życia. Marcin Szewczyk w swojej pracy *Linguistic discrimination against the Roma in the Polish-language Web 2.0 discourse* wskazał na występowanie w polskich *social mediach* pewnej mającej podłoże kulturowe skali dyskryminacji Romów poprzez użycie odpowiednich słów wartościujących³⁰⁰. W wypowiedziach polityków widać dosyć mocno związek

²⁹⁸ J. Grzenia, *Komunikacja Językowa ...*, dz. cyt., s. 119-149.

²⁹⁹ J. Puzynina, *Problemy wartościowania ...*, dz. cyt.

³⁰⁰ M. Szewczyk, *Linguistic discrimination against the Roma in the Polish-language web 2.0 discourse*, „ROCZNIKI NAUK SPOŁECZNYCH”, 2022, nr 50(2).

wartościowania z perswazją.³⁰¹ Dominującą funkcją języka polityki jest funkcja perswazyj³⁰².

W dyskursie politycznym Aleksy Awdiejew podaje 4 różne formy wartościowania: (I) zasada kooperacji – zbiorowa odpowiedzialność, gdzie celem jest współpraca z uczestnikiem dyskusji dla osiągnięcia wspólnej korzyści informacyjnej (II) tematyzacja sądów oceniających w sytuacji, gdy są one sprzeczne z orientacją odbiorcy (przykład: *Wybrany w pośpiechu parlament nie będzie reprezentatywny* - wypowiedź sugeruje że parlament będzie wybrany w pośpiechu); (III) Wymuszone implikatury aksjologiczne (jako ukryte presupozycje *implikowane przez wypowiedzenie w jakimś kontekście nawet wtedy, kiedy nie wynikają logicznie z tego, co zostało wypowiedziane*³⁰³, na przykład wypowiedź: *W miłym sercu Janusza Korwin-Mikkego regionie bałkańskim też wszystko się zaczynało od głupawych przemówień* – sugerująca, że J. Korwin-Mikke gloryfikuje konflikt bałkański oraz, że ten konflikt może być potencjalnym czynnikiem kryzysu w Polsce, mimo że sugerowanie to nie jest wynikiem implikacji logicznej; (IV) tzw. pyskówka polityczna inaczej mówiąc bojowa postawa działaczy partyjnych (lub uczestników komunikowania politycznego)³⁰⁴.

Istnieje również szereg wartościowań typowo negatywnych stosowanych przez polityków. Anna Czelakowska wskazuje m. in: negatywnie wartościowanie oponenta wskazywanie jego błędów i niewłaściwego postępowania, krytykuje się sposób mówienia (np.: jazgot, bełkot), zachowania osoby (np.: „jest Pan bezczelny”), opis sytuacji (np.: naciąganie, kłamstwo)³⁰⁵. Czasami w wypowiedziach niektórych polityków można zaobserwować stosowanie wyrazów o zabarwieniu ekspresywnym, przenoszących wykładnik oceniający (np.: bandyckie napady, kompletna destrukcja)³⁰⁶. W przypadku wypowiedzi polityków ważnym środkiem negatywnego wartościowania, obserwowanym w wielu wypowiedziach polityków, są presupozycje – przez nie można przekazać informację niepewną lub całkowicie fałszywą, a jednocześnie wyprzec się intencjonalności jej

³⁰¹ E. Kapela, *Językowe środki ...*, dz. cyt., s. 158.

³⁰² B. Walczak, *Co to jest język polityki?*, 2017, s. 20.

³⁰³ G. Gazdar, *Pragmatics, Implicature, Presupposition and Logical Form*, New York 1979, s. 38.

³⁰⁴ A. Awdiejew, *Wartościowanie wymuszone, a szacunek dla odbiorcy w dyskursie politycznym*, 2018.

³⁰⁵ A. Czelakowska, *Językowe środki wartościowania negatywnego w wypowiedziach polityków*, *Annales Academiae Paedagogicae Cracoviensis, „Studia Linguistica IV”*, 2008, s. 60-61.

³⁰⁶ S. Grabias, *O ekspresywności języka*, Lublin 1981, s. 33; A. Czelakowska, *Językowe środki ...*, dz. cyt., s. 60-61.

wypowiedzenia. Kolejnym sposobem przekazywania wartościowania negatywnego są metafory, które odwołują się do wiedzy ogólnej, widzenia świata czy stereotypów (np. faryzeizm – jako obłuda przypisywana do postawy polityka)³⁰⁷.

Język to kopalnia wiedzy na temat świata i człowieka. Przenosi emocje i zawiera łańdunki, które mogą wywołać określone odczucia. Jest nie tylko wyrażeniem myśli i doświadczeń, ale zawiera się w nim doświadczenie przeszłych pokoleń światopoglądów, stereotypów zachowań ukutych na przestrzeni wieków. Język pozwala ukazywać jak dana wspólnota językowa interpretuje otaczającą rzeczywistość. Badając język, możemy poznać światopoglądy (jako właściwy charakter narodów), możemy opisać wspólnoty posługujące się tymi językami³⁰⁸. System językowy to w najczęstszym przypadku jedyne i niezbędne narzędzie komunikacji międzyludzkiej. JOS można zrekonstruować badając wypowiedzi językowe (elementy systemu języka które pełnią rolę nośników obrazów świata – m.in. słownictwo, frazeologia, struktura i składnia, stylistyka)³⁰⁹. Każda społeczność, a więc i polityczna mimo używania technicznie tego samego języka może posiadać swój własny *językowy obraz świata*.

1.7. Big Data

Analizując dane z serwisów społecznościowych lub innych platform, w których ilość danych jest ogromna, ciągle rośnie, jest zmienna, możemy mówić o wykorzystaniu wielkich zbiorów danych czyli *Big Data*. Jest to pojęcie, które zaczęło być używane w latach 90. XX wieku. Jedno z pierwszych pojęć opracowali w 1997 Michael Cox i David Ellsworth. Badacze zdefiniowali pojęcie *Big Data* jako duży zbiór danych do analizy, który należy zwiększać w celu wydobycia danych informacyjnych³¹⁰. Kolejną definicję w 2001 stworzył analityk Laney Doug, opierając tę definicję o model 3V³¹¹. W modelu tym, aby opisać termin *Big Data* musimy mieć dane posiadające 3 atrybuty jak *Volume*

³⁰⁷ A. Czelakowska, *Językowe środki...*, dz. cyt., s. 60-64.

³⁰⁸ G. Żuk, *Językowy obraz...*, dz. cyt., s. 239–257.

³⁰⁹ J. Anusiewicz, A. Dąbrowska, M. Fleischer, *Językowy obraz...*, dz. cyt., s. 31.

³¹⁰ M. Cox, D. Ellsworth, *Managing Big Data for Scientific Visualization*, ACM SIGGRAPH '97 Course #4, *Exploring Gigabyte Datasets in Real-Time: Algorithms, Data Management, and Time-Critical Design*, Los Angeles 1997.

³¹¹ L. Doug, *Data Management: Controlling Data Volume, Velocity, and Variety*, „*Application Delivery Strategies*”, META Group (currently with Gartner), 2001.

(Objętość), *Variety* (Różnorodność), *Velocity* (Szybkość przetwarzania). W 2012 firma *Gartner.com* dodała jeszcze dwie charakterystyki jak *Variability* (Zmienność) i *Complexity* (Złożoność).

W dalszych latach pojęcie zostało przekształcane i pojawiły się definicje *Big Data* sformułowane przez firmy, zajmujące się analizą danych.

Firma IBM opisuje *Big Data* jako strukturę różnorodnych danych, które są generowane z różnych źródeł w dużej ilości wraz z dużą ilością. IBM tworzy definicję w oparciu o 4 charakterystyki: *Volume* (Objętość), *Velocity* (Szybkość przetwarzania), *Variety* (Różnorodność), *Veracity* (Wiarygodność).³¹²

SAS tworzy definicję w oparciu o 3 atrybuty *Volume* (Objętość), *Velocity* (Szybkość przetwarzania), *Variety* (Różnorodność) + 2 *Variability* (Zmienność), *Veracity* (Wiarygodność)³¹³.

Patrząc na powyższe definicje można ustandaryzować, że termin *Big Data* charakteryzują następujące atrybuty:

- Objętość – opisywana jako jeden z atrybutów V – *Volume*, pokazująca, że ilość danych jest bardzo dynamiczna i będzie rosnąć;
- Szybkość – opisywana jako jeden z atrybutów V – *Velocity*, pokazująca, że dane napływają strumieniowo, często wymagają odpowiedniej obsługi w postaci wykorzystania nowoczesnych technologii. Często muszą być pobierane porcjami w odpowiedni sposób³¹⁴;
- Różnorodność – opisywana jako jeden z atrybutów V – *Variety* – wskazująca, że dane mogą pochodzić z różnych źródeł, często w różnych formach i różnych strukturach modeli. Przykładem mogą być dane z działów wewnątrz przedsiębiorstwa przechowywane w bazach danych, dane ze źródeł zewnętrznych,

³¹² Big data by IBM, <https://www.redbooks.ibm.com/redpapers/pdfs/redp5070.pdf>, 20.08.2021.

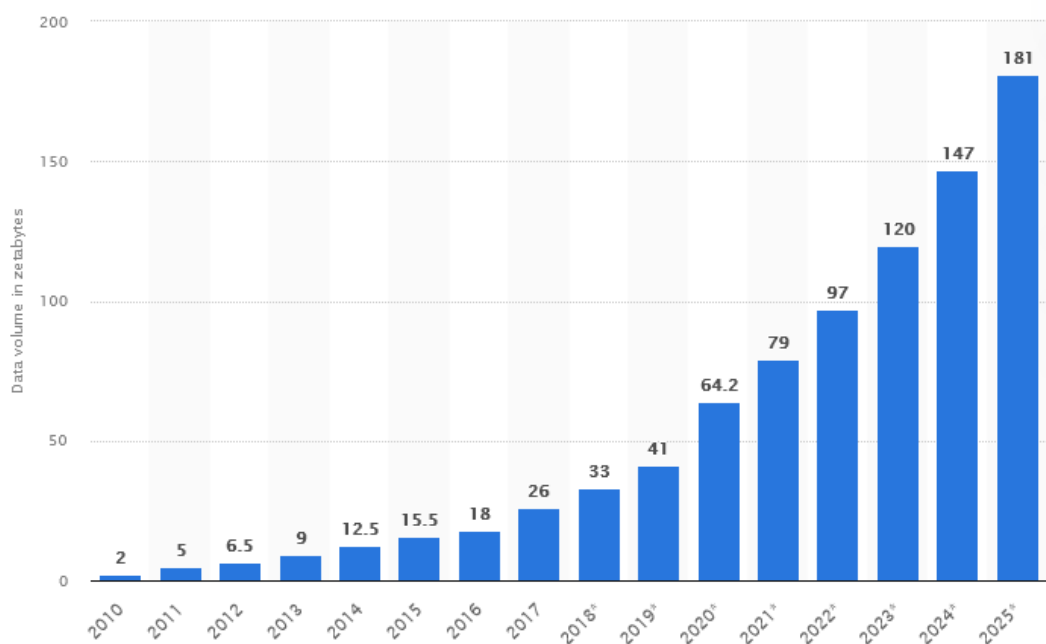
³¹³ What is Big Data, https://www.sas.com/pl_pl/insights/big-data/what-is-big-data.html, 26.08.2021.

³¹⁴ M. Tabakow, J. Korczak, B. Franczyk, *Big data – definicje, wyzwania i technologie informatyczne*, Uniwersytet Ekonomiczny we Wrocławiu, Wrocław 2014.

internetowych (komentarze, fora, blogi, vlogi, itd.), dane z *Deep Webu*, które nie są indeksowane przez większość popularnych wyszukiwarek oraz inne³¹⁵;

- Zmienność – dane, które są zmiennie w czasie, natężenie napływu może być cykliczne lub sezonowe, może być związane z przemianami gospodarczymi lub nadchodzącymi wydarzeniami, na przykład sprzedaż prezentów w okresie świąt³¹⁶;
- Złożoność – jest również powiązana z różnorodnością. Dane mogą mieć różny format, różną strukturę;
- Wartość – prezentują wartość informacyjną, dają możliwość do wyciągania wniosków na podstawie licznych danych. Wykorzystanie tych wartości może dać przedsiębiorstwu ogromne korzyści, jak efektywniejszy model biznesowy, spersonalizowane oferty dla klientów, lepszą strategię marketingową³¹⁷;

Rysunek 25. Ilość danych w zettabajtach (1024^7 bajtów) utworzonych, przechwytanych, skopiowanych i wykorzystanych na całym świecie w latach 2010-2025.



Źródło: Ilość danych w zettabajtach, www.statista.com, 26.08.2021.

³¹⁵ B. He, M. Patel, Z. Zhang, K.C.C Chang, *Accessing the deep web*, *Communications of the ACM*, New York, USA, 2007, nr 50 (5), s. 94-101.

³¹⁶ M. Tabakow, J. Korczak, B. Franczyk, *Big data...*, dz. cyt., s.142.

³¹⁷ M. Tabakow, J. Korczak, B. Franczyk *Big data...*, dz. cyt., s.143.

IBM podaje następujące korzyści z wykorzystania analizy *Big Data*³¹⁸:

- Wspieranie decyzji w czasie rzeczywistym, kiedykolwiek i gdziekolwiek;
- Korzystanie ze wszystkich rodzajów informacji, które można wykorzystać w podejmowaniu decyzji w oparciu o dane;
- Danie ludziom na różnych stanowiskach możliwość odkrywania i analizowania informacji oraz oferowania innym spostrzeżeń;
- Optymalizowanie różnych rodzajów decyzji, niezależnie od tego, czy są podejmowane przez osoby fizyczne, czy w zautomatyzowane systemy z wykorzystaniem wniosków opartych na analityce *Big Data*;
- Zapewnienie wglądu ze wszystkich perspektyw i horyzontów czasowych, od raportów historycznych, do analiz w czasie rzeczywistym oraz do modelowania predykcyjnego;
- Poprawianie wyników biznesowych i zarządzanie ryzykiem.³¹⁹

Niektóre prace wskazują, że wadami analiz *Big Data* może być obawa klientów o swoje dane, bądź powodowanie uzależnień zakupowych po proponowaniu przez algorytm innych produktów, których mogą wcale nie potrzebować³²⁰.

W polityce również istnieje wiele prac związanych z wykorzystaniem analiz *Big Data*. E. Colleoni, A. Rozza, A. Arvidsson w swojej pracy przedstawiają klasyfikację wyborców USA ze względu na poglądy polityczne, z podziałem na Republikanów i Demokratów, wykorzystując do tego dane z serwisu Twitter oraz *Big Data* wraz z budowaniem modeli predykcyjnych³²¹. Inny przykład wykorzystania analiz z danych z social mediów w polityce to prace A. Jungherra (około 115 prac związanych z wykorzystaniem

³¹⁸ Big data by...dz. cyt., 20.08.2021.

³¹⁹ Big data by...dz. cyt., 20.08.2021

³²⁰ T. M. Le, S.Y. Liaw, *Effects of Pros and Cons of Applying Big Data Analytics to Consumers Responses in an E-Commerce Context*, „MDPI”, 2017, nr 9(5).

³²¹ E. Colleoni, A. Rozza, A. Arvidsson, *Echo Chamber*...dz. cyt.

Twittera w polityce)³²². Praca E. Bellio, G. Crestini, A. Arkoudas pokazuje z kolei wykorzystanie Twittera przez polityków w kampanii wyborczej 2016 w USA³²³.

W polskiej literaturze na szczególną uwagę zasługują opracowania W. Gogołka, który w swoich pracach wydobywa wartościowe dane do wyciągania wniosków, tworząc termin *rafinacji danych*³²⁴. P. Kuczma w swoich pracach³²⁵ wskazał, że dane z sieci również te generowane przez użytkowników Internetu, są wiarygodnym źródłem informacji. D. Jaruga wraz z W. Gogołkiem wykorzystując *Big Data* zajmowali się znalezieniem wyrażen – sentymentów, które niosą określony ładunek emocjonalny (pozytywny, negatywny, bądź neutralny).³²⁶ Warto zwrócić uwagę na niezwykle istotne spostrzeżenie: *Fundamentalnym warunkiem powodzenia RI [rafinacji informacji] jest dysponowanie dostateczną ilością informacji, by przekroczyć próg nieufności do uzyskanych wyników pozwalających na uzyskanie wiarygodnych wyników, np. predykcję zjawiska. System musi „się uczyć” – osiągać zakładaną statystyczną istotność*³²⁷.

1.8. Kluczowe pojęcia i definicje

Najważniejsze pojęcia i związane z nimi teorie, używane w niniejszej pracy zostały już zaprezentowane, jest jeszcze jednak kilka pojęć z różnych dziedzin, na które warto zwrócić uwagę.

Media społecznościowe – A. Kaplan, M. Haenlein zdefiniowali media społecznościowe jako *grupę bazujących na internetowych rozwiązaniach aplikacji, które opierają się na ideologicznych i technologicznych podstawach Web 2.0 i które umożliwiają stworzenie i wymianę treści wygenerowanych przez użytkowników w sieci*³²⁸. Encyklopedia PWN definiuje media społecznościowe jako: *technologie internetowe i mobilne, umożliwiające kontakt pomiędzy użytkownikami poprzez wymianę informacji, opinii i wiedzy*.³²⁹

³²² A. Jungherr, *Twitter in Politics* ..., dz. cyt.

³²³ E. Bellio, G. Crestini, A. Arkoudas, *Twitter and...*, dz. cyt.

³²⁴ W. Gogołek, *Rafinacja informacji* ..., dz. cyt., s. 229-238.; W. Gogołek, *Potencjał informacyjny* ..., dz. cyt., s. 253-263.; W. Gogołek, i inni, *Z badań nad* ..., dz. cyt., s. 31-43, http://studiamedioznawcze.pl/article.php?date=2015_3_62&content=gogolek&lang=pl, 10.08.2021.

³²⁵ P. Kuczma, W. Gogołek, *Web Information* ..., dz. cyt., s. 241-263.

³²⁶ W. Gogołek, D. Jaruga, *Z badań* ... dz. cyt, 2016; http://studiamedioznawcze.pl/article.php?date=2016_4_67&content=gogolek&lang=pl, 10.08.2021.

³²⁷ W. Gogołek, D. Jaruga, *Z badań* ..., dz. cyt., s. 110.

³²⁸ A. M. Kaplan, M. Haenlein, *Users of the world, unite! The challenges and opportunities of Social Media*, „Business Horizons”, nr 53(1), 2010, s. 59-68.

³²⁹ Media Społecznościowe, sjp.pwn.pl/slowniki/media-spoecznościowe.html, 13.09.2021.

NLP (ang. *Natural Language Processing*) powstało w latach 50. jako skrzyżowanie sztucznej inteligencji (ang. *Artificial intelligence*) i lingwistyki. NLP różniło się pierwotnie od wyszukiwania informacji tekstowych (IR, ang. *information retrieval*), które używa wysoce skalowalnych technik opartych na statystykach do efektywnego indeksowania i wyszukiwania dużych ilości tekstu. Z czasem jednak NLP i IR nieco się zbliżyły. Obecnie NLP wywodzi się z kilku bardzo różnorodnych dziedzin, co wymaga od dzisiejszych badaczy i programistów NLP znacznego poszerzenia swojej bazy wiedzy.³³⁰

Rafinacja dużych zasobów sieciowych odnosi się do procesu pozyskiwania ze zbiorów danych cennych informacji – niekiedy wymaga to wielu przekształceń i połączeń zbiorów danych³³¹.

Internet i Sieć dwa pojęcia, które często stosowane są zamiennie w celu ułatwienia i tak też będą stosowane w niniejszej pracy. W powszechnym (bardziej praktycznym) użyciu pierwsze pojęcie zwykle odnosi się do kwestii technicznych, drugie zaś do relacji społecznych³³². Internet to *rozbudowany układ światowych sieci komputerowych, który nie posiada uporządkowanej ani hierarchicznej struktury, organu zarządzającego (każda z sieci jest administrowana oddzielnie) oraz formalnego właściciela*³³³. Sieć natomiast (w rozumieniu sieć komputerowa) to według definicji PWN: *zestaw sprzętu i oprogramowania pozwalający przekazywać informacje między komputerami (zwanymi węzłami sieci komputerowej)*³³⁴.

Modelowanie predykcyjne jest powszechnie stosowaną techniką statystyczną do przewidywania przyszłych zachowań³³⁵. Metody i techniki modelowania predykcyjnego to forma technologii eksploracji danych, która działając poprzez analizę historycznych danych oraz danych bieżących, poprzez generowanie modelu pomaga przewidywać przyszłe wyniki. W modelowaniu predykcyjnym zbierane są dane (wejścia modelu), tworzony

³³⁰ P. M. Nadkarni, L. Ohno-Machado, W. W. Chapman, *Natural language processing: an introduction*, Journal of the American Medical Informatics Association, 2011, nr 18(5), s. 544-551.

³³¹ W. Gogolek, *Rafinacja dużej ...*, dz. cyt.

³³² M. Gosek, *Cyfrowy uczeń, Dylematy mediatyzacji*, Akapit, Toruń 2020.

³³³ A. Adamski, *Media w ...*, dz. cyt., s. 28.

³³⁴ Definicja Sieci PWN, <https://encyklopedia.pwn.pl>, 22.03.2022.

³³⁵ Predictive modeling, www.gartner.com/en/information-technology/glossary/predictive-modeling, 13.09.2021.

jest model statystyczny, dokonywane jest prognozowanie, a model jest weryfikowany (lub modyfikowany) w przypadku dostępności dodatkowych danych.

1.9. Podsumowanie rozdziału

Podsumowując, w rozdziale 1. skupiono się na opisie mediów, mediatyzacji oraz szczególną uwagę poświęcono modelom komunikacji w tym modelowi Lazarsfelda i Katza, który wprowadzał podejście sondażowe i koncentrował się na zastosowaniu metod statystycznych, ilościowych oraz jakościowych. Rozwinięto w tym rozdziale aspekty systemu społecznego, medialnego, ale przede wszystkim skupiono się na systemie politycznym, który jest związany z niniejszą pracą. Omówiono kwestie ściśle związane zarówno z językiem, jak i wykorzystaniem technologii umożliwiającymi przetwarzaniem dużych zbiorów danych określanych jako *Big Data*.

Rozdział 2. Mediatyzacja polskiej polityki

2.1. Przemiany społeczno-gospodarcze w Polsce w latach po 1989 roku

Polska przed 1989 r. tak, jak wiele innych państw była pod rządami niedemokratycznymi, komunistycznymi, w którym panowała cenzura, brak wolnych wyborów, konstytucyjne uznaniem *przewodniej roli* partii komunistycznej oraz *umacnianie przyjaźni i współpracy* ze Związkiem Sowieckim jako Polska Rzeczpospolita Ludowa (PRL)³³⁶. W drugiej połowie lat 70 z chwilą pojawienia się w Polsce zorganizowanej opozycji, jednym z głównych jej celów było się demaskowanie fałszerstw i przekłamań dokonywanych przez rządzącą Polską Zjednoczoną Partię Robotniczą (PZPR) sprawująca rządy w PRL w latach 1948-1989³³⁷. Masowe protesty z lat 1956, 1968, 1970 i 1976 oraz opór społeczeństwa spowodował zmiany gospodarcze w Polsce³³⁸. Zmiany polityczne w Polsce zostały uwieńczone kompromisem podczas obrad Okrągłego Stołu pomiędzy władzami komunistycznymi, a opozycją solidarnościową. Po 1989 roku zaczął się długotrwały, mający istotny wpływ na większość aspektów życia społecznego i gospodarczego, proces przechodzenia polskiej gospodarki z systemu gospodarki centralnie sterowanej do systemu gospodarki kapitalistycznej (wolnorynkowej). Transformacji tej towarzyszyła również transformacja ustrojowa.³³⁹

³³⁶ Polska Rzeczpospolita Ludowa, <https://polskiemiesiace.ipn.gov.pl/mie/wszystkie-wydarzenia/czerwiec-1989/polska/historia/117922,PRL-Polska-Rzeczpospolita-Ludowa.html>, 20.06.2022.

³³⁷ A. Dudek, *Historia polityczna Polski 1989–2012*, Znak, Kraków 2013, s. 33-34.

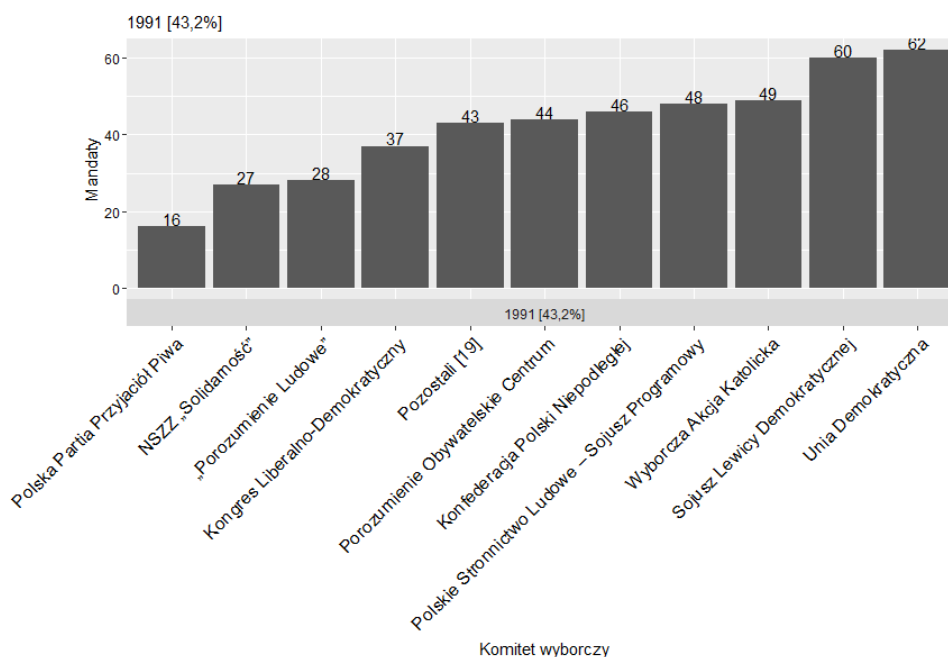
³³⁸ PRL czerwiec 1989, <https://polskiemiesiace.ipn.gov.pl/mie/wszystkie-wydarzenia/czerwiec-1989/polska/historia/117922,PRL-Polska-Rzeczpospolita-Ludowa.html>, 20.06.2022.

³³⁹ J. Kaliński, Z. Landau, *Gospodarka Polski w XX wieku*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 369.

2.2. Polityka w latach od 1989 do aktualnego stanu (2021)

Pod koniec sierpnia 1989 r. ówczesny premier Tadeusz Mazowiecki zaprezentował szereg reform politycznych i gospodarczych. W początkowej fazie nastąpiły zmiany personalne oraz organizacyjne w najważniejszych instytucjach państwowych, których celem było ograniczenie wpływu komunistów. W sprawach gospodarczych kluczową rolę powierzono Leszkowi Balcerowiczowi. Na początku 1990 roku Polska Zjednoczona Partia Robotnicza (PZPR) została rozwiązana – część członków utworzyła partię o nazwie Socjaldemokracja Polska (SdRP). Partia Zjednoczone Stronnictwo Ludowe (ZSL) została przeobrażona w Polskie Stronnictwo Ludowe (PSL). Powstawały nowe partie, które skupiały przedstawicieli dawnej opozycji. Zwycięstwo Lecha Wałęsy i przegrana T. Mazowieckiego w wyborach prezydenckich w 1990 r. pogłębiły proces spolaryzowania obozu antykomunistycznego. T. Mazowiecki złożył rezygnację z funkcji premiera, a na jego miejsce został wybrany Jan Krzysztof Bielecki. Następowwały częste spory w koalicji rządzącej, co wpłynęło na niską frekwencję wyborów parlamentarnych w 1991.

Rysunek 26. Uzyskane mandaty do Sejmu w 1991[frekwencja].

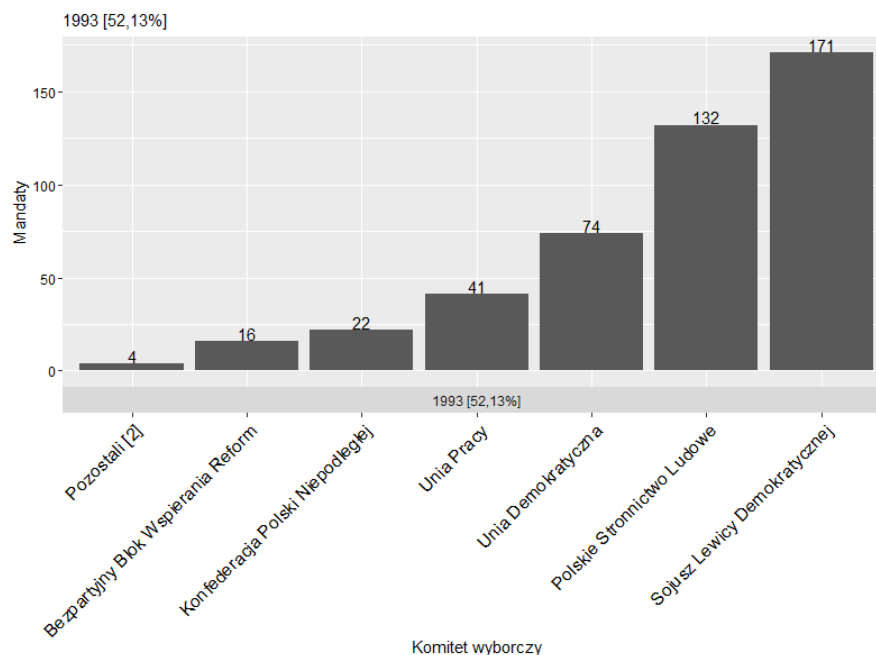


Źródło: opracowanie własne na podstawie danych z pkw.gov.pl (20.11.22) z wykorzystaniem biblioteki ggplot2.

Najwięcej głosów otrzymała partia Unia Demokratyczna (UD) T. Mazowieckiego oraz postkomunistyczny Sojusz Lewicy Demokratycznej (SLD), skoncentrowany wokół Aleksandra Kwaśniewskiego. Prezesem Rady Ministrów został Jan Olszewski z partii

Porozumienie Centrum (PC) – w skład Rady Ministrów nie powołano L. Balcerowicza. Rząd J. Olszewskiego popadł w konflikt z prezydentem L. Wałęsą oraz parlamentem i został odwołany ze stanowiska w czerwcu 1992 r. Po nieudanej próbie stworzenia rządu przez Waldemara Pawlaka (PSL), rząd został sformowany przez Hannę Suchocką (UD). Rząd H. Suchockiej został odwołany w maju 1993 r.

Rysunek 27. Uzyskane mandaty do Sejmu w 1993 [frekwencja].



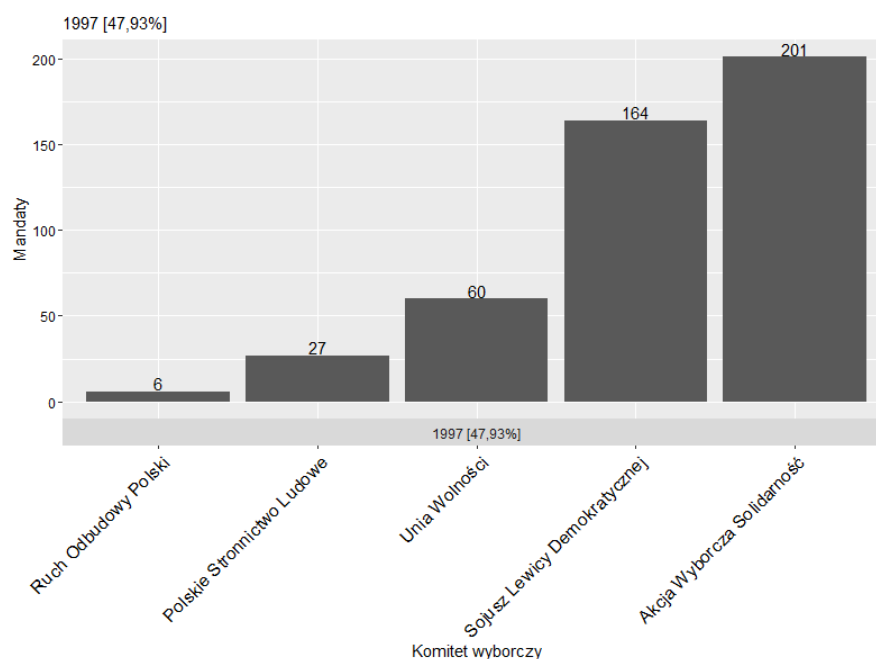
Źródło: opracowanie własne na podstawie danych z pkw.gov.pl (20.11.22) z wykorzystaniem biblioteki ggplot2.

W wyborach we wrześniu 1993 r. zwyciężyło ugrupowanie SLD, a klęskę poniosły ugrupowania postsolidarnościowe. Urząd Prezydenta RP pozostał w rękach dawnej opozycji. Społeczeństwo było zmęczone oraz zdezorientowane niejasnymi sporami politycznymi oraz przeprowadzonymi reformami gospodarczymi, przez co wybrało ugrupowania lewicowe, które przekonywały bezbolesnymi zmianami systemowymi. W 1993 r. W. Pawlak został wybrany na Prezesa Rady Ministrów – rząd został sformowany przez koalicję SLD i PSL. W 1995 r. W. Pawlak po wotum nieufności został odwołany i jego miejsce zajął Józef Oleksy – niecały rok później z powodu pojawienia się zarzutów o szpiegostwo J. Oleksy podał się do dymisji. Jego miejsce zajął Włodzimierz Cimoszewicz. Skuteczna kampania w 1995 r. dała wygraną A. Kwaśniewskiemu na urząd Prezydenta RP. Lewica odzyskała władzę w parlamencie i objęła urząd prezydencki. W 1997

r. uchwalono demokratyczną Konstytucję Rzeczypospolitej Polskiej. Po 6 latach od powstania solidarnościowych rządów, te ponownie znalazły się w opozycji.

W 1996 r. partia Akcja Wyborcza (AWS) zyskiwała poparcie wśród społeczeństwa i w kolejnych wyborach objęła koalicyjną władzę wraz z Unią Wolności (UW).

Rysunek 28. Uzyskane mandaty do Sejmu w 1997[frekwencja].



Źródło: opracowanie własne na podstawie danych z pkw.gov.pl (20.11.22) z wykorzystaniem biblioteki ggplot2.

W 1997 r. wybrano rząd Jerzego Buzka, a funkcję ministra finansów objął ponownie L. Balcerowicz (UW). Pod wpływem pogorszenia się sytuacji gospodarczej, niesprawnie wdrażanych reform i przejawów korupcji, wewnętrzne spory koalicyjne doprowadziły w 2000 r. do ustąpienia L. Balcerowicza z funkcji ministra finansów. W następnych wyborach zmęczone politycznymi niesnaskami oraz trudną sytuacją ekonomiczną społeczeństwo opowiedziało się na kontynuację prezydentury A. Kwaśniewskiego.³⁴⁰ Dnia 2.03.1999 po latach starań i zabiegów dyplomatycznych Polska została przyjęta do NATO.

³⁴⁰ J. Kaliński, Z. Landau, *Gospodarka Polski ...*, dz. cyt., s. 369-372.

Główne pozytywne zmiany polityczno-gospodarcze po 1989 roku ³⁴¹:

- *szybkie zrównoważenie rynków, zwłaszcza dóbr konsumpcyjnych;*
- *złagodzenie hiperinflacji i usunięcie ryzyka zupełnego rozchwiania się systemu gospodarczego;*
- *stabilizacja pieniądza;*
- *uwolnienie cen, konsekwencją czego było przywrócenie stosunków pieniężnych i otwarcie gospodarki, co przyczyniło się do racjonalizacji wyceny zasobów gospodarczych oraz decyzji mikroekonomicznych;*
- *poddanie przedsiębiorstw presji konkurencyjnej w wyniku czego znacząco wzrósł eksport na rynki zachodnie;*
- *zrównoważenie budżetu.*

Natomiast do negatywnych skutków powinno się przede wszystkim zaliczyć³⁴²:

- *spadek poziomu produkcji i Produktu Krajowego Brutto;*
- *destrukcję znacznej części potencjału gospodarczego przedsiębiorstw państwowych oraz upadek wielu obszarów produkcji krajowej;*
- *obniżenie się płac realnych i poziomu konsumpcji;*
- *znaczący wzrost bezrobocia.*

2.3. Sytuacja społeczna po 1989 roku

Mimo, że proces transformacji dokonał się głównie w płaszczyźnie politycznej, przejście z totalitaryzmu i dyktatury do demokracji i państwa prawa spowodowało szereg zmian w społeczeństwie.

W pierwszej połowie lat 90-tych nastąpił szereg zmian społecznych. Przede wszystkim nasiliły się dysproporcje w rozwoju gospodarczym, skutkiem czego był proces

³⁴¹ G. Kołodko, *Transformacja gospodarki polskiej. Sukces czy porażka?*, BGW, Warszawa 1992; M. Dahl, *Wpływ transformacji ekonomicznej na sferę społeczną w Polsce po 1989 roku*, Myśl Ekonomiczna i Polityczna, 2010.

³⁴² G. Kołodko, *Transformacja gospodarki ...*, dz. cyt.; M. Dahl, *Wpływ transformacji ...*, dz. cyt.

materialnego rozwarstwienia społeczeństwa oraz pogorszenie sytuacji wielu gospodarstw domowych. Do 1994 r. poziom bezrobocia wzrósł od 0,3% do 16%, w kolejnych latach nastąpiło odwrócenie tego trendu – do końca 1998 poziom ten spadł do 10,4%, by potem z kolei podnieść się do poziomu 20,6% w roku 2004, na co wpłynęła perspektywa członkostwa w Unii Europejskiej. W 1997 r. 13,3% gospodarstw domowych mogło być sklasyfikowane w granicy ustawowego ubóstwa (posiadało prawo do ubiegania się o świadczenia społeczne) – do końca 2004 r. odsetek ten wzrósł do 19,2%³⁴³. Najgorsza sytuacja panowała na wsiach, szczególnie w słabo rozwiniętych rolniczo częściach północnej i północno-wschodniej Polski. Wzrost bezrobocia spowodował wzrost ubóstwa oraz biedy. Ograniczenia budżetowe na ochronę zdrowia doprowadziły do deficytu usług ochrony zdrowia, co spowodowało szybką komercjalizację usług medycznych, która stała się dostępna tylko dla zamożniejszej części społeczeństwa.³⁴⁴

W sferze społecznej zaobserwowano gwałtowny wzrost aspiracji edukacyjnych oraz zdobycia wykształcenia. Zmniejszył się udział osób z wykształceniem podstawowym, przy zwiększeniu odsetka osób posiadających wykształcenie średnie i wyższe³⁴⁵.

Kolejnym szczególnym osiągnięciem była możliwość realizacji podstawowych praw człowieka i demokratyzacji życia społecznego, o które naród polski od lat walczył³⁴⁶.

Biorąc pod uwagę negatywne i pozytywne aspekty zmian można uznać, że sfera społeczna najbardziej ucierpiała przy transformacji systemu³⁴⁷. Zjawiska, które dotknęły społeczeństwo (często w nierównym stopniu), zarówno w obszarze politycznym, jak i gospodarczym, doprowadziły do nieufności wobec polityków i całego porządku demokratycznego.

³⁴³ E. Mikuła, *Spoleczny wymiar transformacji*, Uniwersytet Rzeszowski, 2004, s. 265.

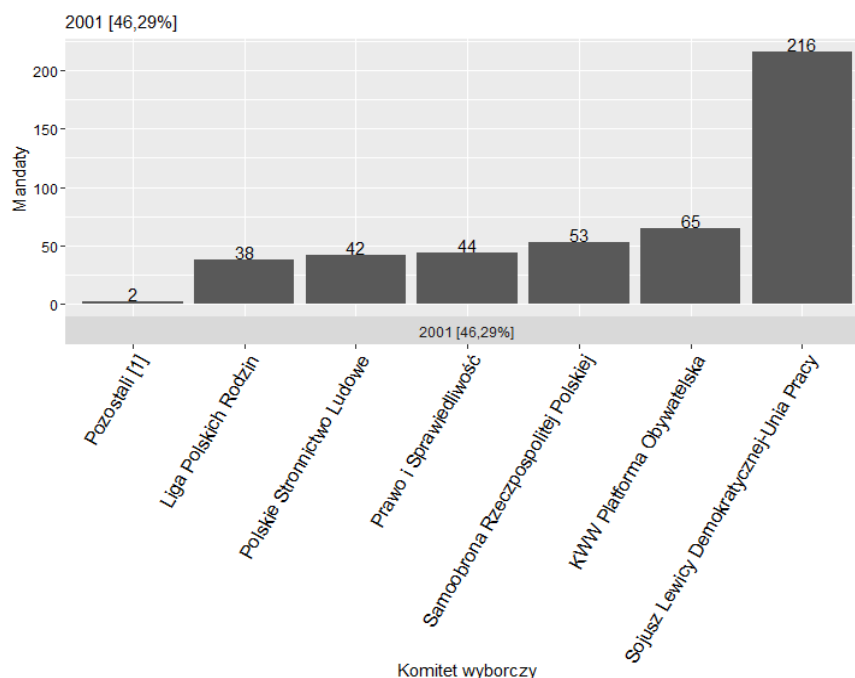
³⁴⁴ M. Dahl, *Wpływ transformacji ...*, dz. cyt.

³⁴⁵ M. Dahl, *Wpływ transformacji...*, dz. cyt., s. 41.

³⁴⁶ H. Skorowski, *Polskie przemiany w perspektywie politologiczno-socjologicznej*, „MAZOWSZE Studia Regionalne”, 2014, nr 15.

³⁴⁷ M. Dahl, *Wpływ transformacji...*, dz. cyt., s. 41.

Rysunek 29. Uzyskane mandaty do Sejmu w 2001[frekwencja].



Źródło: opracowanie własne na podstawie danych z *pkw.gov.pl* (20.11.22) z wykorzystaniem biblioteki *ggplot2*.

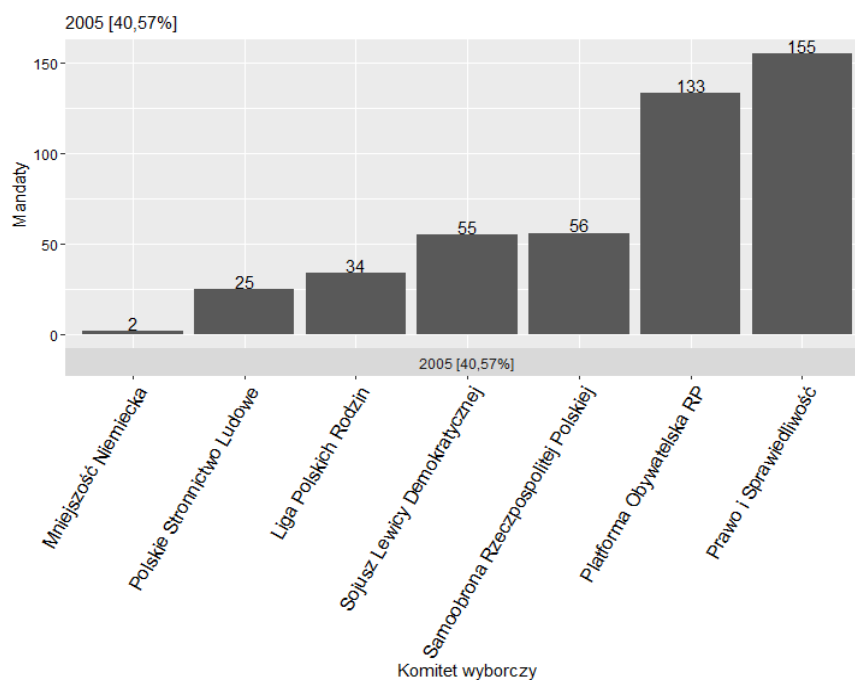
Po marginalizacji w 2000 r. i późniejszym upadku partii AWS oraz UW w 2001 r. powstały nowe organizmy: Platforma Obywatelska (PO) oraz Prawo i Sprawiedliwość (PiS). W wyborach parlamentarnych 23 września 2001 wygrała koalicja SLD i Unia Pracy (UP) uzyskując 41,04% głosów. Do parlamentu weszły Platforma Obywatelska (12,68%), Samoobrona (10,2%), Prawo i Sprawiedliwość (9,5%), Liga Polskich Rodzin (7,97%), Polskie Stronnictwo Ludowe (8,98%) Do parlamentu nie dostały się AWS (5,6%) i Unia Wolności (3,2%), które nie osiągnęły progu wyborczego.³⁴⁸ W 2001 r. został wprowadzony mechanizm subwencji – finansowania partii politycznych z budżetu państwa, przyczyniając się do wzmocnienia partii i po części wyrównania zasobów finansowych. Po 24 latach polskiej demokracji zmieniła się pozycja liderów niektórych partii i zarysowały pewne związane w tym procesy: pierwszy – wzmocnienie pozycji statutowych, drugi – utrzymywanie się lidera przez dłuższy okres, trzeci – silne utożsamianie partii z jej liderem³⁴⁹. Widoczne stało się zjawisko tzw. *ang. leader effect* rozumiane

³⁴⁸ A. Dudek, *Historia polityczna ...*, dz. cyt., s. 441-442.

³⁴⁹ A. Nyzio, *W słabości siła. Bilans ewolucji polskich partii politycznych i w latach 1989-2013*, „Polityka wewnętrzna i bezpieczeństwo”, 2014, nr 2, s. 204-208.

jako wpływ osoby na przywództwo, na poparcie dla jego partii lub głosowanie przeciwko osobie lidera konkurencyjnego podmiotu³⁵⁰. 7-8.06.2003 Polacy w referendum opowiedzieli się za wstąpieniem Polski do Unii Europejskiej. 1.05.2004 Rzeczpospolita Polska przystąpiła do Unii Europejskiej³⁵¹. Dzień później rząd Leszka Millera (SLD) podał się do dymisji. 13.06.2004 odbyły się pierwsze wybory do Parlamentu Europejskiego.

Rysunek 30. Uzyskane mandaty do Sejmu w 2005 [frekwencja].



Źródło: opracowanie własne na podstawie danych z pkw.gov.pl (20.11.22) z wykorzystaniem biblioteki ggplot2.

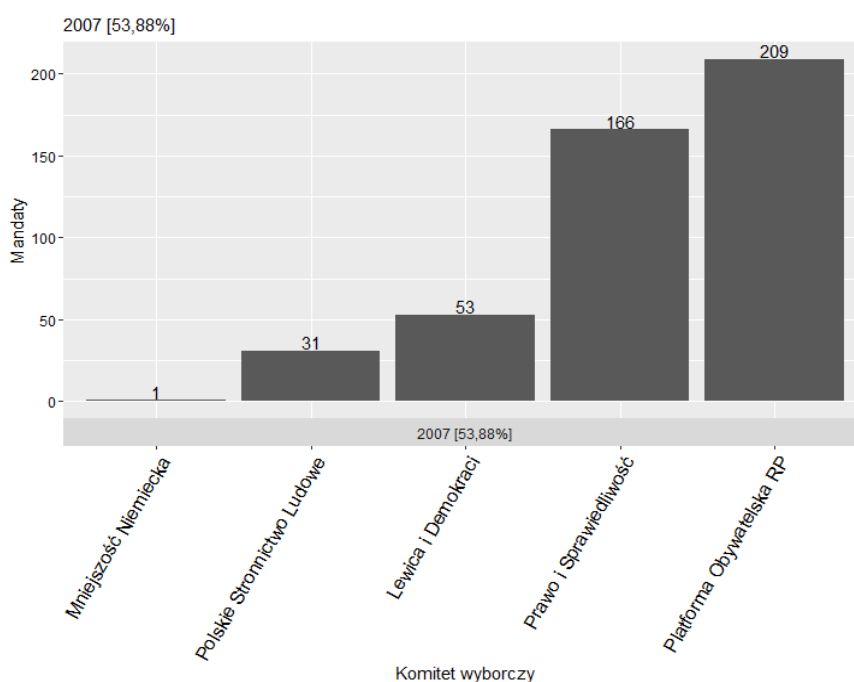
W wyborach parlamentarnych w 2005 r., wygrało Prawo i Sprawiedliwość uzyskując 26,99% głosów i 155 mandatów w Sejmie. Tuż za PiS znalazła się Platforma Obywatelska, która otrzymała 24,15% głosów (133 mandaty). W 2005 wybory prezydenckie wygrał Lech Kaczyński (PiS), pokonując w II turze Donalda Tuska (PO). 5.05.2006 roku powołano koalicję, a funkcje wicepremierów otrzymali politycy – Roman Giertych i Andrzej Lepper, uznawani za kontrowersyjnych. Koalicja rządząca rozpadła się i 7.09.2007 roku Sejm podjął decyzję o skróceniu kadencji³⁵².

³⁵⁰ W. Peszyński, *Personalizacja politycznych preferencji*, „Preferencje Polityczne. Postawy – identyfikacje – zachowania”, 2012, nr 3, s. 181-183.

³⁵¹ A. Dudek, *Historia polityczna...*, dz. cyt., s. 452.

³⁵² A. Dudek, *Historia polityczna...*, dz. cyt., s. 559, 590.

Rysunek 31. Uzyskane mandaty do Sejmu w 2007 [frekwencja].



Źródło: opracowanie własne na podstawie danych z pkw.gov.pl (20.11.22) z wykorzystaniem biblioteki ggplot2.

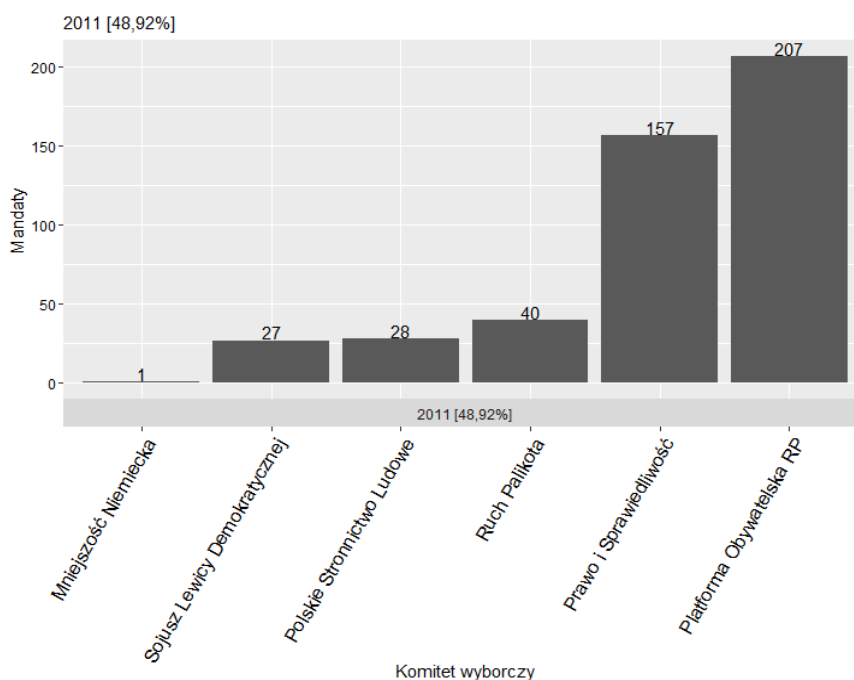
Wybory parlamentarne przeprowadzone 21.10.2007 wygrała PO, uzyskując 41,51% głosów, dało to 210 mandatów w Sejmie. Kolejne miejsce uzyskało PiS (32,11%), co przełożyło się na 166 mandatów poselskich. Na kolejnym miejscu uplasowała się koalicja Lewica i Demokraci (SLD, SDPL, PD, UP), zdobywając 13,2% głosów – 53 mandaty poselskie. Do Sejmu weszło również PSL, uzyskując 8,91% głosów – 31 mandatów. Do Sejmu nie weszły zaś współrządzące w poprzedniej kadencji LPR i Samoobrona, które nie przekroczyły progu wyborczego (5%)³⁵³. Partia PO utworzyła rząd w koalicji z PSL. Premierem rządu został D. Tusk, a wicepremierem i ministrem gospodarki W. Pawlak (PSL). W Senacie zaś bezwzględną większość zdobyła PO³⁵⁴. I tura przedterminowych wyborów prezydenckich była związana z katastrofą samolotu 10.04.2010 i śmiercią Prezydenta RP Lecha Kaczyńskiego wraz z 95 osobami. W wyborach prezydenckich w II turze zwyciężył Bronisław Komorowski uzyskując 53,01% głosów.³⁵⁵

³⁵³ A. Dudek, *Historia polityczna...*, dz. cyt., s. 598-599.

³⁵⁴ A. Dudek, *Historia polityczna...*, dz. cyt., s. 598-599.

³⁵⁵ A. Dudek, *Historia polityczna...*, dz. cyt., s. 628.

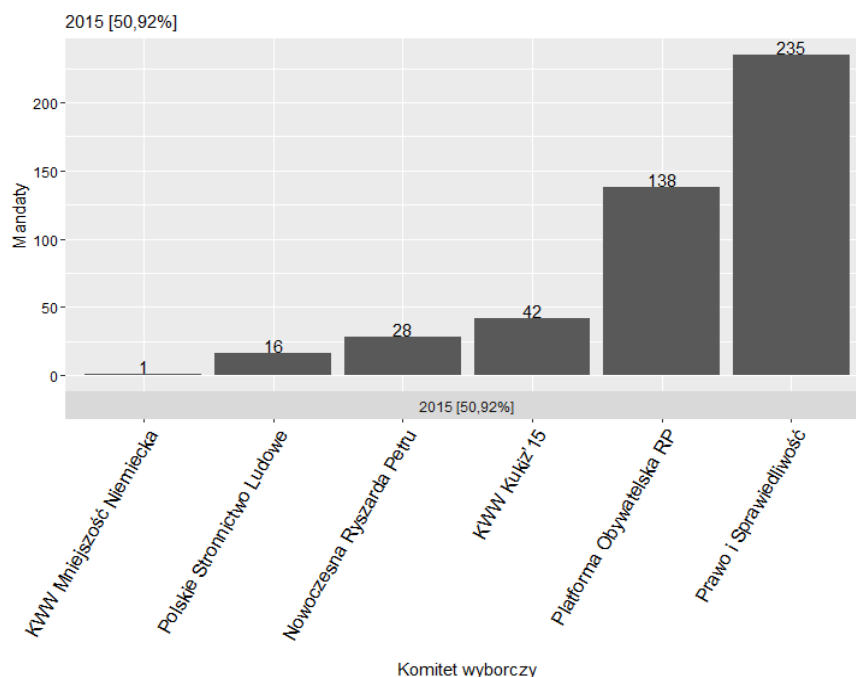
Rysunek 32. Uzyskane mandaty do Sejmu w 2011 [frekwencja].



Źródło: opracowanie własne na podstawie danych z pkw.gov.pl (20.11.22) z wykorzystaniem biblioteki ggplot2.

Wybory parlamentarne przeprowadzone 9.10.2011 wygrała PO z 39,18% głosami, co dało 207 mandatów. Drugie miejsce zajęło PiS z 29,89% głosów, co przyniosło 157 mandatów. Kolejne miejsca zajął Ruch Palikota zdobywając 10,02% głosów i 40 mandatów. Do Sejmu weszło również PSL z 28 mandatami uzyskując 8,36% głosów oraz SLD uzyskując 8,24% głosów, co przełożyło się na 27 mandatów. Partia PO utworzyła ponownie rząd w koalicji z PSL. Premierem nowego rządu ponownie został D. Tusk, a prezes PSL W. Pawlak znów został wybrany wicepremierem i ministrem gospodarki. W Senacie ponownie bezwzględną większość mandatów zdobyła PO. W wyborach prezydenckich w 2015 w II turze zwyciężył Andrzej Duda z obozu PiS uzyskując 51,55% głosów i wygrywając z urzędującym prezydentem B. Komorowskim.

Rysunek 33. Uzyskane mandaty do Sejmu w 2015 [frekwencja].

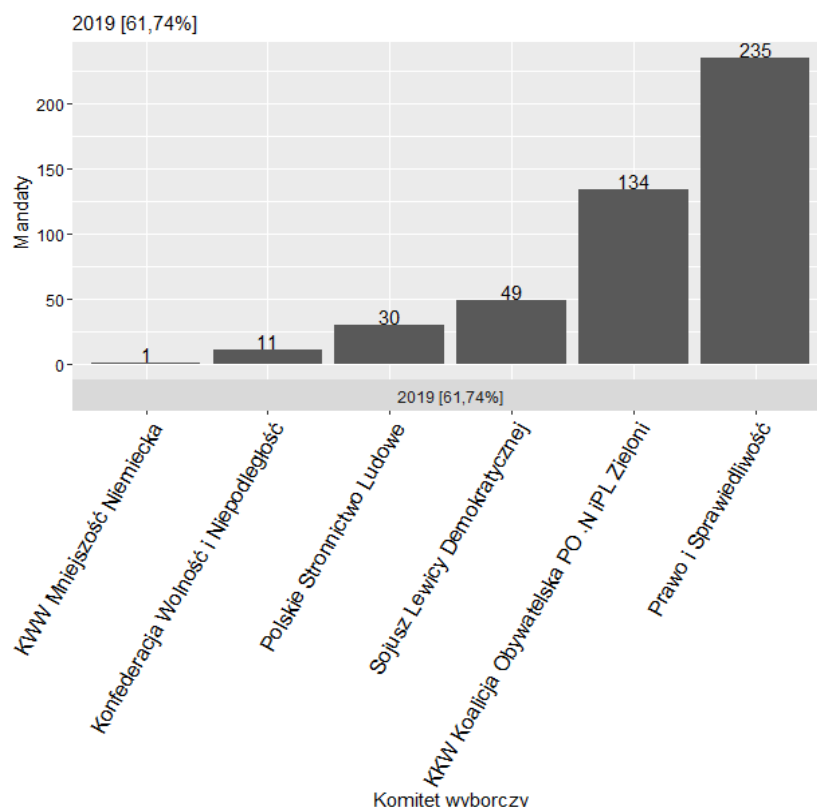


Źródło: opracowanie własne na podstawie danych z pkw.gov.pl (20.11.22) z wykorzystaniem biblioteki ggplot2.

Kolejne wybory parlamentarne zostały przeprowadzone 25.10.2015. Do Sejmu weszło 5 partii. Komitet Prawo i Sprawiedliwość uzyskał najwięcej głosów – 37,58% głosów (235 mandatów poselskich), a ponadto jako pierwszy obóz rządzący po 1989 roku zdobył większość umożliwiającą sformowanie samodzielnego rządu³⁵⁶. Do Sejmu dostały się także: Platforma Obywatelska, Kukiz'15, Nowoczesna i PSL. Poza parlamentem znalazła się Lewica. Premierem została Beata Szydło, która została zastąpiona w 2017 roku przez Mateusza Morawieckiego.

³⁵⁶ L. Garlicki, *Polskie prawo konstytucyjne. Zarys wykładu*, Wolters Kluwer, 2020, s. 37–39, 475–478.

Rysunek 34. Uzyskane mandaty do Sejmu w 2019 [frekwencja].



Źródło: opracowanie własne na podstawie danych z pkw.gov.pl (20.11.22) z wykorzystaniem biblioteki ggplot2.

Kolejne wybory parlamentarne zostały przeprowadzone 13.10.2019. Do Sejmu przeszło 5 partii. Zwycięskim ugrupowaniem okazało się Prawo i Sprawiedliwość, które wystawiło wspólną listę z koalicjantami w postaci Solidarnej Polski i Porozumienia, uzyskując 43,59% głosów (co daje 235 mandatów), po raz drugi zdobywając większość umożliwiającą sformowanie samodzielnego rządu. Do Sejmu dostały się również: Koalicja Obywatelska (składająca się z Platformy Obywatelskiej, Nowoczesnej, Inicjatywy Polskiej oraz partii Zieloni) (27,40%), Sojusz Lewicy Demokratycznej (składający się z Sojuszu Lewicy Demokratycznej, Wiosny oraz Lewicy Razem) (12,56%), Polskie Stronnictwo Ludowe (tworząc projekt polityczny Koalicja Polska, wraz z posłami z Kukiz 15') (8,55%) i Konfederacja Wolność i Niepodległość (KORWiN, Ruch Narodowy, Konfederacja Korony Polskiej) (6,81%). PiS nie zdobył większości w Senacie, pomimo

uzyskania największej liczby mandatów. Frekwencja wyniosła 61,74%.³⁵⁷ Do Sejmu nie weszły następujące komitety wyborcze: KWW Koalicja Bezpartyjni i Samorządowcy, Skuteczni Piotra Liroya-Marca, Akcja Zawiedzionych Emerytów Rencistów, Prawica. W wyborach prezydenckich w 2020 r. II turze najwięcej głosów (51,03%) otrzymał popierany przez PiS urzędujący prezydent A. Duda³⁵⁸.

Tabela 3. Uzyskane mandaty od 1991 oraz rozkład procentowy w Sejmie.

Komitet Wyborczy	Rok [frekwencja]	Mandaty	% mandatów
Unia Demokratyczna	1991 [43,2%]	62	13%
Sojusz Lewicy Demokratycznej		60	13%
Wyborcza Akcja Katolicka		49	11%
Polskie Stronnictwo Ludowe – Sojusz Programowy		48	10%
Konfederacja Polski Niepodległej		46	10%
Porozumienie Obywatelskie Centrum		44	10%
Pozostali [19]		43	9%
Kongres Liberalno-Demokratyczny		37	8%
„Porozumienie Ludowe”		28	6%
NSZZ „Solidarność”		27	6%
Polska Partia Przyjaciół Piwa		16	3%
Sojusz Lewicy Demokratycznej	1993 [52,13%]	171	37%
Polskie Stronnictwo Ludowe		132	29%
Unia Demokratyczna		74	16%
Unia Pracy		41	9%
Konfederacja Polski Niepodległej		22	5%
Bezpartyjny Blok Wspierania Reform		16	3%
Towarzystwo Społeczno-Kulturalne Mniejszości Niemieckiej na Śląsku Opolskim		3	1%
Towarzystwo Społeczno-Kulturalne Niemców Województwa Katowickiego		1	0%
Akcja Wyborcza Solidarność	1997 [47,93%]	201	44%
Sojusz Lewicy Demokratycznej		164	36%
Unia Wolności		60	13%
Polskie Stronnictwo Ludowe		27	6%
Ruch Odbudowy Polski		6	1%
Sojusz Lewicy Demokratycznej-Unia Pracy	2001 [46,29%]	216	47%
KWW Platforma Obywatelska		65	14%

³⁵⁷ Dziennik Ustaw Rzeczypospolitej Polskiej, <http://isap.sejm.gov.pl/isap.nsf/download.xsp/WDU20190001955/O/D20191955.pdf>, 01.09.2021.

³⁵⁸ Obwieszczenia PKW – Wybory Prezydenta Rzeczypospolitej Polskiej w 2020 r., <https://prezydent20200628.pkw.gov.pl/prezydent20200628/pl/obwieszczenia>, 01.09.2021

Samoobrona Rzeczypospolitej Polskiej		53	12%
Prawo i Sprawiedliwość		44	10%
Polskie Stronnictwo Ludowe		42	9%
Liga Polskich Rodzin		38	8%
Komitet Wyborczy Wyborców „Mniejszość Niemiecka”		2	0%
Prawo i Sprawiedliwość	2005 [40,57%]	155	34%
Platforma Obywatelska RP		133	29%
Samoobrona Rzeczypospolitej Polskiej		56	12%
Sojusz Lewicy Demokratycznej		55	12%
Liga Polskich Rodzin		34	7%
Polskie Stronnictwo Ludowe		25	5%
Mniejszość Niemiecka		2	0%
Platforma Obywatelska RP	2007 [53,88%]	209	45%
Prawo i Sprawiedliwość		166	36%
Lewica i Demokraci		53	12%
Polskie Stronnictwo Ludowe		31	7%
Mniejszość Niemiecka		1	0%
Platforma Obywatelska RP	2011 [48,92%]	207	45%
Prawo i Sprawiedliwość		157	34%
Ruch Palikota		40	9%
Polskie Stronnictwo Ludowe		28	6%
Sojusz Lewicy Demokratycznej		27	6%
Mniejszość Niemiecka		1	0%
Prawo i Sprawiedliwość	2015 [50,92%]	235	51%
Platforma Obywatelska RP		138	30%
KWW Kukiz'15		42	9%
Nowoczesna Ryszarda Petru		28	6%
Polskie Stronnictwo Ludowe		16	3%
KWW Mniejszość Niemiecka		1	0%
Prawo i Sprawiedliwość	2019 [61,74%]	235	51%
KKW Koalicja Obywatelska PO .N iPL Zieloni		134	29%
Sojusz Lewicy Demokratycznej		49	11%
Polskie Stronnictwo Ludowe		30	7%
Konfederacja Wolność i Niepodległość		11	2%
KWW Mniejszość Niemiecka		1	0%

Źródło: opracowanie własne na podstawie danych z pkw.gov.pl (20.11.22).

Wyniki wyborów do Sejmu IX kadencji pokazują, że PiS uzyskało bezwzględną większość. W największych miastach: Warszawa, Łódź, Gdańsk, Poznań i Szczecin – Koalicja Obywatelska (KO) uzyskała procentowo lepsze poparcie niż partia rządząca.

Tabela 4. Wyniki głosowania w skali okręgów w procentach – liczba mandatów [uzyskany % poparcia] w 2019.

Okręg	PiS	KO	SLD	PSL	Konfederacja	Inni	MN
Białystok	8 [52%]	3 [21%]	1 [9%]	1 [9%]	1 [7%]	0 [1%]	0 [0%]
Bielsko-Biała	1 [48%]	6 [27%]	2 [11%]	0 [6%]	0 [7%]	0 [2%]	0 [0%]
Bydgoszcz	5 [36%]	4 [31%]	2 [15%]	1 [9%]	0 [7%]	0 [1%]	0 [0%]
Chełm	8 [60%]	2 [15%]	1 [7%]	1 [12%]	0 [6%]	0 [1%]	0 [0%]
Częstochowa	4 [44%]	2 [23%]	1 [16%]	0 [9%]	0 [6%]	0 [3%]	0 [0%]
Elbląg	4 [41%]	2 [28%]	1 [12%]	1 [11%]	0 [6%]	0 [3%]	0 [0%]
Gdańsk	4 [32%]	6 [41%]	1 [13%]	0 [6%]	1 [7%]	0 [0%]	0 [0%]
Kalisz	6 [42%]	3 [25%]	2 [13%]	1 [13%]	0 [7%]	0 [0%]	0 [0%]
Katowice	13 [38%]	12 [33%]	4 [16%]	0 [5%]	1 [7%]	0 [3%]	0 [0%]
Kielce	1 [55%]	3 [17%]	1 [10%]	1 [10%]	1 [6%]	0 [1%]	0 [0%]
Konin	5 [47%]	2 [20%]	1 [15%]	1 [10%]	0 [7%]	0 [1%]	0 [0%]
Koszalin	3 [37%]	3 [32%]	1 [15%]	1 [9%]	0 [6%]	0 [0%]	0 [0%]
Kraków	12 [47%]	6 [27%]	2 [11%]	1 [8%]	1 [8%]	0 [1%]	0 [0%]
Krosno	8 [63%]	2 [16%]	0 [6%]	1 [8%]	0 [7%]	0 [0%]	0 [0%]
Legnica	6 [42%]	3 [25%]	2 [16%]	1 [7%]	0 [6%]	0 [3%]	0 [0%]
Lublin	9 [55%]	3 [19%]	1 [8%]	1 [9%]	1 [7%]	0 [1%]	0 [0%]
Łódź	4 [33%]	4 [36%]	2 [20%]	0 [5%]	0 [7%]	0 [0%]	0 [0%]
Nowy Sącz	8 [66%]	1 [14%]	0 [6%]	1 [7%]	0 [7%]	0 [0%]	0 [0%]
Olsztyn	5 [39%]	3 [26%]	1 [14%]	1 [13%]	0 [7%]	0 [1%]	0 [0%]
Opole	5 [38%]	4 [27%]	1 [12%]	1 [10%]	0 [6%]	0 [0%]	1 [8%]
Piła	4 [36%]	3 [31%]	1 [13%]	1 [14%]	0 [7%]	0 [0%]	0 [0%]
Piotrków Tryb.	6 [56%]	1 [16%]	1 [11%]	1 [10%]	0 [7%]	0 [0%]	0 [0%]
Płock	6 [52%]	2 [17%]	1 [9%]	1 [15%]	0 [5%]	0 [2%]	0 [0%]
Poznań	3 [25%]	5 [45%]	2 [16%]	0 [6%]	0 [7%]	0 [0%]	0 [0%]
Radom	6 [58%]	2 [17%]	0 [7%]	1 [10%]	0 [6%]	0 [1%]	0 [0%]
Rzeszów	1 [62%]	2 [14%]	1 [7%]	1 [8%]	1 [8%]	0 [1%]	0 [0%]
Siedlce	9 [60%]	2 [14%]	0 [6%]	1 [12%]	0 [6%]	0 [1%]	0 [0%]
Sieradz	7 [50%]	3 [20%]	1 [12%]	1 [10%]	0 [6%]	0 [2%]	0 [0%]
Słupsk	5 [36%]	5 [36%]	2 [12%]	1 [8%]	1 [7%]	0 [0%]	0 [0%]
Szczecin	4 [35%]	5 [36%]	2 [15%]	1 [7%]	0 [7%]	0 [0%]	0 [0%]
Tarnów	7 [60%]	1 [14%]	0 [6%]	1 [13%]	0 [7%]	0 [0%]	0 [0%]
Toruń	6 [40%]	4 [26%]	2 [15%]	1 [11%]	0 [6%]	0 [1%]	0 [0%]
Wałbrzych	4 [41%]	3 [32%]	1 [12%]	0 [7%]	0 [5%]	0 [2%]	0 [0%]
Warszawa	12 [34%]	13 [35%]	4 [16%]	2 [7%]	1 [7%]	0 [2%]	0 [0%]
Wrocław	5 [35%]	5 [33%]	2 [15%]	1 [6%]	1 [7%]	0 [3%]	0 [0%]
Zielona Góra	4 [34%]	4 [31%]	2 [16%]	1 [12%]	1 [7%]	0 [0%]	0 [0%]

Źródło: Opracowanie własne na podstawie: Dane dot. wyników do Sejmu 2019, <https://sejmse-nat2019.pkw.gov.pl/sejmsenat2019/pl/wyniki/sejm/pl>, 01.09.2021.

W polskim Sejmie przy podziale mandatów obowiązuje aktualnie metoda liczenia głosów metoda d'Hondta (mimo, że nie zawsze tak było po 1989 r.³⁵⁹). Konieczność jej stosowania wynika z faktu, że wg konstytucji wybory do Sejmu są wyborami proporcjonalnymi³⁶⁰. W przypadku Senatu obowiązuje system jednomandatowych okręgów wyborczych – mandat w danym okręgu zdobywa kandydat, który zebrał najwięcej głosów; 100 senatorów wybiera się w stu jednomandatowych okręgach, przy czym każdy komitet wyborczy może zgłosić jednego kandydata w danym okręgu³⁶¹. W przypadku wyborów do Sejmu sytuacja wygląda inaczej. Z 41 okręgów wyborczych wybiera się 460 posłów – kształt okręgów wyborczych reguluje Ustawa z dnia 5 stycznia 2011 r.³⁶² W okręgach wielomandatowych, liczba mandatów zależy od tego, ilu ludzi w nich mieszka³⁶³. Metoda d'Hondta pozwala na rozstrzygnięcie, do których komitetów trafią mandaty z wynikiem ułamkowym. Metoda d'Hondta, w swojej charakterystyce sprzyja większym partiom, a co za tym idzie zapobiega dużemu rozdrobnieniu partyjnemu – co jest dla niektórych partii zaletą a dla niektórych wadą³⁶⁴. Jednym z głównych postulatów Kukiz 15' jest zmiana ordynacji wyborczej na jednomandatowe okręgi wyborcze (JOWy)³⁶⁵. Przy podziale mandatów, biorą udział tylko kandydaci z komitetów wyborczych, przekraczających w skali całego kraju 5%-owy próg wyborczy oraz w przypadku koalicyjnych komitetów wyborczych, przekraczających 8%-owy próg wyborczy³⁶⁶. Głosując na wybranego kandydata głosujemy na listę (komitet wyborczy). Często wśród polityków ważnym aspektem jest startowanie z pierwszego miejsca na liście.

Po otrzymaniu mandatu zgodnie z regulaminami Sejmu mogą być tworzone frakcje polityczne jak: klub poselski – może być utworzony z co najmniej 15 posłów, lub koło

³⁵⁹ M. Rulka, *Proporcjonalne formuły wyborcze i ich konsekwencje*, „Studia PRAWNICZE”, 2010, nr 2(184).

³⁶⁰ Na czym polega metoda d'Hondta?, rp.pl, 20.11.2022.

³⁶¹ Wybory do Senatu 2019 podział mandatów, <https://forsal.pl/artykuly/1420851,wybory-do-senatu-2019-podzial-mandatow-jak-liczone-sa-glosy.html>, 20.11.2022.

³⁶² Kodeks wyborczy (Dz.U. z 2011 r. nr 21, poz. 113 z późn. zm.), <https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20110210112>, 17.11.2022.

³⁶³ Warszawa będzie mieć więcej posłów? PKW wzywa Sejm do korekt, <https://warszawa.wyborcza.pl/warszawa/7,54420,29066108,warszawa-bedzie-miec-wiecej-poslow-panstwowa-komisja-wyborcza.html?disableRedirects=true>, 18.11.2022.

³⁶⁴ M. Rulka, *Proporcjonalne formuły ...*, dz. cyt., s. 152.

³⁶⁵ Dlaczego warto wprowadzić JOW?, <http://www.kukiz15.org/program/krajowy/jow>, 18.11.2022.

³⁶⁶ Ile mandatów jest do zdobycia w okręgach wyborczych?, <https://www.rp.pl/polityka/art1083171-ile-mandatow-jest-do-zdobycia-w-okregach-wyborczych>, 19.11.2022.

poselskie (co najmniej 3 posłów)³⁶⁷. Posłowie z senatorami mogą tworzyć wspólne kluby, koła lub zespoły parlamentarne³⁶⁸. Zgodnie z regulaminem Sejmu parlamentarzysta może należeć tylko i wyłącznie do jednego klubu poselskiego lub koła poselskiego. Od 2001 r. partie polityczne, które osiągną zadowalający wynik w wyborach, są finansowane przez państwo za pomocą subwencji (jest to finansowa zapomoga, nieodpłatna i bezzwrotna dla ich działalności) oraz dotacji. Prawo do otrzymania subwencji ma partia, która otrzyma 3% poparcia w całym kraju, oraz koalicja jeśli uzyska 6% poparcia³⁶⁹. Partie polityczne są zobowiązane do sprawozdawczości na rzecz Państwowej Komisji Wyborczej (PKW) w związku z otrzymaniem środków z budżetu państwa, rokrocznie sporządzając sprawozdania zawierające informacje o otrzymanej subwencji oraz o poniesionych z subwencji wydatkach³⁷⁰.

Tabela 5. Zestawienie głównych charakterystyk komitetów wyborczych oraz głównych założeń programowych w 2019

Komitet wyborczy 2019	Charakterystyka i krótki opis postulatów
Koalicja Polska, (PSL, Kukiz'15)	Charakterystyka PSL: Agraryzm, Chrześcijańska demokracja, Kukiz' 15: Prawicowy populizm Demokracja bezpośrednia³⁷¹ PSL opowiada się za protekcjonizmem państwa w rolnictwie, sprzeciwia się aborcji, związkom cywilnym osób tej samej płci i eutanazji. Mimo konserwatywnego charakteru w kwestiach społecznych, w przeszłości tworzyła koalicje rządowe z centrolewicowym SLD i PO.³⁷² Głównym postulatem Kukiz' 15 jest doprowadzenie w Polsce do zmian ustrojowych poprzez wprowadzenie m.in.:

³⁶⁷ Regulamin Sejmu – tekst jednolity, <https://www.sejm.gov.pl/prawo/regulamin/kon7.htm>, 19.11.2022.

³⁶⁸ Ustawa o wykonywaniu mandatu posła i senatora, sejm.gov.pl, 19.11.2022.

³⁶⁹ Rozdział 4 ustawy z dnia 27 czerwca 1997 r. o partiach politycznych (Dz.U. Nr 98, poz. 604), https://pkw.gov.pl/uploaded_files/1498132393_ustawa_o_partiach_politycznych_-_wersja_aktualna.pdf, 08.09.2021.

³⁷⁰ Ł. Zamęcki, *Společne podstawy ładu politycznego*, Wydział Dziennikarstwa i Nauk Politycznych Uniwersytet Warszawski, 2011, s. 196-197.

³⁷¹ Parties and Elections, <http://www.parties-and-elections.eu/poland.html>, 08.09.2021.

³⁷² A short guide to the Polish political scene, <https://notesfrompoland.com/2019/11/13/a-short-guide-to-the-polish-political-scene>, 08.09.2021.

	ordynacji większościowej w formie jednomandatowych okręgów wyborczych. ³⁷³
PiS	Charakterystyka: Narodowy konserwatyzm W swoim programie wyborczym PiS obiecuje podniesienie płacy minimalnej do 4000 zł miesięcznie do 2023 roku oraz dodatkowych miesięcznych wypłat emerytur dla emerytów. Partia zobowiązuje się również do ochrony tradycyjnego modelu rodziny i wartości katolickich oraz kontynuowania reformy sądownictwa, które doprowadziły do konfliktu z Unią Europejską. Niektórzy członkowie partii obiecują także „repolonizację” i „odzyskanie kontroli” nad mediami po zwycięskich wyborach.³⁷⁴
Konfederacja	Charakterystyka: Narodowy konserwatyzm, Paleolibertarianizm, Eurosceptycyzm Konfederacja chce całkowicie zlikwidować podatek dochodowy, a także wprowadzić dobrowolność składek na ubezpieczenie społeczne ZUS. Jest przeciwna przyjęciu przez Polskę waluty euro, „ideologii LGBT i gender”. Zobowiązuje się również do liberalizacji dostępu do broni, ochrony Polski przed bez spadekowymi roszczeniami majątkowymi oraz wzmocnienia wartości chrześcijańskich i tradycyjnego modelu rodziny.³⁷⁵
Koalicja Obywatelska (KO)	Charakterystyka: Liberalny konserwatyzm Koalicja Obywatelska zobowiązuje się do przywrócenia demokracji i praworządności po rządach PiS, a także do dalszej integracji europejskiej. W swoim programie przedstawia obniżenie podatku dochodowego, a także wprowadzenie dodatkowego świadczenia finansowego dla osób otrzymujących płacę minimalną. Zobowiązuje się też do zlikwidowania wprowadzonego przez PiS zakazu handlu w niedzielę i przywrócić państwowe finansowanie procedur in vitro, które zostały zlikwidowane przez obecnie rządzącą partię. Pod względem energetycznym obiecuje wyeliminować węgiel do 2040 r., ma w planie zalegalizowanie związków cywilnych osób tej samej płci.³⁷⁶
SLD	Charakterystyka: Socjaldemokracja Sojusz Lewicy obiecuje liberalizację na froncie społecznym: poluzowanie prawa aborcyjnego w Polsce; poprawa dostępu do antykoncepcji i edukacji seksualnej;

³⁷³ Postulaty Kukiz '15, <https://pl.wikipedia.org/wiki/Kukiz%E2%80%9915>, 08.09.2021.

³⁷⁴ A short...dz. cyt., 08.09.2021.

³⁷⁵ A short...dz. cyt., 08.09.2021.

³⁷⁶ A short...dz. cyt., 08.09.2021

	wzmocnienie praw LGBT, jak również dążenie do wprowadzenia możliwości małżeństwa osób tej samej płci. Partia zapowiada również bezpłatne żłobki, bezpłatne posiłki dla dzieci w szkołach oraz bezpłatny transport publiczny dla osób starszych. ³⁷⁷
--	--

Źródło: Opracowanie własne na podstawie: *A short guide to the Polish political scen*, <https://notesfrom-poland.com/2019/11/13/a-short-guide-to-the-polish-political-scene; Parties and Elections>
<http://www.parties-and-elections.eu/poland.html>; wikipedia.org, 08.09.2021.

2.3. Rozwój mediów w Polsce po 1989 roku

W 1990 Sejm przyjął *Ustawę o uchyleniu ustawy o kontroli publikacji i widowisk, zniesieniu organów tej kontroli oraz zmianie ustawy Prawo prasowe* – popularnie ustawa nazywana była zniesieniem cenzury. Wolny rynek stanął otworem przed wieloma nowymi tytułami, m.in. *Gazetą Wyborczą*, która tę szansę wykorzystała. Przejęcia własnościowe po 1989 spowodowały wypełnienie wolnorynkowej niszy zagranicznymi koncernami, zaś regionalne, mniej wspierane pisma odchodziły w zapomnienie. Na rynku polskich mediów pojawiły się nowe medialne koncerny, stacje radiowe oraz telewizje. Z pewnością duży wpływ miała telewizja Polsat, która rozpoczęła nadawanie z końcem 1992 roku. Jako pierwsza telewizja również wprowadziła program typu *talk show* w zachodnim stylu (choć niektórzy badacze mediów za pierwszy program podobnego typu uznają *Tele Echo* z 1956 roku³⁷⁸). W 1996 r. zadebiutowała stacja TVN (program informacyjny *Fakty*, program rozrywkowy *Big Brother*) i również zaczęła wprowadzać nowe standardy w Polsce. W 1994 r. powołano Radio Zet. Prasa zaś po 1989 r. stała się popularna i ukształtowała się po likwidacji Robotniczej Spółdzielni Wydawniczej Prasa-Książka-Ruch. Powstało wiele dzienników m.in. *Super Express* po 1989, *Fakt* w 2003.³⁷⁹

Okres 1989-2000 możemy podzielić na pewne podokresy³⁸⁰:

- Faza I (okres od maja 1989 r. do połowy 1991 r.) – okres entuzjazmu nowych wydawców i wymuszonych przekształceń niektórych starych tytułów;

³⁷⁷ A short...dz. cyt., 08.09.2021

³⁷⁸ J. Kończak, *Od Tele-Echa do Polskiego Zoo. Ewolucja Programu TVP*, Warszawa 2008, s. 28

³⁷⁹ B. Czechowska-Derkacz, *Inforozrywka i tabloidyzacja w polskich mediach po 1989 roku w kontekście komunikacji politycznej*, [w:] J. Kreft, R. Stopikowski (red.), *Media, biznes, kultura*, Marpress, 2011, nr 2, s. 290–298.

³⁸⁰ R. Filas, *Dwadzieścia lat przemian polskich mediów (1989–2009) w ujęciu periodycznym*, Zeszyty prasoznawcze, Kraków 2010, s. 30.

- Faza II (okres od połowy 1991 r. do 1992 r.) – okres pozornej stabilizacji i podskórnych zmian w prasie oraz w radiu;
- Faza III (okres od 1993 r. do sierpnia 1994 r.) – okres walki o rynek mediów, szczególnie audiowizualnych;
- Faza IV (okres od września 1994 r. do 1996 r.) – okres zagospodarowania rynku po procesie koncesyjnym oraz szturmie tygodników niemieckich;
- Faza V (okres od 1997 do 2000 r.) – okres podziału rynku mediów i rozwijającej się specjalizacji.

Można wyodrębnić również kolejne fazy po roku 2000³⁸¹:

- Faza VI (okres od 2001 r. do 2003 r.) – okres rozwijającej dominacji mediów elektronicznych w okresie kryzysu ekonomicznego;
- Faza VII (okres od 2004 r. do 2007 r.) – okres przebudowy oferty tradycyjnych mediów wobec przewidywanej inwazji nowych technologii medialnych;
- Faza VIII (okres od 2008 do dziś³⁸²) – okres realnej konfrontacji mediów tradycyjnych i nowych technologii w okresie spowolnienia gospodarczego.

Prasa codzienna w fazach I-V doznała rewolucji w okresie między 1989 a 2000 r. Nastąpiła zmiana nawyków czytelnich, czytanych tytułów, znaczące zmniejszenie liczby dzienników regionalnych i koncentracja wydawców, zwykle z udziałem zagranicznego kapitału. Nakład gazet codziennych spadł z 7,5 mln egzemplarzy w 1989 r. do 3,9 mln w 1999 r. Faza VI była okresem stagnacji zarówno pod względem liczby tytułów, jak i łącznych nakładów. W fazie VII nastąpiła zmiana struktury oferty dzienników (kryzys płatnych regionalnych dzienników, rozkwit bezpłatnych gazet np. Metro). W fazie VIII trwał trudny okres dla dzienników zarówno ogólnopolskich, jak i regionalnych.³⁸³

Czasopisma w fazie I-V w porównaniu do dzienników codziennych przeżywały dynamiczny rozwój oferty i czytelnictwa. W latach 1992-1993 wprowadzano

³⁸¹ R. Filas, *Dwadzieścia lat ...*, dz. cyt., s. 37.

³⁸² Artykuł był opublikowany w 2010 r., stąd data końcowa jest nieznana.

³⁸³ R. Filas, *Dwadzieścia lat...*, dz. cyt., s. 30-50.

miesięczniki, od 1995 nastąpił rozwój tygodników lub dwutygodników. Często na licencjach zagranicznych wprowadzano kalki czasopism dużych domów wydawniczych. W fazie VI następowała koncentracja dużych domów wydawniczych. Liczba tytułów przekroczyła 6 tys. w 2003 roku. W fazie VII liczba ta doszła do 6,5 tys.. Główny przyrost odnotowały tytuły o niższej częstotliwości publikacji. W fazie VIII liczba tytułów przekroczyła 7 tys.³⁸⁴

Radiofonia w fazie I-V była reprezentowana od lat 90. przez monopol anten Polskiego Radia, przez nieład okresu przedkoncesyjnego (pirackie bądź na granicy legalności lokalne stacje), przez wprowadzenie od 1994 r. w ramy publicznego nadawcy 4 programowo zróżnicowanych anten ogólnokrajowych składających się z 17 regionalnych rozgłośni Polskiego Radia, do ogólnopolskich stacji komercyjnych jak *Radio Zet*, *RMF FM*, niekomercyjnego katolickiego *Radia Maryja* i setek innych stacji lokalnych. Faza VI była dla radia okresem odnowienia koncesji. W fazie VII słabła silna pozycja *Radia Zet* oraz *RMF FM* na korzyść rozgłośni lokalnych. W fazie VIII rosła liczba słuchających radia codziennie do 81,4% badanych w 2009 z 76,2% w 2007 r.³⁸⁵

Telewizja w fazach I-V podlegała podobnym przemianom, jak radio w Polsce w tym okresie. Od 1994 r. państwowa telewizja została przekształcona w telewizję publiczną: TVP1 oraz TVP2. Stopniowo przybywało nowych konkurentów, również docierających z zagranicy. W fazie VI Polacy stali się fanami oglądania telewizji – według badań TNS/OBOP było to średnio 225 min. dziennie. W okresie kryzysu ekonomicznego stacje rywalizowały między sobą. W fazie VII średni czas oglądania telewizji wzrósł do 241 min. dziennie. Nastąpił wzrost liczby sieci kablowych np.. *Vectra*, *Multimedia*, które często miały w ofercie pakiet z Internetem. W fazie VIII nastąpił rozwój telewizji cyfrowej – liczba abonentów platform cyfrowych wyniosła ponad 5 mln.³⁸⁶ W 2020 *Cyfrowy Polsat* miał nieco ponad 5 mln aktywnych usług płatnej telewizji w systemie abonamentowym cyfrowej, *Canal +* około 2,71 mln klientów i subskrybentów (od 2020 nie podaje

³⁸⁴ Tamże.

³⁸⁵ R. Filas, *Dwadzieścia lat...*, dz. cyt., s. 30-50.

³⁸⁶ Tamże.

tej liczby), *Orange* miało około 995 tys. klientów usług telewizyjnych³⁸⁷. Jeśli chodzi o telewizję cyfrową, prognozy, szczególnie na świecie, mogą spadać³⁸⁸.

Internet i nowe technologie komunikowania należą głównie do fazy V. W tym okresie rozpowszechniła się telefonia mobilna (GSM) – liczba abonentów z 0,8 mln użytkowników w 1997 roku wzrosła do 6,7 mln w roku 2000. Gwałtowny rozwój przeszedł Internet. W Polsce jako symboliczną datę jego powstania w komercyjnej wersji przyjmuje się 17.08.1991. Wraz z udostępnieniem narzędzi takich, jak protokół IP, strony WWW i przeglądarki (na początku *Internet Explorer*) zmieniał on swoją funkcję z wymiany poczty elektronicznej stając się platformą i środkiem komunikowania masowego. Pojawiły się pierwsze portale Wirtualna Polska (1995 r.), Onet.pl (1996 r.), Interia (1999 r.). Rozwijały się gazety w wersji elektronicznej: Gazeta.pl – e-wersja Gazety Wyborczej, Super Express (1996 r.), Rzeczpospolita (1997 r.).

Głównym problemem w rozwoju był ograniczony dostęp do Internetu. Internet był też powolny, głównie poprzez drogie impulsy wykorzystywane przez modemy telefoniczne. W fazie VI regulacje prawne pozwoliły na pojawienie się nowych dostawców Internetu. Coraz szybciej wkraczał on do polskich domów. W 2003 było 21,7% internautów. Równocześnie rozwijały się sieci telefonii komórkowej. W fazie VII rosła rola Internetu, w 2007 roku odsetek internautów sięgał już 41,5 %. Rozwijał się Internet szerokopasmowy. Następował rozwój komunikatorów jak *Gadu-Gadu*, *Skype* oraz portali społecznościowych, jak *Nasza klasa*, portali zakupowych jak *Allegro* oraz portali rozrywkowych. W fazie VIII odsetek internautów przekraczał 15 mln użytkowników w wieku 15+ (ponad 50%). Następował wyścig o szybkość dostarczania Internetu między dostawcami takimi, jak *Neotrada*, *Netia* i innymi. Rozwijał się Internet bezprzewodowy (mobilny).³⁸⁹

W 2020 r. według GUS ponad 90,4% gospodarstw domowych posiadało dostęp do Internetu³⁹⁰. Ze względu na okres pandemii część społeczeństwa musiała

³⁸⁷Do 2027 roku ubędzie 7 mln abonentów płatnej telewizji w Europie, <https://www.wirtualnemedia.pl/arttykul/do-2027-roku-ubedzie-7-mln-abonentow-platnej-telewizji-w-europie>, 22.03.2022.

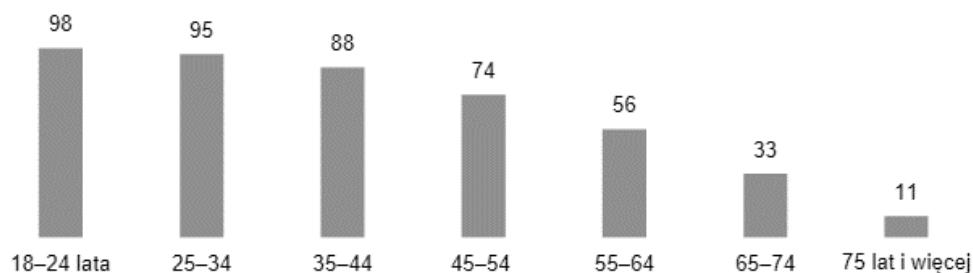
³⁸⁸ Tamże, 22.03.2022.

³⁸⁹ R. Filas, *Dwadzieścia lat...*, dz. cyt., s. 30-50.

³⁹⁰ M. Szutiak, Prawie 90% gospodarstw domowych w Polsce ma dostęp do szerokopasmowego Internetu - podaje GUS, <https://www.telepolis.pl/wiadomosci/prawo-finanse-statystyki/gus-xszerokopasmowy-internet-90-procent-gospodarstw-d-mowych-w-polsce>, 01.09.2021.

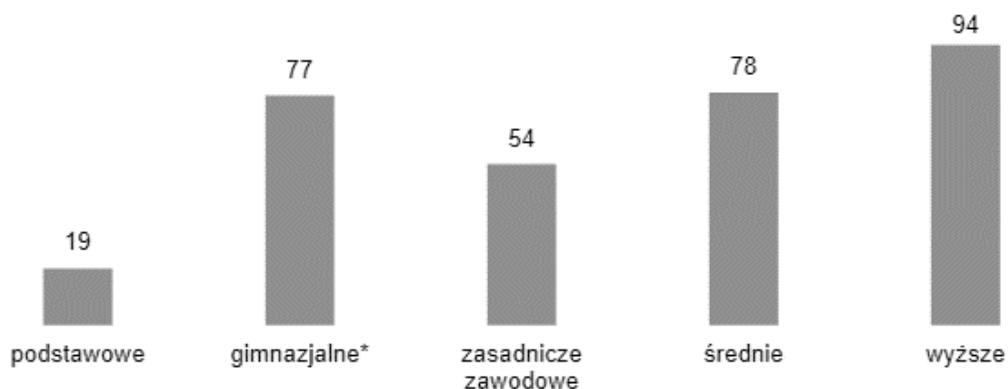
wykorzystywać rozwiązania internetowe. Według badań CBOS w 2020 r. blisko 98% korzystało z Internetu w przedziale wiekowym 18-24 lata i 95% w przedziale 25-34 lata. 94% użytkowników miało wykształcenie wyższe.

Rysunek 35. Odsetek korzystających z Internetu z podziałem na wiek w 2020 roku



Źródło: CBOS Korzystanie z Internetu, Komunikat z badań, Nr 85/2020, ISSN 2353-5822, 01.09.2021.

Rysunek 36. Odsetek korzystających z Internetu z podziałem na wykształcenie w 2020 roku



Źródło: CBOS Korzystanie z Internetu, Komunikat z badań, Nr 85/2020, ISSN 2353-5822, 01.09.2021.

W grupie 18-34 lat jest aktualnie blisko 100% korzystających z mobilnego, bezprzewodowego Internetu. Coraz chętniej Polacy kupują przez Internet – w 2020 roku to już 85% użytkowników Internetu.³⁹¹

³⁹¹ Tamże, 01.09.2021.

Patrząc na wykorzystywanie mediów w konkretnych celach, wartości przekraczające 50% użytkowników Internetu w 2020 roku dotyczyły czytania internetowych wersji gazet i czasopism 51% (spadek z 52% w IV 2018) oraz oglądania przez Internet telewizji, seriali i nagrań wideo (56% – podobna wartość w V 2019, ale można zauważyć, że wartości te od 2006 sukcesywnie rosły). Jeśli chodzi o czytanie gazet i czasopism w Internecie to wartości te od 2006 zawsze oscylowały w okolicy 50% użytkowników Internetu. Od 2018 sukcesywnie rośnie też liczba użytkowników oglądających transmisje realizowane i nadawane przez innych użytkowników na żywo. Dodatkowo od 2020 roku – w związku z pandemią internauci (25%) oglądali msze i nabożeństwa przez Internet.

Rysunek 37. Odsetek korzystających z Internetu z podziałem na cel w 2020 roku oraz wcześniej

Czy w ciągu ostatniego miesiąca:	Odsetki odpowiedzi twierdzących według terminów badań														
	III '06	III '07	III '08	VI '09	IV '10	VII '11	V '12	V '13	IV '14	V '15	V '16	IV '17	IV '18	V '19	VI '20
– czytał(a) Pan(i) internetowe wersje gazet, czasopism*															
Użytkownicy internetu	52	45	58	57	68	54	60	56	59	56	51	48	52	-	51
Ogół dorosłych	16	17	26	28	35	30	34	34	37	35	33	32	33	-	34
– słuchał(a) Pan(i) przez Internet radia															
Użytkownicy internetu	25	32	37	41	40	33	31	34	34	31	30	29	25	28	29
Ogół dorosłych	8	12	16	20	20	18	18	21	21	20	19	19	16	19	20
– oglądał(a) Pan(i) przez Internet telewizję, filmy, serie, nagrania wideo**															
Użytkownicy internetu	14	13	17	38	40	40	37	45	47	50	46	51	52	56	56
Ogół dorosłych	4	5	8	18	21	22	21	27	30	32	30	34	33	38	38
– czytał(a) Pan(i) blogi															
Użytkownicy internetu	-	-	-	-	39	28	29	30	30	34	33	31	30	29	33
Ogół dorosłych	-	-	-	-	20	15	17	18	19	21	22	21	19	20	22
– oglądał(a) Pan(i) w internecie transmisje realizowane i nadawane na żywo przez innych użytkowników															
Użytkownicy internetu	-	-	-	-	-	-	-	-	-	-	-	-	25	26	37
Ogół dorosłych	-	-	-	-	-	-	-	-	-	-	-	-	16	18	25
– oglądał(a) Pan(i) przez Internet transmisję mszy św. lub innego nabożeństwa															
Użytkownicy internetu	-	-	-	-	-	-	-	-	-	-	-	-	-	-	25
Ogół dorosłych	-	-	-	-	-	-	-	-	-	-	-	-	-	-	17

* W latach 2006 i 2007 użyliśmy sformułowania „internetowe wersje gazet codziennych”

** W latach 2006–2008 użyliśmy sformułowania „oglądał(a) Pan(i) przez Internet telewizję”

Źródło: CBOS Korzystanie z Internetu, Komunikat z badań, Nr 85/2020, ISSN 2353-5822, 01.09.2021.

Biorąc pod uwagę ostatni rok widać, że pod względem zamieszczania zdjęć i filmów, największą grupę stanowią użytkownicy w przedziale wiekowym 18-24 lata (56%). Nie ma widocznego związku między oglądaniem mszy i innych nabożeństw nadawanych na żywo od wieku – wartości oscylują w zakresie 21%-35%. Korelacja odwrotna z wiekiem ma miejsce w przypadku oglądania przez Internet telewizji, seriali i nagrań wideo – najwięcej (ponad 83% udziału w grupie 18-24 lata – wartość do grupy obejmującej 65+

spada wraz z wiekiem). Wartości oscylujące w przedziale 45%-57% internautów dotyczą czytania internetowych wersji gazet lub czasopism – Rysunek poniżej.

Rysunek 38. Odsetek korzystających z Internetu z podziałem na cel i wiek w 2020 roku

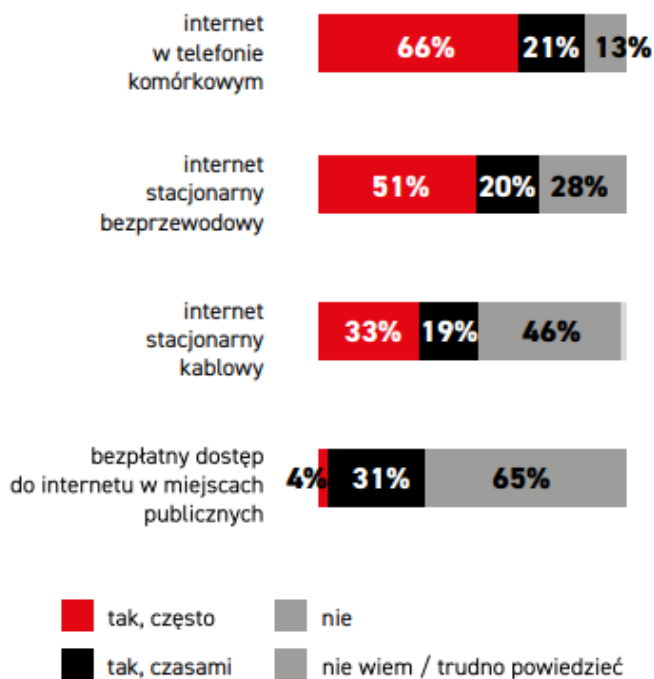
Wiek	Użytkownicy internetu, którzy w ciągu ostatniego miesiąca:						
	czytali internetowe wersje gazet, lub czasopism	słuchali przez internet radio	oglądali przez internet telewizję, filmy, seriale, nagrania wideo	czytali blogi	zamieszczali w internecie zrobione przez siebie zdjęcia, filmy	oglądali w internecie transmisje realizowane i nadawane na żywo przez innych użytkowników	oglądali przez internet transmisję mszy św. lub innego nabożeństwa
	w procentach						
18–24 lata	51	27	83	46	56	63	22
25–34	50	38	72	42	40	47	21
35–44	51	32	54	30	24	37	25
45–54	57	29	53	31	15	29	26
55–64	45	17	26	17	8	21	26
65 lat i więcej	51	19	32	26	19	19	35

Źródło CBOS Korzystanie z Internetu, Komunikat z badań, Nr 85/2020, ISSN 2353-5822, 01.09.2021.

Wśród Polaków korzystanie z Internetu w telefonie stanowi 87% ankietowanych wg badania ePolacy (raport z 2019³⁹²) z Internetu stacjonarnego bezprzewodowego korzysta 71% ankietowanych, a z Internetu kablowego 52% badanych. Z Internetu, który jest udostępniany bezpłatnie w miejscach publicznych 35%.

³⁹² Raport ePolacy 2019 , <https://www.gov.pl/documents>, 02.05.2022.

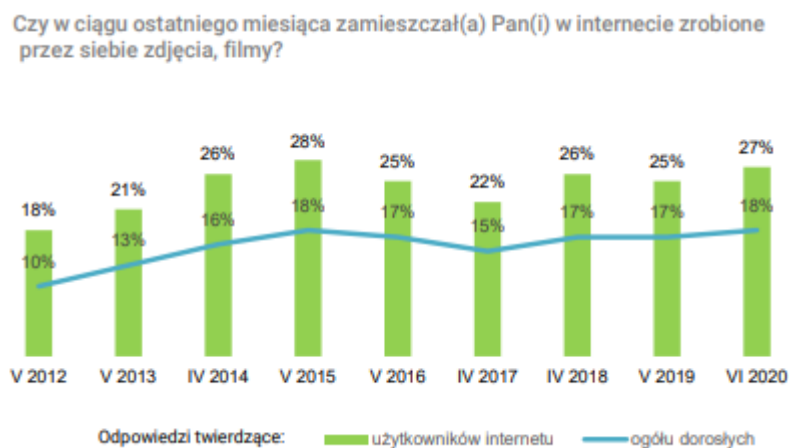
Rysunek 39. Najpopularniejszy wśród Polaków rodzaj dostępu do Internetu



Źródło ePolacy, <https://www.gov.pl/documents>, 02.05.2022.

Jeśli chodzi o odsetek internautów zamieszczających zdjęcia bądź filmy w Internecie w przeciągu ostatnich lat, w ostatnich dwóch latach pojawił się trend wzrostowy i aktualnie odsetek ten wynosi 27% użytkowników Internetu i 18% ogółu dorosłych.

Rysunek 40. Odsetek zamieszczających zdjęcia bądź filmy w Internecie w 2020 roku oraz wcześniej



Źródło: CBOS Korzystanie z Internetu, Komunikat z badań, Nr 85/2020, ISSN 2353-5822, 01.09.2021.

Analizując odsetek używających Internetu w celu komunikacji z innymi w 2020 roku oraz wcześniej, wzrost odsetka w 2020 roku widoczny jest w prawie każdym aspekcie. Ma to związek z pandemią i czasami niejako wymuszoną komunikacją w sposób internetowy – wysoki skok (o 13 pp.) można zauważyć wśród użytkowników Internetu, którzy przez ostatni rok prowadzili rozmowy telefoniczne wykorzystując Internet. Niektóre kanały komunikacji w ostatnim roku nie uległy znacznym zmianom (+/- 2 pp.) należą do nich: wpisy na formach, grupach dyskusyjnych, stronach serwisów społecznościowych (32% użytkowników Internetu) oraz granie w sieci z innymi (15% użytkowników Internetu).

Rysunek 41. Odsetek używających Internetu w celu komunikacji z innymi w 2020 roku oraz wcześniej

Czy w ciągu ostatniego miesiąca:	Odsetki odpowiedzi twierdzących według terminów badań															
	III '05	III '06	III '07	III '08	VI '09	IV '10	VII '11	V '12	V '13	IV '14	V '15	V '16	IV '17	IV '18	V '19	VI '20
rozmawiał(a) Pan(i) ze znajomymi przez komunikator, np. Messenger (Facebook), WhatsApp, Skype, Hangouts (Google), GG (Gadu-Gadu)																
Użytkownicy internetu	57	63	65	68	68	66	51	47	55	56	53	56	40	68	74	79
Ogół dorosłych	16	19	24	30	33	34	28	26	33	35	34	37	26	43	51	53
– prowadził(a) Pan(i) przez internet rozmowy telefoniczne (np. przez Skype)																
Użytkownicy internetu	-	27	25	38	41	39	35	36	33	39	35	29	27	31	34	47
Ogół dorosłych	-	8	9	17	20	20	19	20	20	25	22	19	18	20	24	32
– dokonywał(a) Pan(i) wpisów na forach, grupach dyskusyjnych, stronach serwisów społecznościowych																
Użytkownicy internetu	20	25	28	31	37	34	24	25	32	33	34	31	33	32	32	32
Ogół dorosłych	5	8	10	14	18	18	13	14	19	21	21	20	22	20	22	22
– grał(a) Pan(i) w coś w sieci z innymi ludźmi																
Użytkownicy internetu	-	-	-	-	28	21	17	14	20	20	16	12	12	14	16	15
Ogół dorosłych	-	-	-	-	14	11	9	8	12	13	10	8	8	9	11	10
– uczestniczył(a) Pan(i) w internetowej rozmowie wideo (wideokonferencji), w której udział brało kilka lub wiele osób?																
Użytkownicy internetu	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	31
Ogół dorosłych	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	21

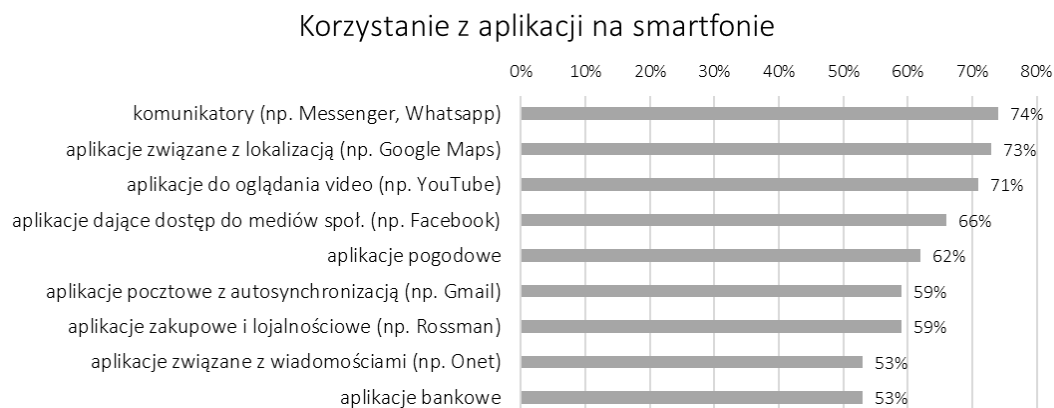
* W 2007 roku pytanie w ogóle nie zawierało przykładowych nazw komunikatorów. W pozostałych badaniach przykłady komunikatorów mogły się różnić

Źródło: CBOS Korzystanie z Internetu, Komunikat z badań, Nr 85/2020, ISSN 2353-5822, 01.09.2021.

Według raportu *ePolacy 2019* najwięcej ankietowanych korzysta ze smartfona korzystając z Internetu (61%), drugie urządzenie w kolejności to laptop (18%), a trzecie miejsce plasuje (z wynikiem 15%) komputer stacjonarny. Jeśli chodzi aplikacje to pierwsze miejsce w najczęściej używanych to komunikatory (74%), drugie miejsce stanowią aplikacje związane z lokalizacją (73%) oraz trzecie miejsce z wynikiem (71%) plasują

aplikacje do oglądania video. Co ciekawe kolejną pozycję zajmują aplikacje dające dostęp do mediów społecznościowych (66%).³⁹³

Rysunek 42. Odsetek używających aplikacji na smartfonie Internetu



Źródło: opracowanie własne na podstawie Raportu ePolacy 2019, <https://www.gov.pl/documents,02.05.2022>.

Patrząc na sposób komunikacji w ostatnim roku, z podziałem na wiek, widzimy niższe wartości w przypadku starszych osób w każdym typie komunikacji. Najmniejszy udział użytkowników Internetu w przedziale powyżej 65 lat (7%) widzimy w wykorzystywaniu Internetu do grania w Sieci – najwyższą wartością (54%) w wykorzystywaniu Internetu do rozmów przez komunikator internetowy. W przypadku osób najmłodszych w grupie 18-24 lata i 25-34 odsetek rozmawiających przez komunikator wynosi kolejno 95% i 93% użytkowników Internetu – mimo tego że odsetek spada dla każdej wyższej grupy wiekowej to nie spada nigdy poniżej 54%, co pokazuje że głównym kanałem komunikacji jest wykorzystywanie komunikatorów internetowych do rozmowy z innymi.

³⁹³ Raport ePolacy..., dz. cyt., 02.05.2022

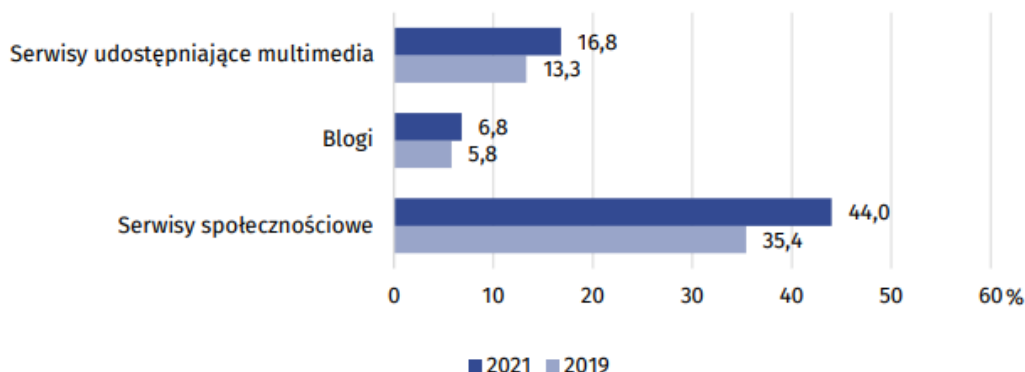
Rysunek 43. Odsetek używających Internetu w celu komunikacji z innymi w podziale na wiek w 2020 roku

Wiek	Użytkownicy internetu, którzy w ciągu ostatniego miesiąca:				
	rozmawiali przez komunikator internetowy	przewodili rozmowy telefoniczne przez internet	dokonywali wpisów na forach, w serwisach społecznościowych	uczestniczyli w internetowej rozmowie wideo (wideo konferencji), w której udział brało kilka lub wiele osób	grali w sieci
	w procentach				
18–24 lata	95	63	48	55	45
25–34	93	58	41	39	21
35–44	82	46	35	31	10
45–54	72	38	28	26	7
55–64	62	36	14	15	3
65 lat i więcej	54	32	19	14	7

Źródło: CBOS Korzystanie z Internetu, Komunikat z badań, Nr 85/2020, ISSN 2353-5822, 01.09.2021.

Jeśli chodzi o przedsiębiorstwa w Polsce w 2021 r. wg danych z GUS blisko co drugie przedsiębiorstwo korzystało z mediów społecznościowych (44%), blisko 17% korzystało z serwisów udostępniających multimedia, a około 7 % korzystało z blogów³⁹⁴.

Rysunek 44. Przedsiębiorstwa korzystające z mediów społecznościowych według rodzajów (w % ogółu przedsiębiorstw)

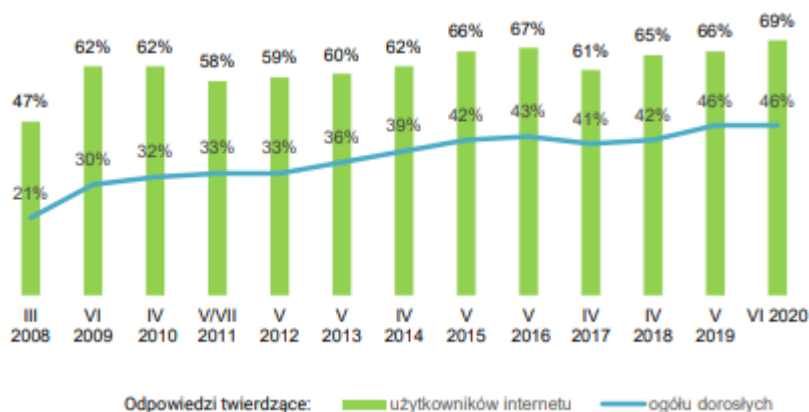


Źródło: GUS, Raport Społeczeństwo informacyjne w Polsce w 2021, 06.06.2022

³⁹⁴ GUS, Raport Społeczeństwo informacyjne w Polsce w 2021, 06.06.2022

W ciągu kilku lat widać, że Polacy chętniej zakładają również konta w serwisach społecznościowych w połowie 2020 wartość oscyluje w granicach 69% użytkowników Internetu i 46% ogółu dorosłych.

Rysunek 45. Odsetek posiadających konto w serwisach społecznościowych w 2020 roku oraz wcześniej

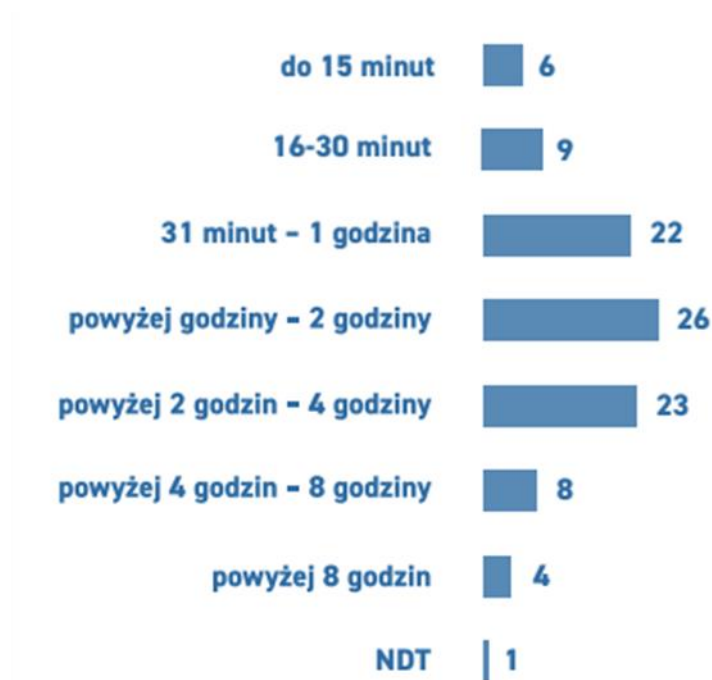


Źródło: CBOS Korzystanie z Internetu, Komunikat z badań, Nr 85/2020, ISSN 2353-5822, 01.09.2021.

Warto zwrócić na czas spędzany w Internecie wg Raportu *ePolacy 2019*, najczęściej korzystamy z Internetu pomiędzy 30 min, a 4 godzin dziennie (łącznie ponad 71% ankietowanych), co wg twórców raportu jest uznane jako *Za długo*. Jak podają autorzy mimo wszystko czas aktywnego korzystania z Internetu jest trudny do oszacowania, bo ludzie łączą się z Internetem wielokrotnie w ciągu dnia, na krótko, nawet nie myśląc, że korzystają właśnie z Internetu³⁹⁵.

³⁹⁵ Raport ePolacy..., dz. cyt., 02.05.2022

Rysunek 46. Procent korzystających biorąc pod uwagę czas poświęcany na korzystaniu z Internetu



Źródło Raport ePolacy 2019, <https://www.gov.pl/documents>, 02.05.2022

Podsumowując, osoba korzystająca z Internetu z wykorzystaniem smartfona będzie miała konto w serwisie społecznościowym (około 69% osób korzystających z Sieci ma konto), zwykle osoba komunikująca się w Internecie z pośrednictwem komunikatorów będzie w przedziale 18-34 lata (wartości 93-95% użytkowników Internetu) oraz w podobnej grupie wiekowej będzie używała Internetu w celu wypowiadała się w serwisie społecznościowym (w około 41-48% udziału osób korzystających z Internetu).

2.4. Przemiany komunikacji politycznej pod wpływem mediów cyfrowych i rola mediów cyfrowych

Na świecie za początek wykorzystania Internetu w kampaniach wyborczych uznaje się rok 1995, gdy w USA jeden z kandydatów na urząd prezydenta Lamar Alexander ogłosił kandydaturę na informacyjnej tablicy w serwisie *America Online*. W 1996 Internet został użyty jako narzędzie marketingu – za pioniera można uznać kandydata Patricka Buchanana, który w prawyborach stworzył swoją stronę WWW. W 1996 r. podczas

debaty prezydenckiej Bob Doyle ogłosił swoją kampanię w Sieci prosząc o wejście na jego stronę i otrzymał 2 mln wejść bezpośrednio po debacie.³⁹⁶

W Polsce rozwój wykorzystania Internetu w polityce postępował w podobny sposób. W 1997 r. partie uczestniczące w wyborach do parlamentu miały swoje strony internetowe – mimo ubogiej grafiki pełniły funkcję głównie broszur o ograniczonej treści³⁹⁷. Wybory prezydenckie w 2000 roku były kolejnym impulsem w rozwijaniu internetowej komunikacji politycznej³⁹⁸. W okresie tych wyborów pod względem treści i zawartości było dużo lepiej: umieszczano zdjęcia, pliki wideo z konferencji, spoty oraz potrzebne adresy kontaktowe. Wszyscy kandydaci (13) posiadali swoje strony WWW. Jednym z czynników prowadzących do personalizacji polityki jest jej mediatyzacja³⁹⁹. W 2001 w okresie wyborów parlamentarnych strony WWW partii promowały twarze swoich liderów budując wizerunek⁴⁰⁰. W latach 2002-2003 w czasie wyborów samorządowych i europejskich partie pracowały nad marketingiem politycznym. W 2005 roku odbyły się w Polsce wybory parlamentarne i prezydenckie, w których również zastosowano innowacyjne w Polsce i zapożyczone z zachodu rozwiązania internetowe w polityce np. na stronie partii PiS pojawił się „kalkulator podatkowy” umożliwiający obliczenie oszczędności w wydatkach w propozycjach PiS i PO⁴⁰¹. Każda z partii wykorzystywała pocztę elektroniczną. W 2007 zaczęto także wykorzystywać reklamy wysyłane na pocztę zwane *direct-mail*⁴⁰². W 2011 pręźnie korzystano z rozwiązań marketingu internetowego – największe ówczesne partie (PiS i PO) używały serwisów społecznościowych, umieszczały pliki wideo, strony były przystosowane na urządzenia mobilne i dla osób niedowidzących, korzystano z doświadczeń kampanii z USA⁴⁰³.

³⁹⁶ A. Falkowski, W. Cwalina, *Marketing polityczny...*, dz. cyt., s. 443-444.

³⁹⁷ A. K. Piasecki, *Wybory. Parlamentarne, samorządowe, prezydenckie 1989–2002*, Wydawnictwo Adam Marszałek, Toruń 2004, s. 76. ; J. Żurawski, *Internet jako współczesny środek elektronicznej komunikacji wyborczej*, Księgarnia Akademicka, Kraków 2010, s. 124-125.

³⁹⁸ J. Żurawski, *Internet jako współczesny środek elektronicznej komunikacji wyborczej*, 2010, s. 124-125.

³⁹⁹ J. Żurawski, *Internet jako...*, dz. cyt., s. 120-121.

⁴⁰⁰ T. Gajowniczek, *Internet w komunikowaniu politycznym*, Wydział Nauk Społecznych Uniwersytet Warmińsko-Mazurski w Olsztynie, 2019, s. 57.

⁴⁰¹ T. Gajowniczek, *Internet w...*, dz. cyt., s. 57.

⁴⁰² Tamże.

⁴⁰³ Tamże, s. 58.

Aktualnie partie polityczne i politycy są aktywni w Internecie, a rola Internetu w polityce i kampaniach rośnie⁴⁰⁴. Korzystanie z mediów społecznościowych, jak Facebook czy Twitter przez polityków jest czymś naturalnym. Trend wzrostowy jest również widoczny w instytucjach publicznych. Jedni chwalą takie rozwiązania widząc w tym rozwój współczesnego komunikowania z obywatelami, drudzy zaś widzą zagrożenia w dyskryminacji osób nie mających kont w mediach społecznościowych⁴⁰⁵. Internet i media społecznościowe są również wykorzystywane w promocji miast, czy gmin. Urzędy korzystają z kanałów dystrybucji marketingu zintegrowanego (*ang. inbound marketing*), aby promować miejscowości czy regiony wśród osób czy potencjalnych turystów nie mieszkających tam – stąd wykorzystanie marketingu internetowego⁴⁰⁶. Z zaangażowaniem i mediatyzacją polityki wiąże się termin *e-partycypacja*. Według J. van Dijka jest to zjawisko zaangażowania obywateli w życie politycznym wykorzystując ICT (*ang. Information and Communication Technologies*).⁴⁰⁷

W Polsce żadne z medium nie rozwinęło się tak szybko jak Internet, w ciągu 10 lat, stając się globalnym narzędziem do przekazywania i odbierania informacji⁴⁰⁸. Andrzej Falkowski i Wojciech Cwalina przedstawiają ważne czynniki, które przyspieszyły wykorzystanie Internetu do celów komunikacji politycznej⁴⁰⁹. Pierwszym z nich jest fakt, że działanie w Sieci może generować odzew w realnym świecie – *poszukiwanie nowych sposobów komunikowania się z wyborcami zostało niejako wymuszone przez zmiany następujące w populacji osób uprawnionych do głosowania*⁴¹⁰. Dodatkowo obecni i przyszli wyborcy oraz ich przyzwyczajenia w wykorzystaniu Internetu jako źródła komunikacji mogą tę dynamikę zwiększyć⁴¹¹. Drugi czynnik wiąże się z sukcesem ich wykorzystania w branży handlowej i promocyjnej, będącej często wzorcem dla polityków⁴¹². Trzecim

⁴⁰⁴ J. N. Kocks, *Political Media Relations Online as an Elite Phenomenon*, Springer, Wiesbaden 2016, s. 65.

⁴⁰⁵ T. Gajowniczek, *Internet w ...*, dz. cyt., s. 60.

⁴⁰⁶ A. Raszkowski, *Przesłanki promocji miast i regionów*. „Zeszyty Naukowe Uniwersytetu Szczecińskiego. Ekonomiczne Problemy Usług”, 2010, nr 56, s. 377-387.

⁴⁰⁷ J. van Dijk, *Models of Democracy and Concepts of Communication*, [w:] K. Hacker, J. van Dijk (red.), *Digital Democracy: Issues of Theory and Practice*, Sage, London – Thousand Oaks – New Delhi 2004, s. 30-53.

⁴⁰⁸ O. Annusewicz, *Funkcje komunikowania politycznego w Internecie*, „Studia politologiczne”, 2020, nr. 14.

⁴⁰⁹ A. Falkowski, W. Cwalina, *Marketing polityczny. Perspektywa psychologiczna*, Gdańsk 2005.

⁴¹⁰ A. Falkowski, W. Cwalina, *Marketing polityczny...*, dz. cyt., s. 439.

⁴¹¹ O. Annusewicz, *Funkcje komunikowania...*, dz. cyt., s. 269.

⁴¹² O. Annusewicz, *Marketing polityczny – pytania i kontrowersje* [w:] E. Pietrzyk-Zieniewicz (red.), *Scena polityczna i media, Miraże sukcesu – ryzyko autoprezentacji*, Warszawa 2004, s. 104.

czynnikiem jest wzrost osób posiadających dostęp do Internetu oraz firm informatycznych, które takie rozwiązania informatyczne oraz nowe technologie wprowadzają⁴¹³.

Według Michaela Cornfielda media interaktywne stanowią zagrożenie dla demokratycznych procesów – dostarczają wyborcom *faux population*, czyli zjawiska odczuwania dobrego samopoczucia i zadowolenia wraz z brakiem odpowiedzialności uczestniczenia w życiu politycznym⁴¹⁴. Obywatele mogą przestać uczestniczyć w realnym działaniu na rzecz partii, stwarzając jedynie iluzję partycypacji w zapleczu politycznym.

O. Annusewicz wyróżnia sześć funkcji komunikowania politycznego w Internecie:

- *Funkcja informacyjno-autoprezentacyjna*, mówi o tym, że politycy, kandydaci starają się przekazywać wyborcom informacje o swoim programie i dodatkowo internetowe instrumenty takie jak własne strony WWW pełnią funkcję autoprezentacyjną;
- *Funkcja perswazyjna*, mówi o tym, że informacje mogą również być nacechowane w emocjonalny sposób z intencją wzbudzenia określonego wrażenia w odbiorcy przekazu;
- *Funkcja edukacyjna*, mówi o tym, że Internet i miejsca do tego przeznaczone, jak strony internetowe i profile w mediach, mogą posłużyć do publikowania treści edukacyjnych – artykułów i informacji;
- *Funkcja integracyjna*, strona, portal czy forum mogą być „miejscem spotkań”, gdzie kandydaci mogą porozmawiać ze swoimi zwolennikami online z wykorzystaniem oprogramowania umożliwiającego przesyłanie obrazu i dźwięku;
- *Funkcja komunikowania dwukierunkowego*, Internet umożliwia komunikowanie zwrotne nie tylko wobec politycznych zwolenników, ale też przeciwników. Niewłaściwy *feedback* polityczny może spotkać się z ostracyzmem na forum większej społeczności i nasilić negatywny elektorat⁴¹⁵;

⁴¹³ A. Falkowski, W. Cwalina, *Marketing polityczny...*, dz. cyt., s. 441-442.

⁴¹⁴ A. Falkowski, W. Cwalina, *Marketing polityczny...*, dz. cyt., s. 438.

⁴¹⁵ P. Gotkowski, *Szeptanie w komunikacji marketingowej*, „Marketing w Praktyce”, 2009, nr 42, s.

- *Funkcja ekonomiczna*, mówi o tym że poprzez Internet możemy pozyskać współpracę biznesową, bądź fundusze na kampanie wyborcze.⁴¹⁶

W badaniu Żaneta Wełna pokazuje, że liderzy poszczególnych partii politycznych w wyborach parlamentarnych w 2019 (Władysław Kosiniak-Kamysz (PSL), Mateusz Morawiecki (PiS), Włodzimierz Czarzasty (SLD), Grzegorz Schetyna (KO), Janusz Korwin-Mikke (Konfederacja)) w mediach społecznościowych kładli nacisk na *funkcję autoprezentacyjną w połączeniu z funkcją perswazyjną*⁴¹⁷. Wyjątek stanowił lider KO – w tym przypadku posty w większym stopniu stawiały na informowanie. Przeprowadzone przez autorkę badanie: *potwierdza zatem obecną w literaturze tezę, że politycy, dostrzegając potencjał komunikacyjny i wizerunkowy mediów społecznościowych, wiodącym celem tej formy komunikacji czynią pozyskiwanie przez siebie elektoratu*⁴¹⁸

Agencja Grayling zaprezentowała europejskie badanie na temat sposobu używania mediów społecznościowych przez polityków w poszczególnych państwach Europy (17 krajów). Na tle wartości średniej liczby opublikowanych postów na osobę Polska (607 postów na osobę) zajmuje 2 miejsce ustępując miejsca Hiszpanii (690 postów na osobę). Biorąc pod uwagę ogół społeczeństwa, we wskaźniku procenta populacji stanowiącego liczbę użytkowników mediów społecznościowych zajmujemy 12 z 17 pozycji. Na podstawie badania można stwierdzić również, że politycy o skrajnych poglądach są bardziej aktywni na wszystkich najpopularniejszych platformach social media w skali Europy. Posty skrajnej prawicy osiągają wysokie wskaźniki zaangażowania wśród odbiorców na Twitterze, podczas gdy zarówno skrajna prawica i lewica mają najniższe wskaźniki zaangażowania na platformie *Instagram*.⁴¹⁹ Mimo tego, że Twitter wydaje się być bardziej preferowanym narzędziem w komunikacji politycznej, to więcej postów posiada konto na platformie Facebook (444 osób) niż na platformie Twitter (372 osób). Co ciekawe, mimo to, liczba publikowanych przez nich treści na Twitterze jest ponad dwa razy wyższa niż na platformie Facebook – (Facebook generuje 31% treści, Twitter 66% treści). Udział platformy Instagram w komunikacji polityków pozostaje marginalny – na 257 postów

⁴¹⁶ O. Annusewicz, *Funkcje komunikowania...*, dz. cyt.

⁴¹⁷ Ż. Wełna, *Polityk aktywny – polityk interaktywny. Obecność liderów politycznych w mediach społecznościowych na przykładzie polskiej kampanii parlamentarnej 2019*, "MEDIA I SPOŁECZEŃSTWO", 2020, nr 12.

⁴¹⁸ Ż. Wełna, *Polityk aktywny ...*, dz. cyt., s. 201.

⁴¹⁹ Grayling's report ..., dz. cyt., 20.04.2022.

posiadających konto na platformie Instagram, ich posty stanowią jedynie 2,7 % wszystkich treści stworzonych przez posłów w social media.

Trzy czynniki, które przyspieszyły wykorzystanie Internetu do celów komunikacji politycznej: (I) działanie w Sieci może generować odzew w realnym świecie. (II) Sukces wykorzystania Internetu w branży handlowej i promocyjnej, jest często wzorcem dla polityków. (III) Wzrost odsetka osób posiadających dostęp do Internetu oraz firm informatycznych, które dostarczają rozwiązania informatyczne⁴²⁰.

Na dzień 15.06.2022 dane na temat oficjalnych partyjnych kont oraz liczby obserwujących przedstawione partie polityczne znajdujące się w Sejmie są w tabeli poniżej:

Tabela 6. Dane dotyczące oficjalnych profili partii/koła Sejmu IX kadencji i liczby obserwujących

Partia/Koło	Twitter	Facebook	Youtube
Prawo i Sprawiedliwość	277 tys. (@pisorgpl)	324 tys. (/pisorgpl/)	13 tys. (/pisorgpl/)
Koalicja Obywatelska/Platforma Obywatelska	107,2 tys. (@Platforma_org)	325 tys. (/PlatformaObywatelska/)	8,92 tys. (/channel/UCT7iDAFxSIN1UHw86s42bFg)
Lewica	111 tys. (@__Lewica)	146 tys. (/NowaaLewica/)	4,37 tys. (/TVSLD)
Koalicja polska=> PSL	36 tys. (@nowePSL)	50 tys. (/nowePSL/)	Poniżej 1 tys. (nowePSL)
Konfederacja	151 tys. (@KONFEDERACJA_)	5,4 tys. (wcześniejszy profil posiadający ponad 670 tys. został usunięty 05.01.22 przez FB przez naruszenie regulaminu) (people/Konfederacja/100077050784448/)	178 tys. (/UCn0TKgJb9EV6COKd8vIzJZQ)

Źródło: opracowanie własne na podstawie danych w serwisach.

⁴²⁰ A. Falkowski, W. Cwalina, *Marketing polityczny...*, dz. cyt., s. 439.

Dodatkowo trzeba brać pod uwagę konta indywidualne partii wewnątrz koalicji oraz indywidualnych polityków. Zarządzanie wizerunkiem polityka z wykorzystaniem Internetu staje się coraz bardziej powszechne, a Twitter zyskuje na popularności – w lipcu 2019 roku z Twittera w Polsce korzystało 5,1 miliona użytkowników, a w lipcu 2020 roku było ich 5,8 miliona (+14% użytkowników)⁴²¹. Twitter dla wielu polityków jest ważnym elementem interakcji z elektoratem, dla części pozostaje nośnikiem komunikatów politycznych skierowanych głównie do mediów⁴²². Najczęściej jednak głównym celem są działania auto promocyjne, które uzupełniają inne narzędzia z zakresu marketingu politycznego i *public relations*⁴²³.

Warto uwypuklić pracę Wojciecha Magusia, gdzie autor badał wykorzystanie Twittera w wyborach do Parlamentu Europejskiego w 2019. Analizuje on aktywność oraz wykorzystanie serwisu wśród przedstawicieli 6 ogólnopolskich komitetów wyborczych (Komitet Wyborczy Prawo i Sprawiedliwość (PiS), który zdobył 45,38% głosów. Kolejny wynik – 38,47% osiągnął Komitet Wyborczy Koalicja Europejska PO PSL SLD .N Zieloni (KE). Następnie Komitet Wyborczy Wiosna Roberta Biedronia (Wiosna), uzyskując 6,06% głosów. Mandatów nie uzyskali kandydaci Komitetu Wyborczego Wyborców Konfederacja KORWiN Braun Liroy Narodowcy (Konfederacja) (z rezultatem 4,55%), Komitetu Wyborczego Wyborców Kukiz'15 (Kukiz'15) (uzyskując wynik 3,69%) oraz Koalicyjnego Komitetu Wyborczego Lewica Razem (Lewica) (wynik 1,24%).⁴²⁴

W. Maguś w swojej pracy zbadał liczbę kont aktywnych na Twitterze dla poszczególnych komitetów wyborczych:

- PiS – liczba kandydatów: 130, procent aktywnych kont 61,5%, uzyskanych mandatów 27 (w tym 22 posiadających konto na Twitterze);

⁴²¹ W. Maguś, *Twitter jako narzędzie zarządzania wizerunkiem*, "Zarządzanie mediami", 2021, nr 9(3), s. 507.

⁴²² W. Maguś, *Twitter jako ...*, dz. cyt., s. 508.

⁴²³ A. Frame, G. Brachotte, *Le tweet stratégique: Use of Twitter as a PR Tool by French Politicians*, „Public Relations Review”, nr 41(2), s. 278–287.

⁴²⁴ W. Maguś, *Twitter jako ...*, dz. cyt., s. 515.

- KE – liczba kandydatów: 130, procent aktywnych kont 66,9%, uzyskanych mandatów 22 (w tym 20 posiadających konto na Twitterze);
- Wiosna – liczba kandydatów: 130, procent aktywnych kont 59,2%, uzyskanych mandatów 3 (w tym 3 posiadających konto na Twitterze);
- Konfederacja – liczba kandydatów: 130, procent aktywnych kont 45,4%, uzyskanych mandatów 0 - (w tym 0 posiadających konto na Twitterze);
- Kukiz'15 – liczba kandydatów: 129, procent aktywnych kont 27,9%, uzyskanych mandatów 0 - (w tym 0 posiadających konto na Twitterze);
- Lewica – liczba kandydatów: 129, procent aktywnych kont 45%, uzyskanych mandatów 0 - (w tym 0 posiadających konto na Twitterze) ⁴²⁵.

Partie, które dostały się do Parlamentu Europejskiego używały serwisu w około 59-67% biorąc pod uwagę liczbę użytkowników Twitter. W dzisiejszych czasach serwis ten stał się ważnym narzędziem *media relations*, które pozwala w czasie rzeczywistym na dotarcie do szerokiego audytorium⁴²⁶. W. Maguś w swojej pracy udowadnia, że nie istnieje korelacja pomiędzy liczbą obserwujących danego kandydata na Twitterze, a wynikiem wyborczym, ale istnieje korelacja między średnią liczbą polubień, a procentowym wynikiem w okręgu ($R = 0,52$) oraz średnią liczbą *retweetów* pojedynczego *tweeta* ($R = 0,52$) a procentowym wynikiem w okręgu⁴²⁷. Autor pokazuje z jakim natężeniem kandydaci z poszczególnych komitetów wyborczych zakładali konta na platformie Twitter od 2008 do 2019⁴²⁸.

Na podstawie wspomnianych danych została przygotowana wizualizacja skumulowanej liczby kont w serwisie Twitter oraz został wykonany model predykcyjny przewidujący liczbę kont w dalszych okresach czasowych (5 lat w przód). Do

⁴²⁵ W. Maguś, *Twitter jako ...*, dz. cyt., s. 515.

⁴²⁶ A. Evans, J. Twomey, S. Talan, *Twitter as a Public Relations Tool*, „Public RelationsJournal”, nr 5(1), s. 1-20.

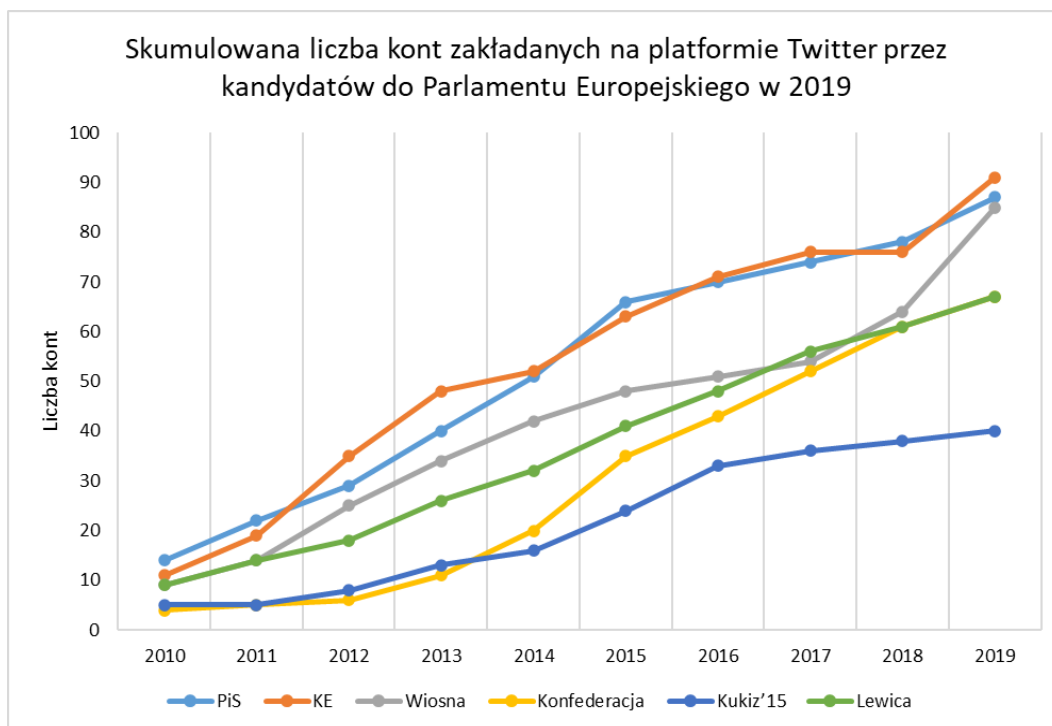
⁴²⁷ W. Maguś, *Twitter jako ...*, dz. cyt., s. 530.

⁴²⁸ W. Maguś, *Twitter jako ...*, dz. cyt., s. 519.

wykonania posłużono się metodą ARIMA (ang. *Auto-Regressive Integrated Moving Average*)⁴²⁹ z użyciem biblioteki R⁴³⁰.

Z rysunku poniżej wynika, że liczba zakładanych kont dla kandydatów do Parlamentu Europejskiego rosła z każdym rokiem od 2010. W 2015 oraz 2019 liczba zakładanych kont rosła nieco mocniej – są to lata okołowyborecze.

Rysunek 47. Skumulowana liczba kont zakładanych na platformie Twitter



Źródło: opracowanie własne na podstawie danych W. Maguś, *Twitter jako narzędzie zarządzania wizerunkiem*, "Zarządzanie Mediami", 2021, nr 9(3), s. 519.

Na podstawie sumy wszystkich zakładanych kont dla kandydatów ze wszystkich partii politycznych stworzono model ARIMA z parametrami (0,1,0) – w modelu ustawiono parametry p , d , q w sposób zoptymalizowany. Model ARIMA zasadniczo tworzy równanie liniowe, które opisuje i prognozuje dane szeregów czasowych. To równanie jest generowane przez trzy oddzielne części, które można opisać jako:

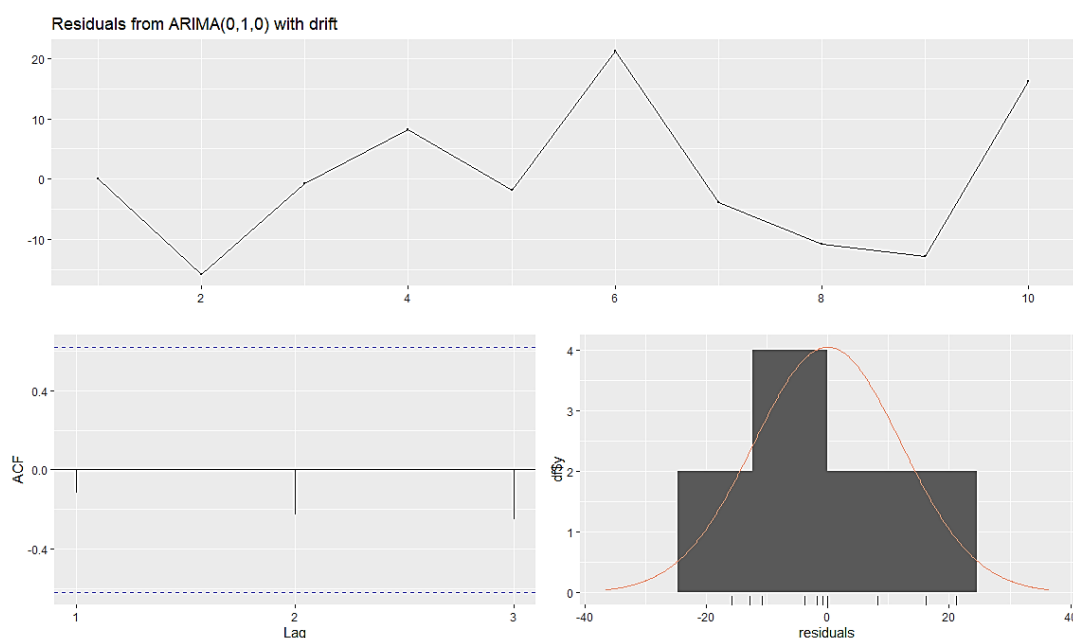
⁴²⁹ L. Ho, M. Xie, *The use of ARIMA models for reliability forecasting and analysis*, "Computers & Industrial Engineering", 1998, nr 35(1-2), s. 213-216.

⁴³⁰ Biblioteka forecast, <https://www.rdocumentation.org/packages/forecast/versions/8.16>, 20.02.2022.

- AR — autoregresja: terminy równania utworzone na podstawie przeszłych punktów danych;
- I — integracja lub różnicowanie: uwzględnienie ogólnego trendu w danych, inaczej mówiąc informuje czy szereg jest stacjonarny(0), niestacjonarny (1);
- MA — średnia ruchoma: warunki równania błędu lub szumu na podstawie przeszłych punktów danych⁴³¹.

Rysunek poniżej pokazuje, że reszty modelu przypominają rozkład normalny, stąd możemy uznać, że model będzie przewidywał przyszłe wartości poprawnie.

Rysunek 48. Algorytm ARIMA (0,1,0) – sprawdzenie jakości modelu



Źródło: opracowanie własne z wykorzystaniem biblioteki *forecast*⁴³²

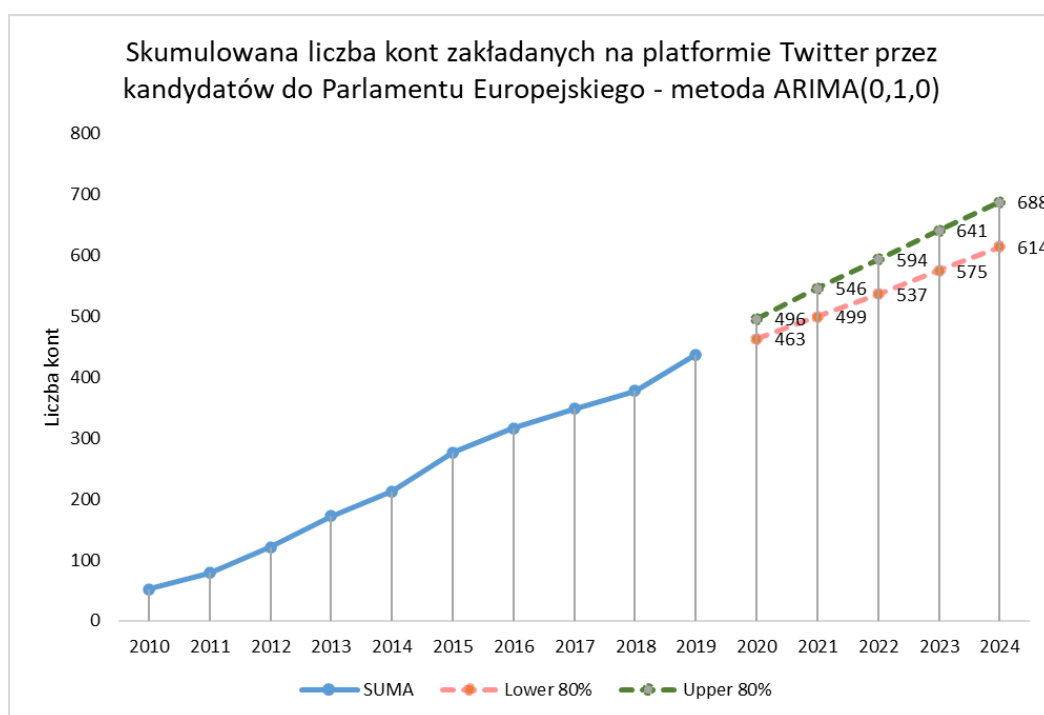
Model pokazuje predykcję na 5 lat w przód. Na rysunku poniżej wyraźnie widać, że interwały (dolne i górne) zawierają komponent trendu – wartości przyszłe z czasem osiągać wyższe wartości w przypadku pesymistycznym (*Lower 80%*) oraz w przypadku optymistycznym (*Upper 80%*). Ze względu na to, że oryginalne dane dotyczą kandydatów do Parlamentu Europejskiego można wnioskować, że podobnie zjawisko może dotyczyć

⁴³¹ What is an ARIMA model, <https://towardsdatascience.com/what-is-an-arima-model-9e200f06f9eb>, 17.06.22.

⁴³² Biblioteka *forecast*...dz. cyt., 20.02.2022

kandydatów do polskiego parlamentu i ogólnie rzecz biorąc sceny politycznej. Na wahania predykcji mogą prowadzić zewnętrzne czynniki m. in. należy brać pod uwagę fakt przejęcia w 2022 Twittera przez Elona Muska oraz zapowiadane przez niego wolnościowe podejście, co może prowadzić do zaburzeń w predykcji tego modelu⁴³³. Dodatkową kwestią która może wpłynąć na wynik predykcji jest sama liczba kandydatów do PE. W 2019 wynosiła ona 866 gdzie w poprzednich latach wyborczych oscylowała na poziomie 1279-1293 wg oficjalnych danych z pkw.gov.pl. Nie miała ona jednak widocznego wpływu na bardziej dynamiczny wzrost zakładanych kont kandydatów do PE w serwisie Twitter w roku wyborczym w 2019.

Rysunek 49. Skumulowana liczba kont zakładanych na platformie Twitter – model predykcji ARIMA



Źródło: opracowanie własne ze zastosowaniem modelu predykcyjnego z wykorzystaniem biblioteki *forecast*⁴³⁴ na podstawie danych W. Maguś, *Twitter jako narzędzie zarządzania wizerunkiem*, "Zarządzanie Mediami", 2021, nr 9(3), s. 519.

⁴³³ Elon Musk sfinalizował zakup Twittera, <https://dorzeczy.pl/ekonomia/362503/elon-musk-sfinalizowal-zakup-twittera-agrawal-zwolniony.html>, 02.02.2023

⁴³⁴ Auto ARIMA, <https://www.rdocumentation.org/packages/forecast/versions/8.16/topics/auto.arima>, 20.07.2022.

2.5. Podsumowanie rozdziału

Nie da się ukryć że rola nowych mediów rośnie, również jeśli chodzi o komunikację polityczną. Według badań CBOS w 2020 r. blisko 98% korzystało z Internetu w przedziale wiekowym 18-24 lata i 95% w przedziale 25-34 lata. 94% użytkowników miało wykształcenie wyższe. Politycy również dostrzegają rosnące możliwości komunikacji jak i budowanie elektoratu. Już ponad 20 lat temu, od wyborów prezydenckich w Polsce w 2000 roku nastąpił rozwój pod względem umieszczania treści i zawartości w Internecie (wszyscy kandydaci na prezydenta posiadali w tym czasie swoje strony WWW). Aktualnie możliwości technologiczne oraz aspekty związane z pandemią znacznie ten proces przyspieszyły. Konto na platformie Facebook posiada 444 aktualnych postów, na Twitterze zaś 372. Liczba publikowanych przez posłów treści na platformie Twitter stanowi 66% treści wszystkich publikowanych przez nich treści w mediach społecznościowych.⁴³⁵ Dodatkowo z poprzedniego rozdziału wiemy, że tematy polityczne były poruszane w *tweetach* i w Polsce występował największy odsetek *tweetów* politycznych wraz z wzmiankami politycznymi⁴³⁶. Komunikacja polityków i aktorów politycznych w Internecie jest kluczowym aspektem współczesnej formy polityki. Czynniki, które przyspieszyły użycie Internetu w celów komunikacji politycznej to przede wszystkim to, że działanie w Internecie może generować odzew w realnym świecie; sukces wykorzystania działań internetowych w branży handlowej i promocyjnej stanowi często wzór dla polityków oraz nastąpił wzrost odsetka osób posiadających dostęp do Internetu i firm, które go dostarczają⁴³⁷. Autor zbudował model predykcyjny do 2024 z wykorzystaniem algorytmów autoregresyjnych przewidujący liczbę kont w serwisie Twitter na podstawie danych dotyczących kandydatów do Parlamentu Europejskiego z którego wynika, że liczba kont zakładanych przez kandydatów do PE w 2024 będzie wahać się w przedziale: 614-688.

Stanowi to jasny sygnał, że platformy mediów społecznościowych posiadają potencjał, jeśli chodzi o analizę danych politycznych oraz tematów politycznych komentowanych w społeczeństwie.

⁴³⁵ Grayling's report...dz. cyt., 13.09.2021

⁴³⁶ C. Ducombe, *The Politics ...*, dz. cyt.

⁴³⁷ A. Falkowski, W. Cwalina, *Marketing polityczny...*, dz. cyt., s. 439.

Rozdział 3. Część Praktyczna – Metody analizy tekstu

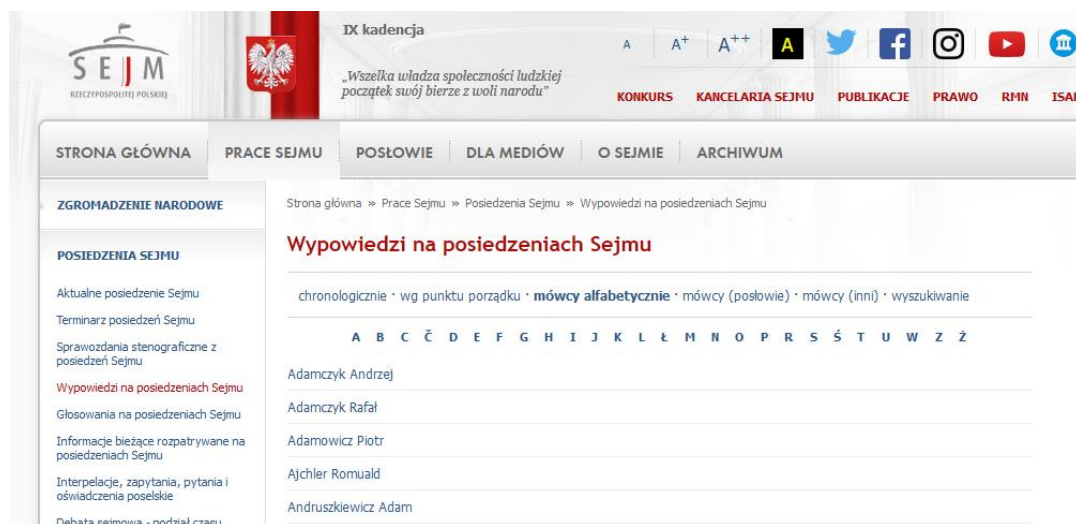
Rozdział został poświęcony badaniom własnym na danych (I) ze stenogramów sejmowych oraz (II) serwisu Twitter.

Dane do analiz wypowiedzi polityków zostały pobrane z urzędowej strony Sejmu (sejm.gov.pl) z wykorzystaniem autorskiego programu stworzonego w języku *Python* na cel niniejszej pracy w celu pobierania publicznych danych i zapisywania go w odpowiednio przygotowaną bazę danych.

3.1. Opis danych

Dane prezentują wystąpienia posłów na mównicy sejmowej za okres od 01.11.2019 do 31.12.2020.

Rysunek 50. Strona <https://sejm.gov.pl/>



Źródło: Wypowiedzi sejmowe, <https://sejm.gov.pl>, 20.04.2021.

Rysunek 51. Dane zapisane w formacie CSV

id	data	kto	tresc_format
1	2021-08-11	Rafał Adamczyk	Pani Marszałek! Wysoka Izbo! Niektóre zawarte w projek...
2	2021-08-11	Rafał Adamczyk	Pani Marszałek! Wysoka Izbo! W dzisiejszym oświadczeni...
3	2021-09-15	Rafał Adamczyk	Pani Marszałek! Wysoka Izbo! Panie Ministrze! Wysłuchał...
4	2021-07-21	Rafał Adamczyk	Panie Marszałku! Wysoka Izbo! Prezydencki projekt odb...
5	2021-07-22	Rafał Adamczyk	Panie Marszałku! Wysoka Izbo! Państwo Ministrowie! Ma...
6	2021-07-22	Rafał Adamczyk	Panie Marszałku! Panie Ministrze! Wysoka Izbo! Pytanie ...
7	2021-07-22	Rafał Adamczyk	Panie Marszałku! Panowie Ministrowie! Wysoka Izbo! Zaj...

Źródło opracowanie własne na podstawie danych ze strony: <https://sejm.gov.pl/>, 20.04.2021.

Posłów przyporządkowano do partii według barw partyjnych, z których kandydowali do Sejmu⁴³⁸. Otrzymano dane, które sklasyfikowane zostały z podziałem na 5 klubów parlamentarnych (wcześniej komitetów wyborczych) – Prawo i Sprawiedliwość (PiS), Koalicja Obywatelska (KO), Lewica, Koalicja Polska (KP), Konfederacja⁴³⁹.

Do celu pobrania danych napisane zostało oprogramowanie na bazie języka *Python* oraz biblioteki *Selenium*⁴⁴⁰, automatycznie sterowane, zapisujące wystąpienia posłów z publicznie dostępnych stenogramów.

3.2. Oczyszczenie i przygotowanie danych

Dane do analizy zostały odpowiednio przygotowane w języku *R*. Wykonane zostały standardowe przekształcenia występujące w analizie tekstu jak: ujednolicenie do małych liter, nawiasów i innych symboli⁴⁴¹. Dodatkowo usunięto imię i nazwisko posła, tytuł wypowiedzi, wszelkie wypowiedzi marszałka lub wicemarszałka, wypowiedzi innych posłów z sali – tak aby otrzymać czystą wypowiedź posła, który zabiera głos.

Następnie usunięto z wypowiedzi spójniki oraz słowa mające marginalne znaczenie takie jak: *który, ze, że, także, również, więc, jeśli, jeżeli, może, którzy, nad, tam, ich, dla-tego, aby, w, nie, czym, jak, tej, przez, po, bo, dla, za, po, są, ma, tego, które, od, te, żeby,*

⁴³⁸ Nie uwzględniano pojedynczych migracji posłów (zmiany barw partyjnych) – przyjęto przypisanie do partii (listy), w której startowali w wyborach.

⁴³⁹ Ze względu na licznosc posłów Konfederacji w Sejmie formalnie jest to koło partyjne.

⁴⁴⁰ Biblioteka Selenium, <https://pypi.org/project/selenium/> 02.02.2021.

⁴⁴¹ Symboli takich jak: przecinek, cudzysłów, znak procenta, cyfry.

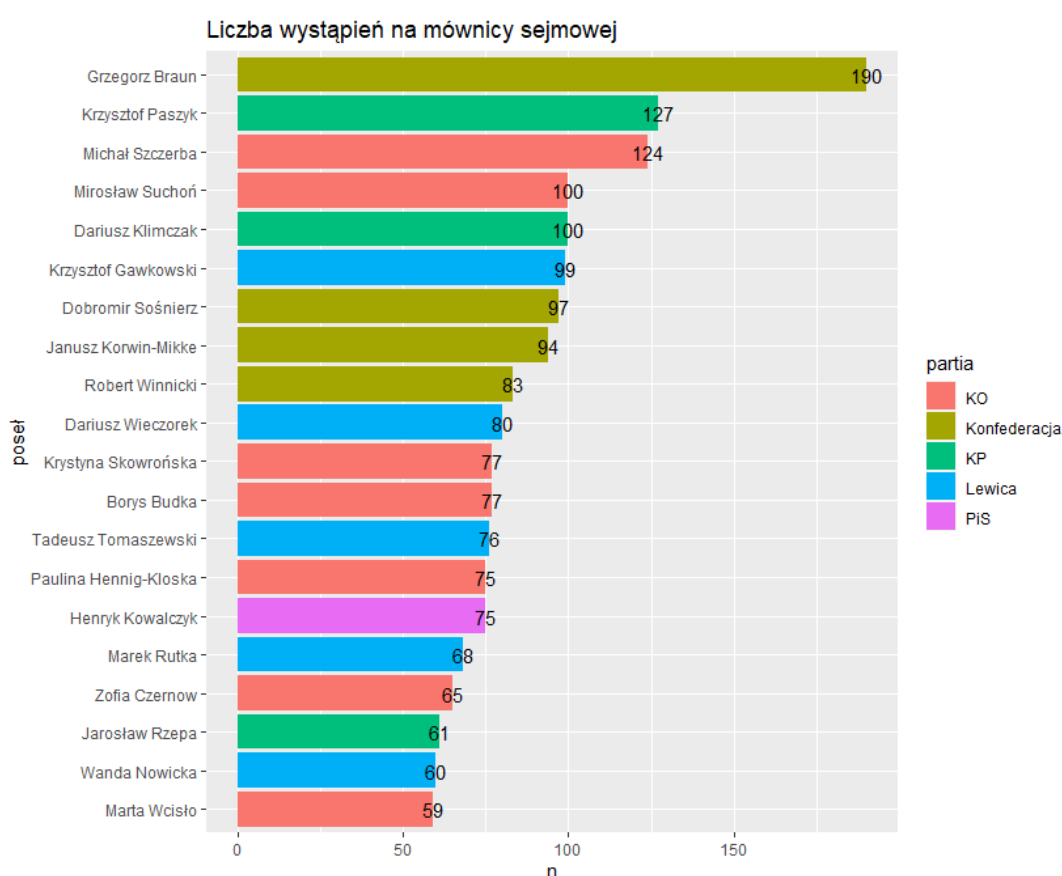
tylko, ta, i, to, na, z, jest, że, się, o, do, a, ale, co, tym, czy, r, tych, ten, już, też, która, zł oraz.

Aby dokonać tych przekształceń posłużono głównie bibliotekami jak: *tidytext*⁴⁴², *tm*⁴⁴³, *dplyr*⁴⁴⁴. Do wizualizacji wyników posłużyła biblioteka *ggplot2*⁴⁴⁵.

3.3. Liczba wystąpień według posłów i partii

Do najaktywniejszych posłów jeśli chodzi o wystąpienia na mównicy sejmowej należą: G. Braun (Konfederacja), K. Paszyk (KP), M. Szczerba (KO).

Rysunek 52. Liczba wystąpień według posłów



Źródło: opracowanie własne z wykorzystaniem biblioteki *ggplot2*.

⁴⁴² Biblioteka *tidytext*, <https://www.rdocumentation.org/packages/tidytext/versions/0.3.3>, 02.05.2021.

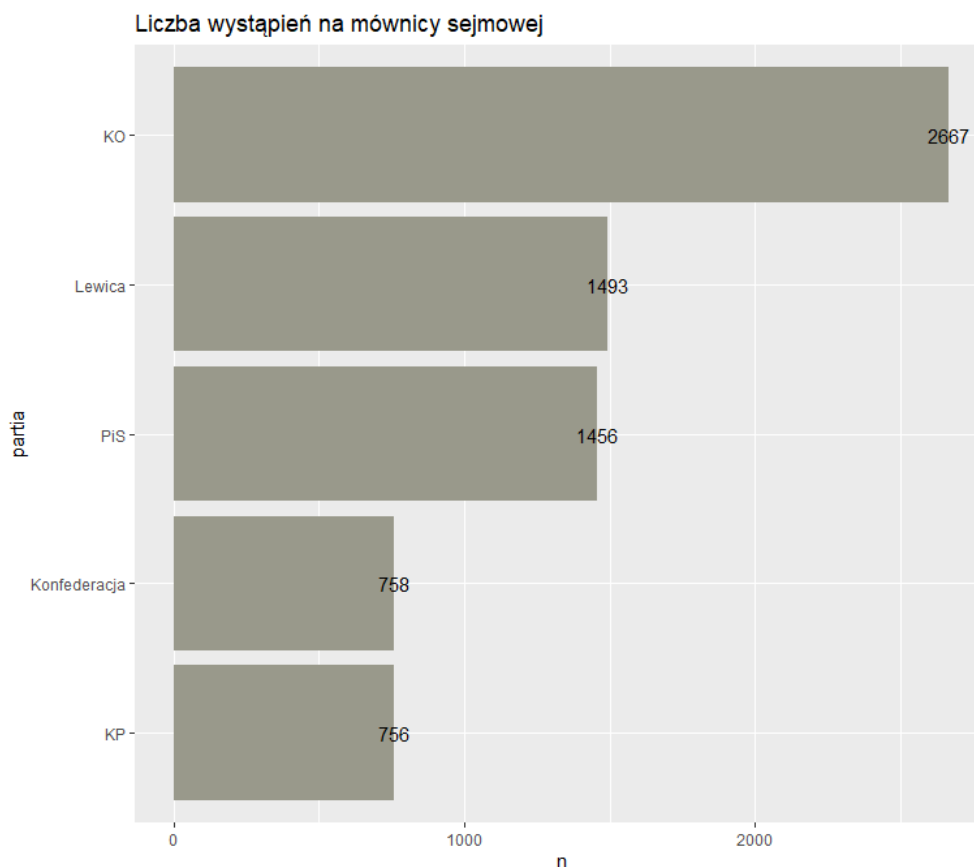
⁴⁴³ Biblioteka *tm*, <https://www.rdocumentation.org/packages/tm/versions/0.7-8>, 02.05.2021.

⁴⁴⁴ Biblioteka *dplyr*, <https://www.rdocumentation.org/packages/dplyr/versions/0.7.8>, 02.05.2021.

⁴⁴⁵ Biblioteka *ggplot2*, <https://www.rdocumentation.org/packages/ggplot2/versions/3.3.6>, 02.05.2021.

Największą liczbą wystąpień charakteryzuje się Koalicja Obywatelska (2667 wystąpień), na drugim miejscu jest Lewica (1493 wystąpień), na trzecim miejscu Prawo i Sprawiedliwość (1456 wystąpień).

Rysunek 53. Liczba wystąpień wg partii



Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Jeśli chodzi o licznosc wystąpień od początku IX kadencji Sejmu (od listopada 2019), widać spadek liczności wystąpień w okresie do kwietnia 2020, co mogło mieć związek z pandemią w Polsce oraz pierwszym lockdownem wprowadzonym od 11 marca 2020⁴⁴⁶. Od maja 2020 do sierpnia 2020 sytuacja się odwraca i wszystkie partie notują tendencję wzrostową. Następnie od września większość partii notuje stały poziom lub delikatny spadek liczności wystąpień. W całym okresie pod względem liczby wystąpień przoduje

⁴⁴⁶ COVID-19 w Polsce, https://pl.wikipedia.org/wiki/Pandemia_COVID-19_w_Polsce, 28.12.2021.

Koalicja Obywatelska. Należy pamiętać, że okres listopad-grudzień jest wliczany do analizy dwukrotnie (dwa lata).

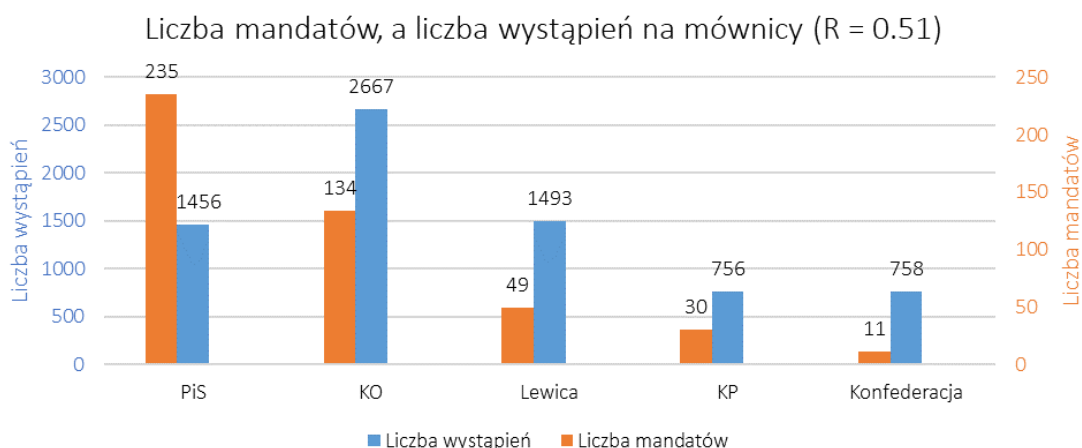
Tabela 7. Liczba wystąpień w badanym okresie, wynik wyborczy i liczba uzyskanych mandatów.

Partia	Liczba wystąpień	Wynik wyborczy 2019 [%]	Liczba mandatów
PiS	1456	43,59	235
KO	2667	27,4	134
Lewica	1493	12,56	49
KP	756	8,55	30
Konfederacja	758	6,81	11

Źródło: opracowanie własne z danych na podstawie <https://pkw.gov.pl/> 10.11.2022.

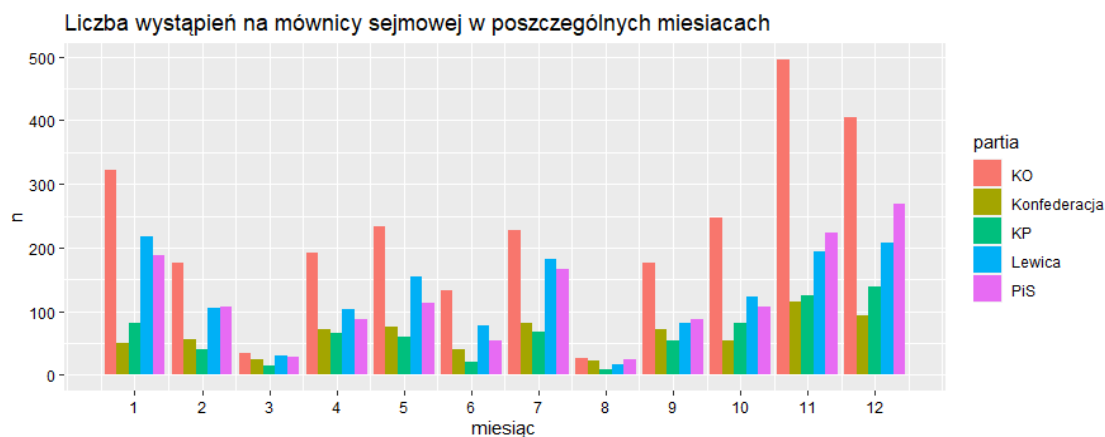
Wskaźnik współczynnika korelacji Pearsona pomiędzy liczbą wystąpień, a liczbą mandatów wynosi: $R = 0.51$ – jest to umiarkowana korelacja dodatnia.

Rysunek 54. Liczba wystąpień w badanym okresie, liczba uzyskanych mandatów oraz obliczony współczynnik korelacji.



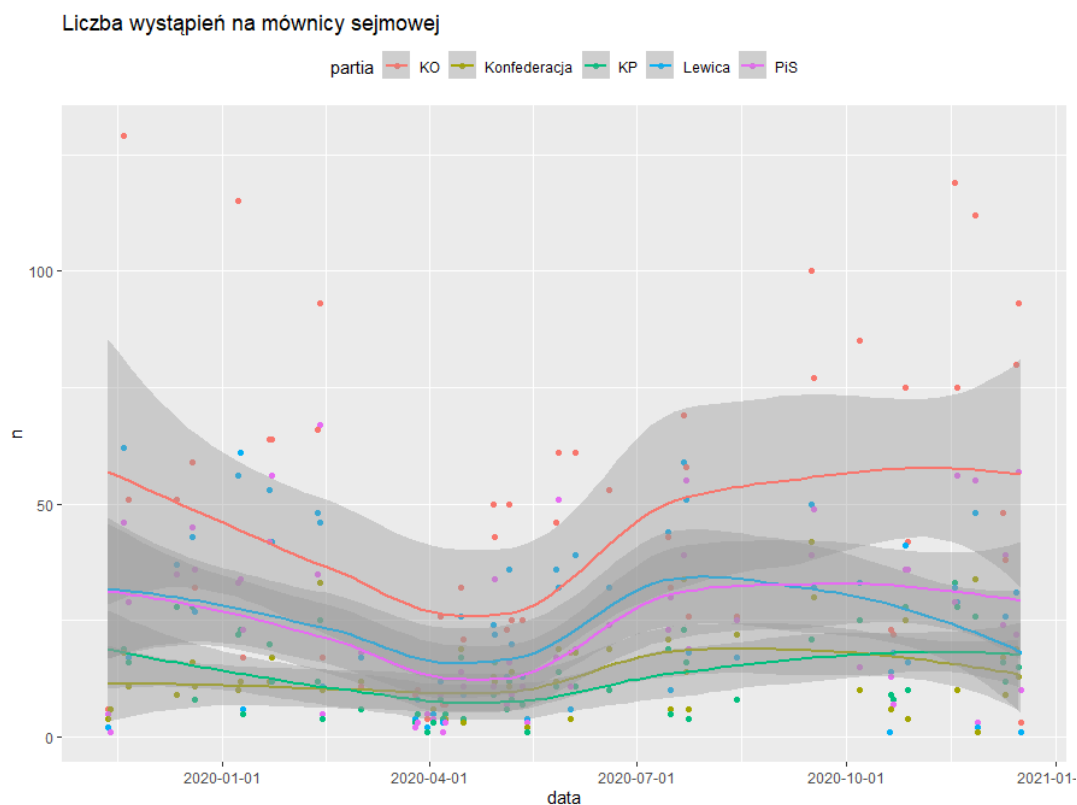
Źródło: opracowanie własne.

Rysunek 55. Liczba wystąpień w badanym okresie w poszczególnych miesiącach z podziałem na partie.



Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Rysunek 56. Liczba wystąpień wg partii w czasie



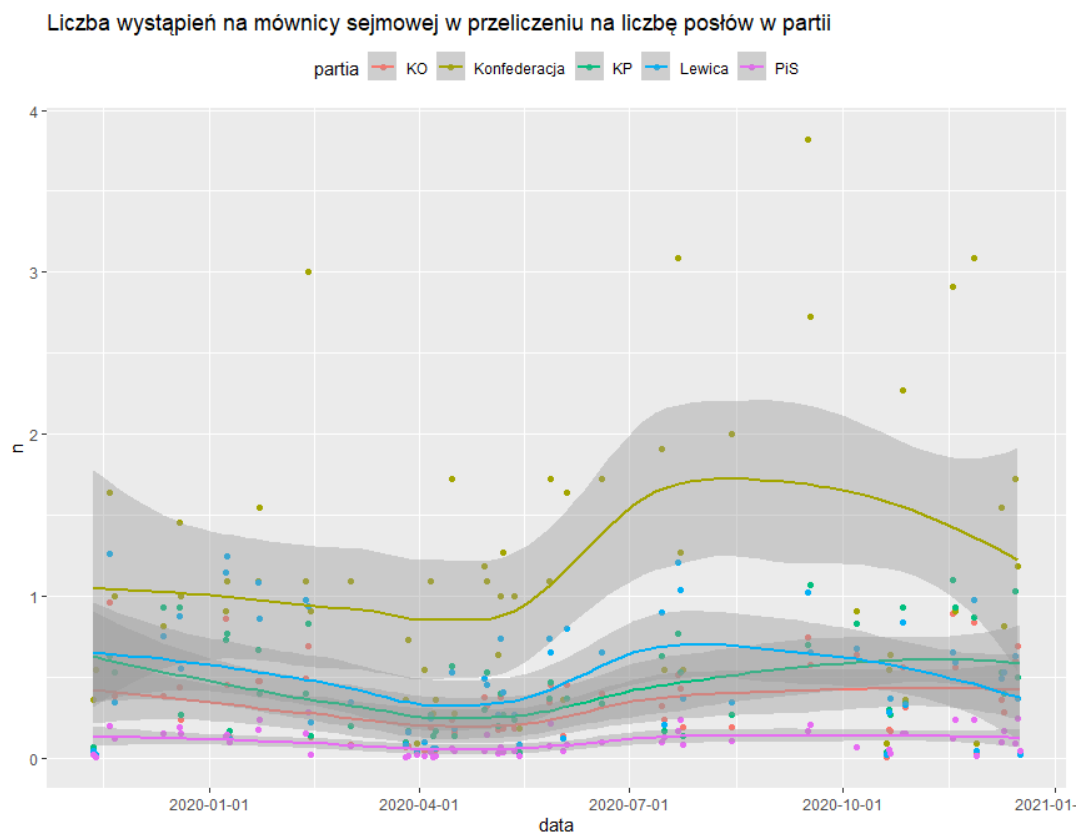
Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Liczba wystąpień według partii w czasie w procentowym ujęciu na liczbę posłów

Pod względem liczności wystąpień od początku IX kadencji Sejmu w przeliczeniu na ogólną liczbę posłów każdej z partii (od listopada 2019) tak, jak w poprzednim ujęciu widać spadek w okresie do kwietnia 2020 r.. W całym okresie najbardziej aktywną partią

(kołem poselskim) jest Konfederacja. Najmniej aktywnie prezentuje się w tym zestawieniu PiS.

Rysunek 57. Liczba wystąpień wg partii w czasie w procentowym ujęciu na liczbę posłów



Źródło: opracowanie własne z wykorzystaniem biblioteki *ggplot2*.

3.4. Liczność słów oraz podział na klastry

W tym ujęciu zwrócono uwagę na poszczególne słowa najczęściej wypowiadane na mównicy sejmowej z podziałem na partie. Dodatkowo z wykorzystaniem bibliotek *dendextend*⁴⁴⁷, *cluster*⁴⁴⁸, *factoextra*⁴⁴⁹ wykonano grupowanie (klasteryzację) słów najczęściej wypowiadanych oraz wykonano dendrogramy wraz z porównaniem wyników klasteryzacji pomiędzy partiami. W wykresach porównujących wyniki grupowania

⁴⁴⁷ Biblioteka *dendextend*, <https://www.rdocumentation.org/packages/dendextend/versions/1.16.0>, 28.12.2021.

⁴⁴⁸ Biblioteka *cluster*, <https://www.rdocumentation.org/packages/cluster/versions/2.1.3>, 28.12.2021.

⁴⁴⁹ Biblioteka *factoextra*, <https://www.rdocumentation.org/packages/factoextra/versions/1.0.7>, 28.12.2021.

kolorowymi liniami oznaczono te klastry, które są identyczne w przypadków dwóch porównywanych partii. W każdym przypadku zbadano typy połączeń w utworzonych klastrach⁴⁵⁰. Do testu użyto 4 metody połączeń: *average*, *single*, *complete*, *Ward*. Wybrano najlepszą posługując się współczynnikiem aglomeracji (*ang. Agglomerative coefficient*) w funkcji `agnes()`⁴⁵¹.

Average linkage – Powiązanie średniej polega na tym, że odległość między każdą parą obserwacji w każdym skupieniu jest sumowana i dzielona przez liczbę par w celu uzyskania średniej odległości między skupieniami. Jest to jedna z najpopularniejszych technik miar odległości w hierarchicznym grupowaniu.⁴⁵²

Single linkage – Pojedyncze wiązanie (najbliższy sąsiad) bazuje na obliczeniu najkrótszej odległości między parą obserwacji w dwóch klastrach. Może tworzyć klastry, w których obserwacje w różnych klastrach są bliżej siebie niż obserwacje w ich własnych klastrach.⁴⁵³

Complete linkage – Całkowite połączenie (najdalszy sąsiad) to sposób, w którym mierzy się odległość między najdalszą parą obserwacji w dwóch skupieniach. Ta metoda zwykle tworzy ciaśniejsze klastry niż pojedyncze wiązanie, ale te ciasne klastry mogą kończyć się bardzo blisko siebie. Jest to jedna z najpopularniejszych metryk odległości.⁴⁵⁴

Ward linkage – Metoda minimalnej wariancji Warda: minimalizuje całkowitą wariancję wewnątrz klastra. Na każdym kroku łączy się para klastrów o minimalnej odległości między klastrami.⁴⁵⁵

⁴⁵⁰ Introduction to hierarchical clustering, <https://towardsdatascience.com/introduction-hierarchical-clustering-d3066c6b560e>, 28.12.2021.

⁴⁵¹ Biblioteka cluster, <https://www.rdocumentation.org/packages/cluster/versions/2.1.2/topics/agnes>, 28.12.2021.

⁴⁵² Tamże, 28.12.2021

⁴⁵³ Tamże, 28.12.2021

⁴⁵⁴ Tamże, 28.12.2021

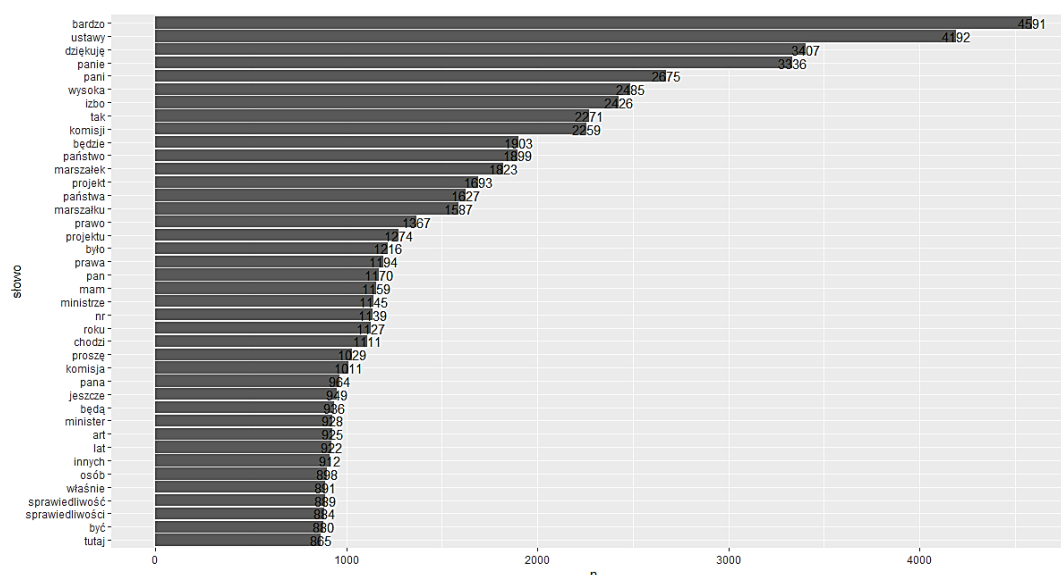
⁴⁵⁵ Ward linkage, www.datanovia.com, 28.12.2021.

Jako metodę klasteryzacji wybrano metodę hierarchicznego klasteringu (*ang. Hierarchical clustering*)⁴⁵⁶. Jest ona również aplikowalna do analizy tekstu⁴⁵⁷. Posłużono się metodą *Average silhouette method* do oceny optymalnej liczby klastrów⁴⁵⁸.

3.4.1. Prawo i Sprawiedliwość (PiS)

Do najczęściej wypowiedzianych słów posłów partii PiS należą: *bardzo* – 4591 razy, *ustawy* – 4192 razy oraz *dziękuję* – 3407 razy.

Rysunek 58. Liczba najczęściej występujących słów – PiS



Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Ze względu na osiągnięcie najwyższej wartości we współczynniku aglomeracji przez metodę Warda (współczynnik przyjmuje wartości 0-1, gdzie 1 oznacza najlepsze wiązanie), ta metoda została wybrana.

⁴⁵⁶ B. Everitt, *Cluster Analysis*, Heinemann Educ Books, London 1975; M. R. Anderberg, *Cluster Analysis for Applications*, Academic Press, New York 1973; Biblioteka hclust, <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/hclust>, 28.12.2021.

⁴⁵⁷ Text Clustering with R, <https://medium.com/@SAPCAI/text-clustering-with-r-an-introduction-for-data-scientists-c406e7454e76>, 28.12.2021.

⁴⁵⁸ Silhouette Method, <https://towardsdatascience.com/silhouette-method-better-than-elbow-method-to-find-optimal-clusters-378d62ff6891>, 28.12.2021.

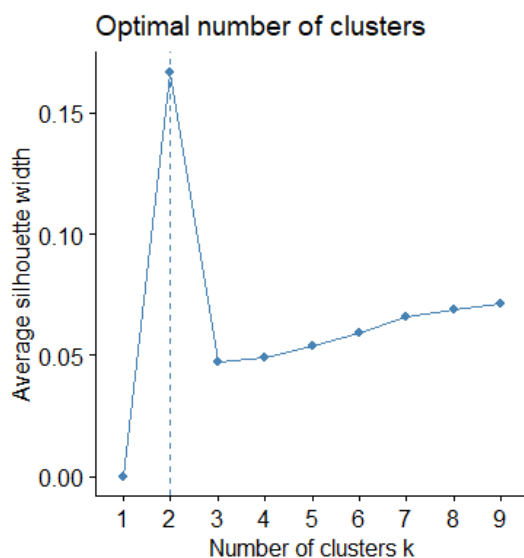
Tabela 8. Współczynnik aglomeracji

average	single	complete	ward
0.3525713	0.3208500	0.5082245	0.6160712

Źródło: opracowanie własne.

W przypadku danych dla tej partii optymalna liczba klastrów to 2.

Rysunek 59. Optymalna liczba klastrów – PiS

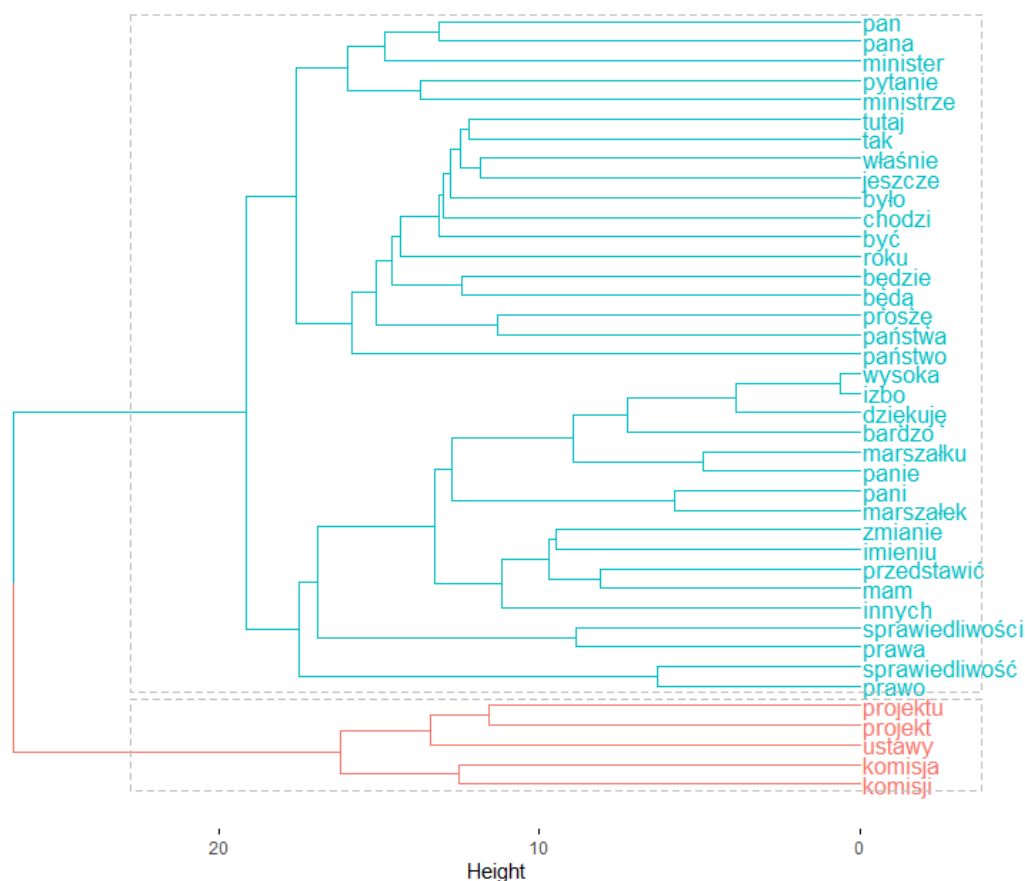


Źródło: opracowanie własne z wykorzystaniem biblioteki factoextra.

Słowa podzielone na klastry tworzą dwie grupy; jedna z nich jest bardziej ogólna i szeroka (przywitanie się, przedstawienie partii oraz intencji wypowiedzi), druga z nich tworzy klaster ze słowami dotyczącymi prac z ustawami w komisjach.

Rysunek 60. Dendrogram. Liczba najczęściej występujących słów z podziałem na klastry – PiS

Cluster Dendrogram

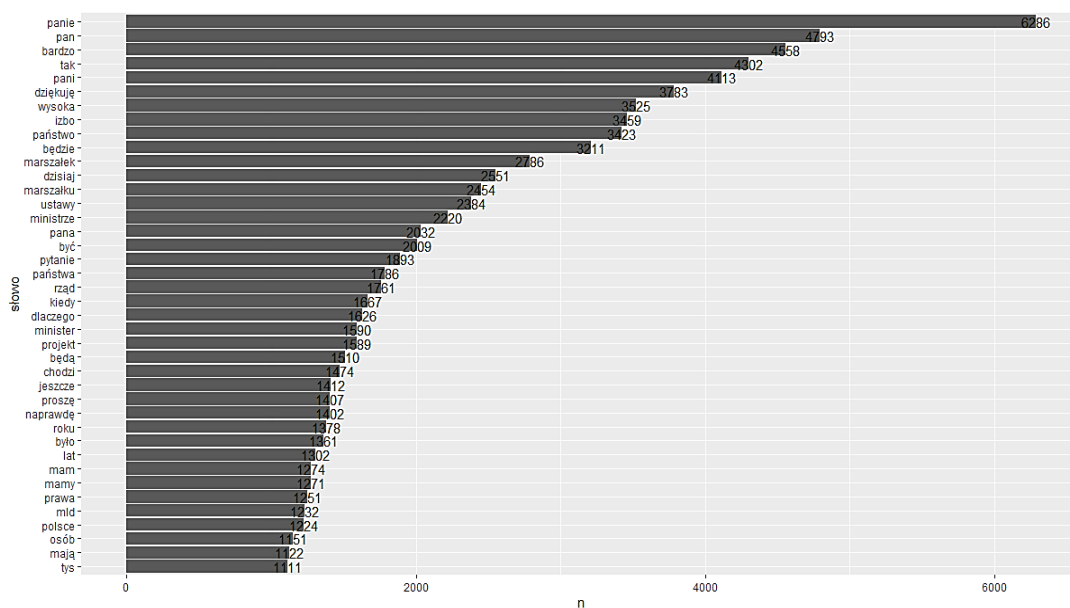


Źródło: opracowanie własne z wykorzystaniem biblioteki factextra.

3.4.2. Koalicja Obywatelska (KO)

Do najczęściej wypowiedzianych słów posłów partii KO należą: *panie* – 6286 razy, *pan* – 4793 razy oraz *bardzo* – 4558 razy.

Rysunek 61. Liczba najczęściej występujących słów – KO



Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Podobnie jak w poprzednim przypadku najwyższym współczynnikiem aglomeracji cechuje się metoda Warda. Metoda Warda została wybrana do tworzenia powiązań.

Tabela 9. Współczynnik aglomeracji

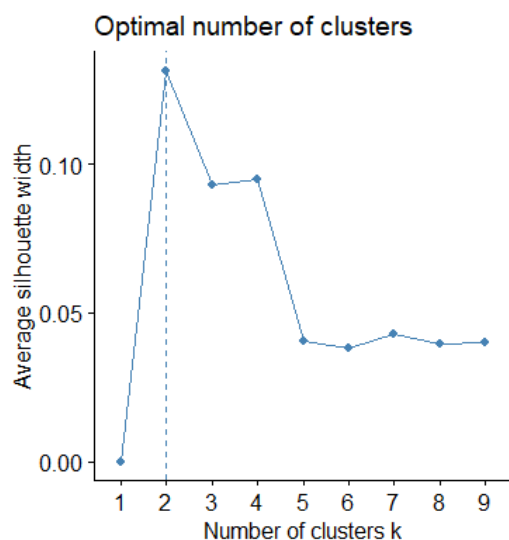
average	single	complete	ward
0.3201875	0.3022944	0.3817483	0.3965639

Źródło: opracowanie własne.

W przypadku danych dla tej partii optymalna liczba klastrów to 2.

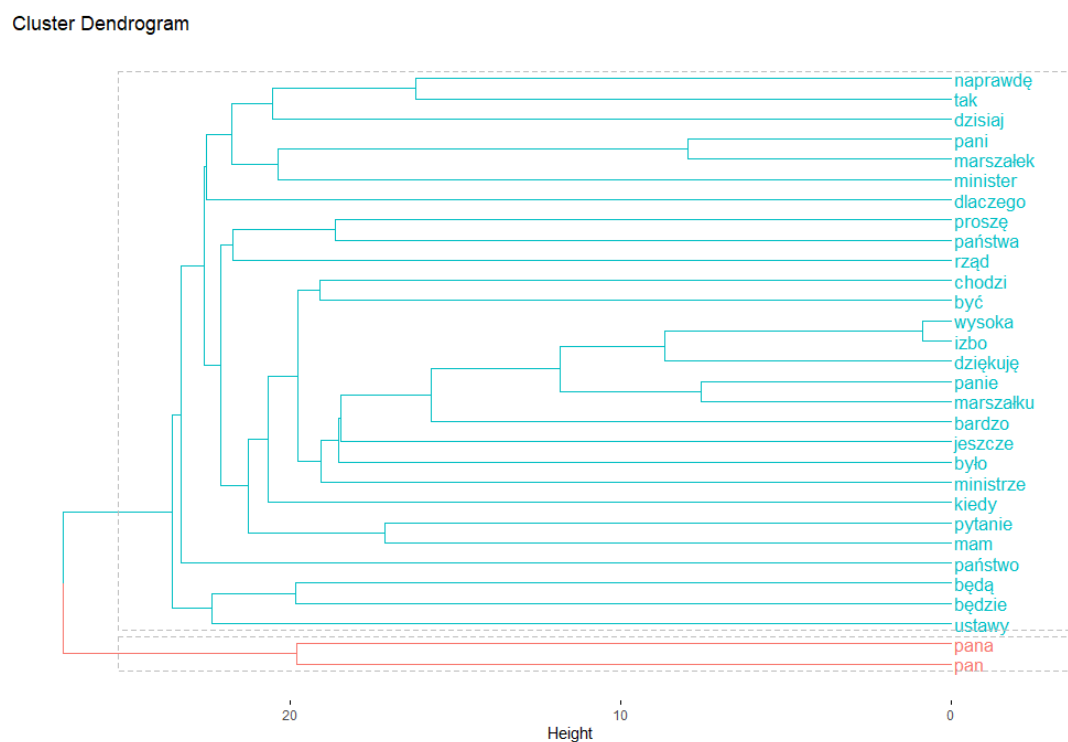
Słowa podzielone na klastry tworzą dwie grupy i podobnie jak przypadku PiS jedna z nich jest bardziej ogólna i szeroka (przywitanie się oraz przedstawienie intencji wypowiedzi), druga z nich tworzy klaster ze słowami dotyczącymi zwrotów *pana, pan*.

Rysunek 62. Optymalna liczba klastrów – KO



Źródło: opracowanie własne z wykorzystaniem biblioteki factextra.

Rysunek 63. Dendrogram. Liczba najczęściej występujących słów z podziałem na klastry – KO

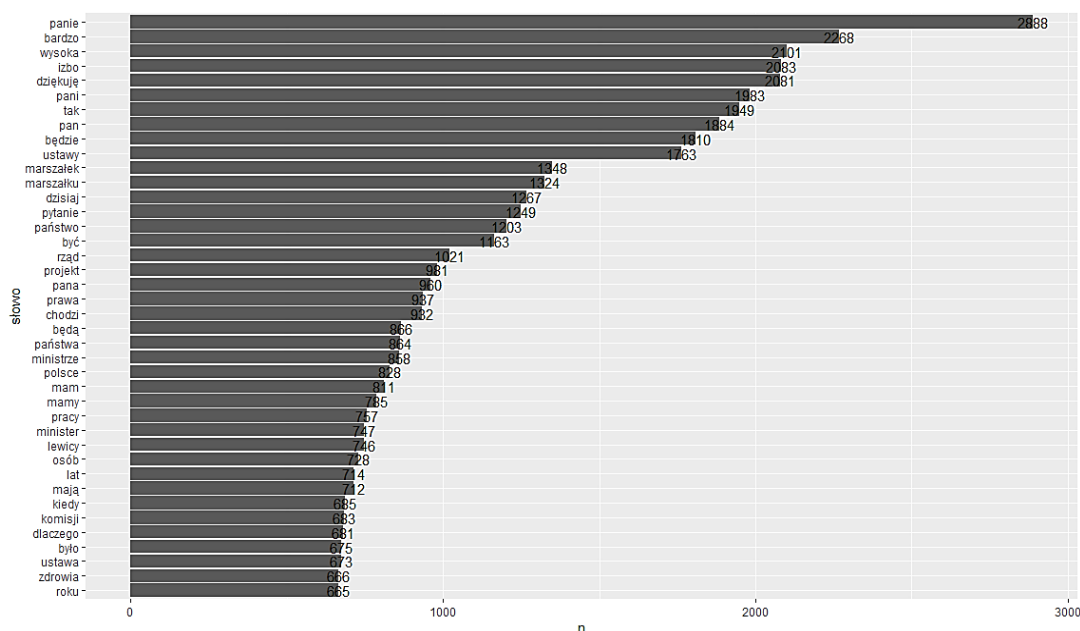


Źródło: opracowanie własne z wykorzystaniem biblioteki fviz()

3.4.3. Lewica

Do najczęściej wypowiadanych słów posłów partii Lewica należą: *panie* – 2888 razy, *bardzo* – 2268 razy oraz *wysoka* – 2101 razy.

Rysunek 64. Liczba najczęściej występujących słów – Lewica



Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Znów najwyższym współczynnikiem aglomeracji cechuje się metoda Warda. Taka metoda została wybrana.

Tabela 10. Współczynnik aglomeracji

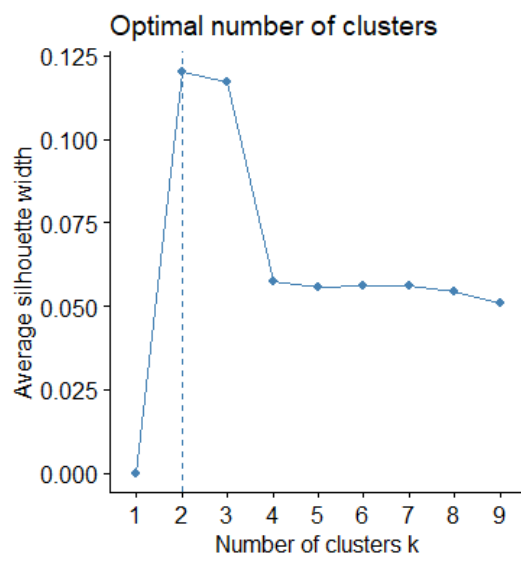
average	single	complete	ward
0.3201616	0.3080442	0.3915774	0.4015468

Źródło: opracowanie własne.

W przypadku danych dla tej partii optymalna liczba klastrów to 2.

Słowa podzielone na klastry tworzą dwie grupy i podobnie jak przypadku KO i PiS jedna z nich jest bardziej ogólna i szeroka (przywitanie się oraz przedstawienie intencji wypowiedzi), druga z nich tworzy klaster ze słowami dotyczącymi zwrotów *projekt*, *ustawy*.

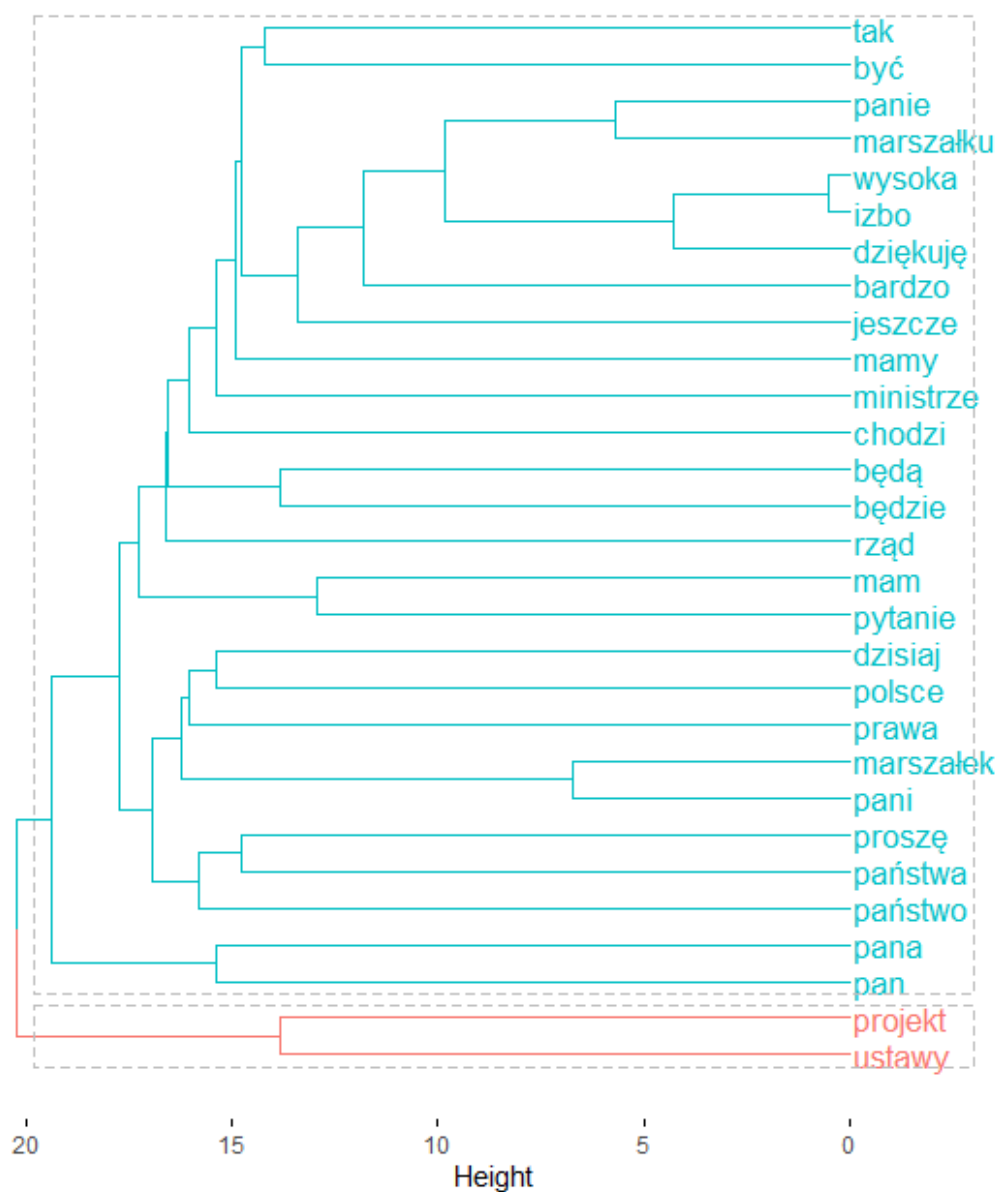
Rysunek 65. Optymalna liczba klastrow – Lewica



Źródło: opracowanie własne z wykorzystaniem biblioteki factoextra.

Rysunek 66. Dendrogram. Liczba najczęściej występujących słów z podziałem na klastry – Lewica

Cluster Dendrogram

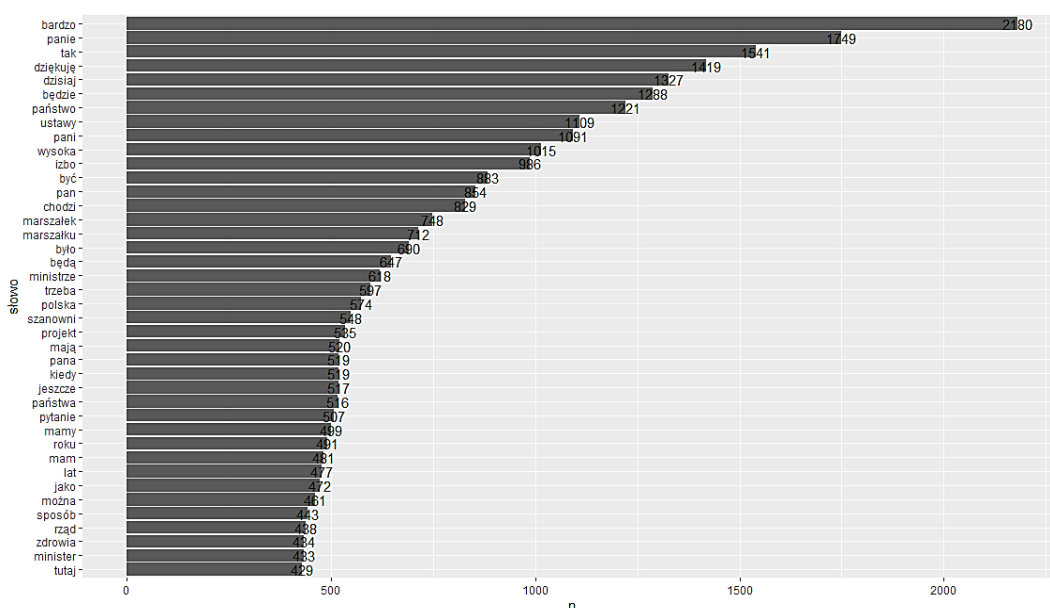


Źródło: opracowanie własne z wykorzystaniem biblioteki factextra.

3.4.4. Koalicja Polska (KP)

Do najczęściej wypowiedzianych słów posłów partii Koalicja Polska należą: *bardzo* – 2180 razy, *panie* – 1749 razy oraz *tak* – 1541 razy.

Rysunek 67. Liczba najczęściej występujących słów – KP



Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Tak jak w każdym innym poprzednio wybrana została metoda Warda, ze względu na najwyższą wartość współczynnika aglomeracji.

Tabela 11. Współczynnik aglomeracji

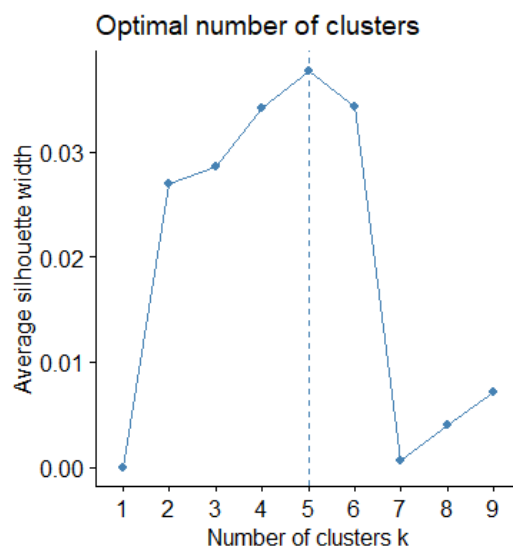
average	single	complete	ward
0.3126768	0.3140001	0.3981944	0.4527342

Źródło: opracowanie własne.

W przypadku danych dla tej partii optymalna liczba klastrów to 5.

Słowa podzielone na klastry tworzą pięć grup. Pierwsza z nich pierwszy z nich tworzą słowa: *panie*, *ministrze*. Drugi klaster tworzy słowo: *państwo*. Trzeci osobny klaster tworzą zwroty *pan*, *pana*. Czwarty klaster należy do szerszego worka słów grupującego przywołanie się, intencję wypowiedzi. Ostatni piąty klaster grupuje słowa o informacji o projekcie ustawy oraz nazwę partii – podobna liczba użytych słów pogrupowała je w pięć osobnych grup.

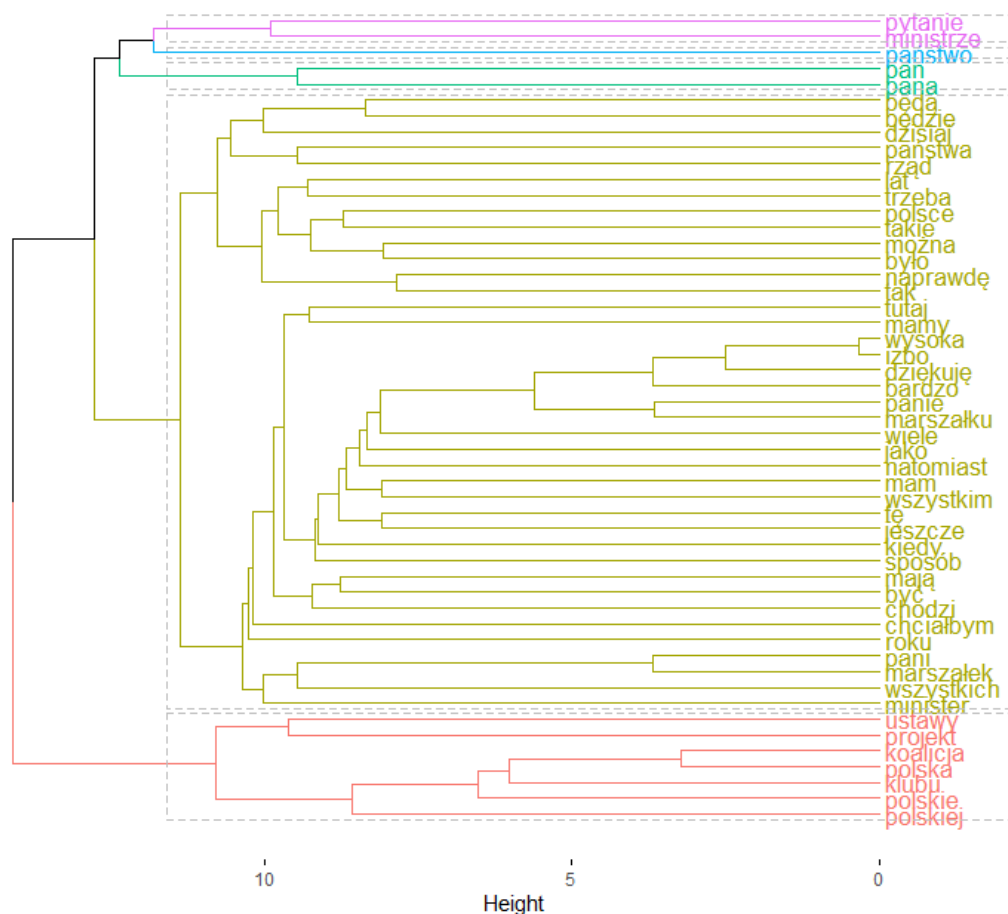
Rysunek 68. Optymalna liczba klastrow – KP



Źródło: opracowanie własne z wykorzystaniem biblioteki factoextra.

Rysunek 69. Dendrogram. Liczba najczęściej występujących słów z podziałem na klastry – KP

Cluster Dendrogram

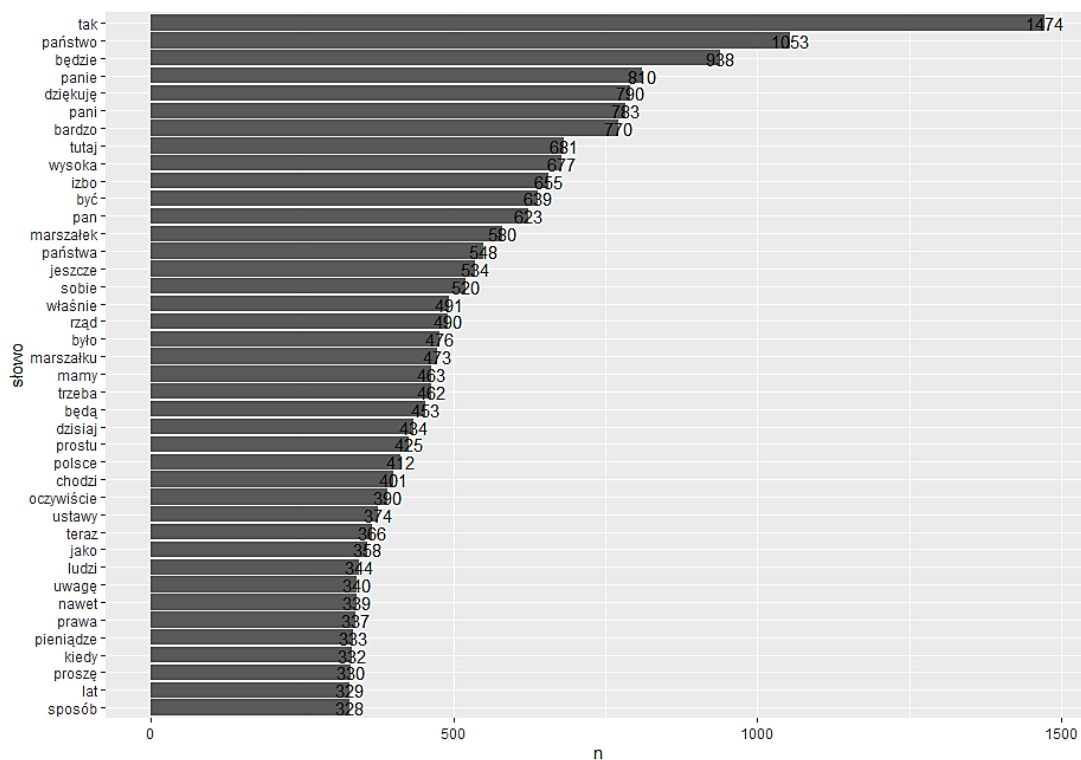


Źródło: opracowanie własne z wykorzystaniem biblioteki factextra.

3.4.5. Konfederacja

Do najczęściej wypowiadanych słów posłów koła Konfederacja należą: *tak* – 1474 razy, *państwo* – 1053 razy oraz *będzie* – 988 razy.

Rysunek 70. Liczba najczęściej występujących słów – Konfederacja



Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Tabela 12. Współczynnik aglomeracji

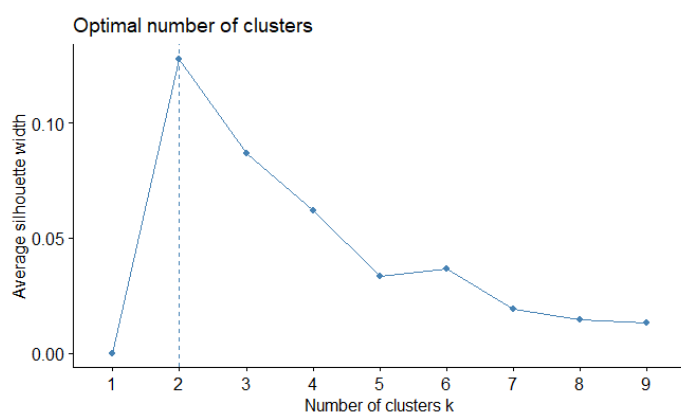
average	single	complete	ward
0.2757901	0.2744020	0.3312011	0.3972712

Źródło: opracowanie własne.

W przypadku danych dla tej partii optymalna liczba klastrow to 2, potwierdza to wartość metody Warda dla współczynnika aglomeracji.

Słowa podzielone na klastry tworzą dwie grupy i podobnie jak przypadku KO, PiS i Lewicy jedna z nich jest bardziej ogólna i szeroka (przywitanie się oraz przedstawienie intencji wypowiedzi), druga z nich tworzy klaster ze słowami dotyczącymi zwrotów *pani*, *marszałek*, *pan*.

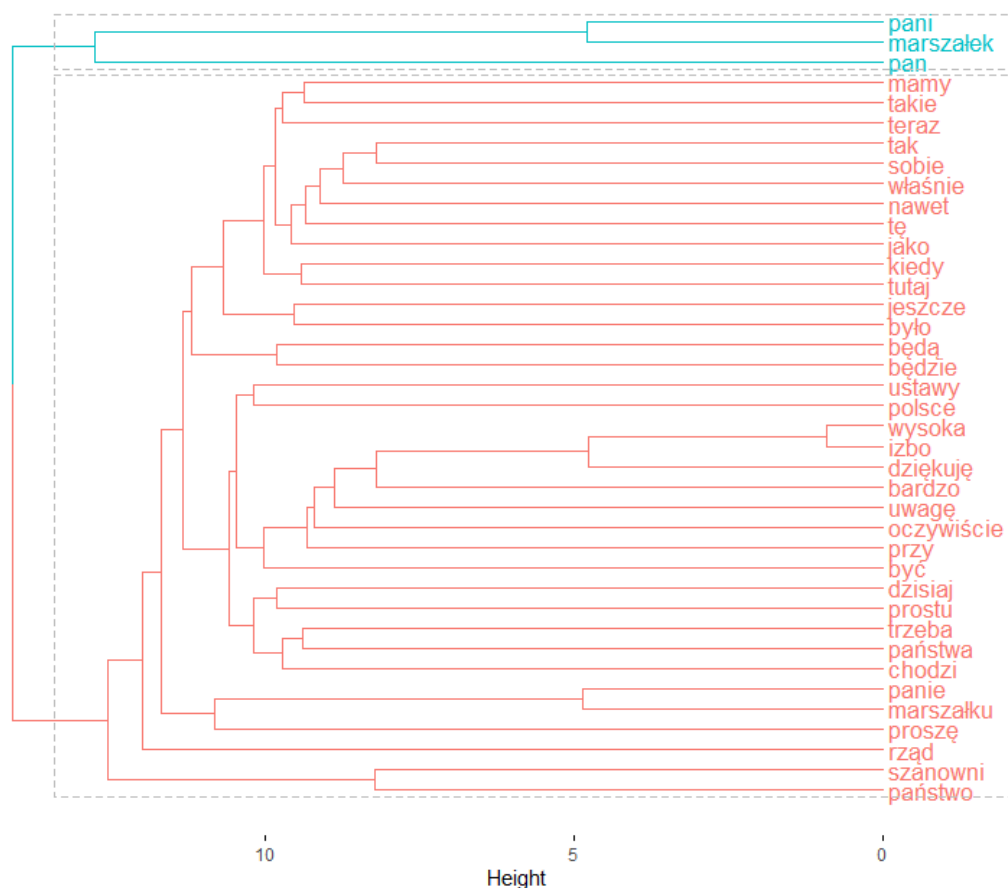
Rysunek 71. Optymalna liczba klastrów – Konfederacja



Źródło: opracowanie własne z wykorzystaniem biblioteki factextra.

Rysunek 72. Dendrogram. Liczba najczęściej występujących słów z podziałem na klastry – Konfederacja

Cluster Dendrogram



Źródło: opracowanie własne z wykorzystaniem biblioteki factextra.

3.4.6. Podsumowanie

Biorąc pod uwagę słowa o najwyższej liczności są to zwroty grzecznościowe typu: *pan, pani, panie, marszałku, dziękuję, bardzo*. We wszystkich przypadkach oprócz KP utworzone zostały po 2 klastry (jeden z zawartością słów o mniejszej, drugi o większej liczności). W przypadku KP utworzono 5 klastrow (Pierwszy tworzą słowa: *panie, mini-strze*. Drugi klaster tworzy słowo: *państwo*. Trzeci klaster tworzą zwroty *pan, pana*. Czwarty klaster grupuje przywitanie się oraz intencję wypowiedzi. Ostatni piąty klaster grupuje słowa o informacji o projekcie ustawy oraz nazwę partii). W przypadku Lewicy, PiS i KP mamy podobne klastry zawierające słowa: *projekt, ustawy*.

3.5. Porównanie dendrogramów oraz pomiar indeksem Jaccarda

W wykresach (dendrogramach) porównujących wyniki klastrowania metodą hierarchicznego klasteringu najczęściej używanych słów, kolorowymi liniami oznaczono te klastry, które są identyczne w przypadkach dwóch porównywanych partii. Ze względu na różną zawartość liczności słów w słownikach nie jest możliwe obliczenie współczynnika korelacji w klastrach (*ang. Correlation Coefficient*) oraz współczynnika splątania (*ang. Entanglement Coefficient*). Z tych względów ograniczono się tylko do interpretacji graficznej dendrogramów. Do wizualizacji posłużono się biblioteką *dendextend*⁴⁵⁹.

Zdecydowano się za to użycie indeksu *Jaccarda* (*ang. Jaccard index*) pokazującego stopień podobieństwa zbioru elementów. Współczynnik przyjmuje wartości 0 (brak podobieństwa) do 1 (idealne podobieństwo). Obliczany ze wzoru: $J(A, B) = |A \cap B| / |A \cup B|$. Zdefiniowany jest jako wielkość przecięcia podzielona przez wielkość sumy dwóch zbiorów⁴⁶⁰.

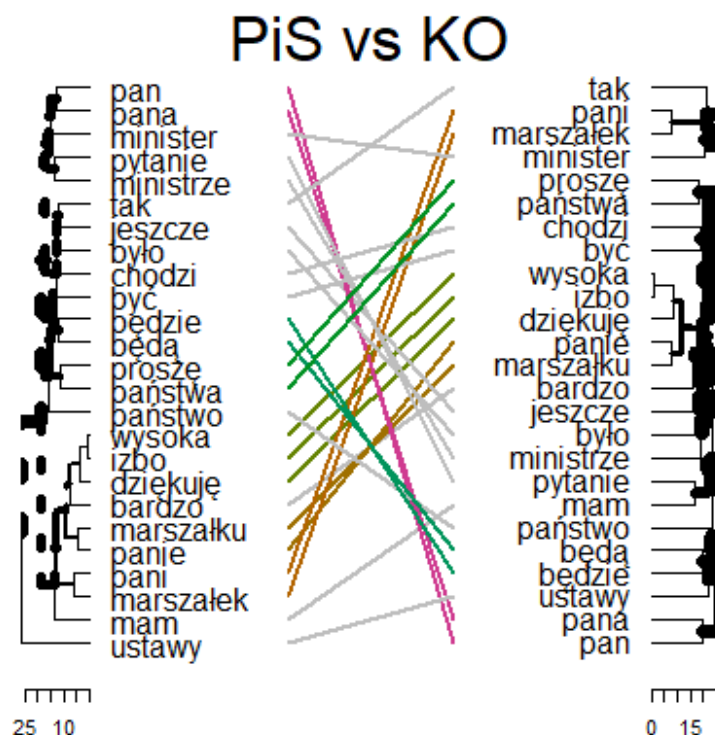
3.5.1. Prawo i Sprawiedliwość (PiS) vs Koalicja Obywatelska (KO)

Porównując obydwie dendrogramy widać podobieństwa, jeśli chodzi przede wszystkim o zwroty typu: (pan, pana), (pani, marszałek), (panie, marszałku), (proszę, państwa), (wysoka, izbo, dziękuję), (będzie, będą).

⁴⁵⁹ Biblioteka *dendextend*, <https://www.rdocumentation.org/packages/dendextend/versions/1.16.0>, 04.05.2021.

⁴⁶⁰ S. Fletcher, M. Z. Islam, *Comparing sets of patterns with the Jaccard index*, „Australasian Journal of Information Systems”, 2018, nr 22, s. 4.

Rysunek 73. Porównanie dendrogramów. Kolorami ujęto te same grupy klastrów. Po lewej: PiS, po prawej: KO

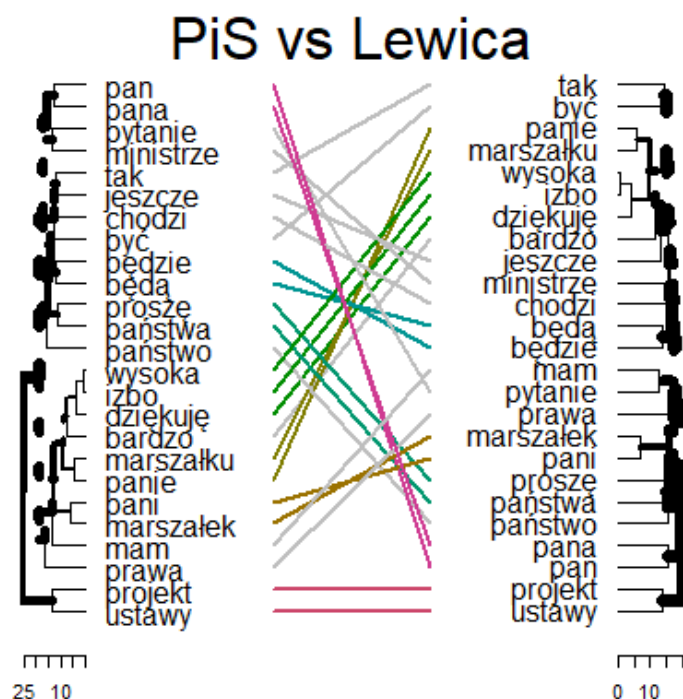


Źródło: opracowanie własne z wykorzystaniem biblioteki dendextend

3.5.2. Prawo i Sprawiedliwość (PiS) vs Lewica

Porównując obydwa dendrogramy widać podobieństwa głównie w zakresie zwrotów typu: (pan, pana), (pani, marszałek), (panie, marszałku), (proszę, państwa), (wysoka, izbo, dziękuję), (projekt, ustawy), (będzie, będą).

Rysunek 74. Porównanie dendrogramów. Kolorami ujęto te same grupy klastrów. Po lewej: PiS, po prawej: Lewica

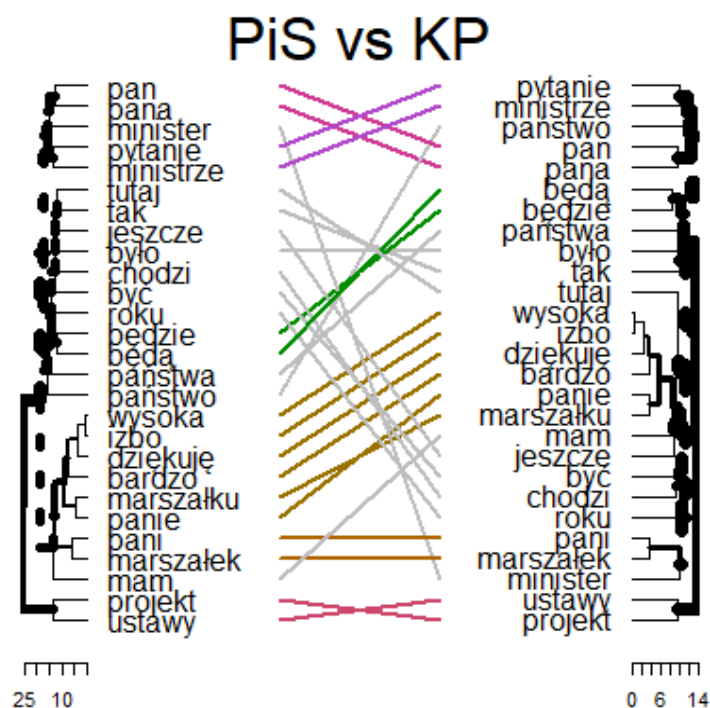


Źródło: opracowanie własne z wykorzystaniem biblioteki dendextend

3.5.3. Prawo i Sprawiedliwość (PiS) vs Koalicja Polska (KP)

W porównaniu obydwóch dendrogramów można dostrzec podobieństwa głównie w zwrotach typu: (pan, pana), (pani, marszałek), (wysoka, izbo, dziękuję, bardzo, panie, marszałku), (projekt, ustawy), (będzie, będą), (pytanie, ministże).

Rysunek 75. Porównanie dendrogramów. Kolorami ujęto te same grupy klastrów. Po lewej: PiS, po prawej: KP

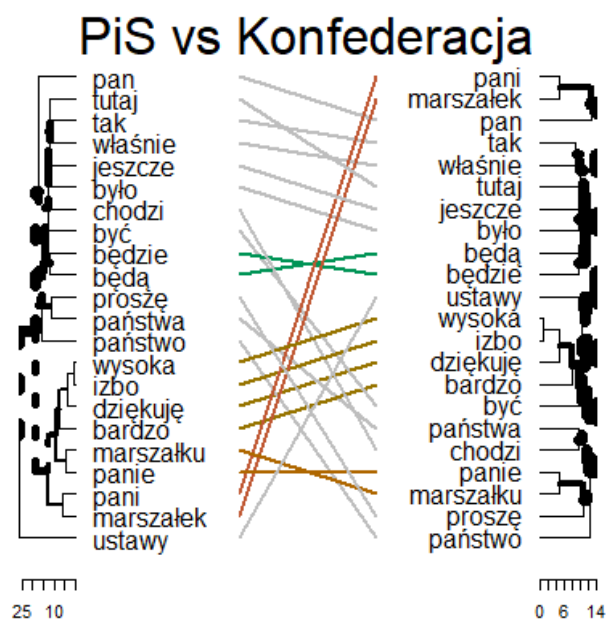


Źródło: opracowanie własne z wykorzystaniem biblioteki dendextend

3.5.4. Prawo i Sprawiedliwość (PiS) vs Konfederacja

Porównując obydwa dendrogramy widać podobieństwa, jeśli chodzi przede wszystkim o zwroty typu: (pani, marszałek), (panie, marszałku), (wysoka, izbo, dziękuję, bardzo), (będzie, będa).

Rysunek 76. Porównanie dendrogramów. Kolorami ujęto te same grupy klastrów. Po lewej: PiS, po prawej: Konfederacja

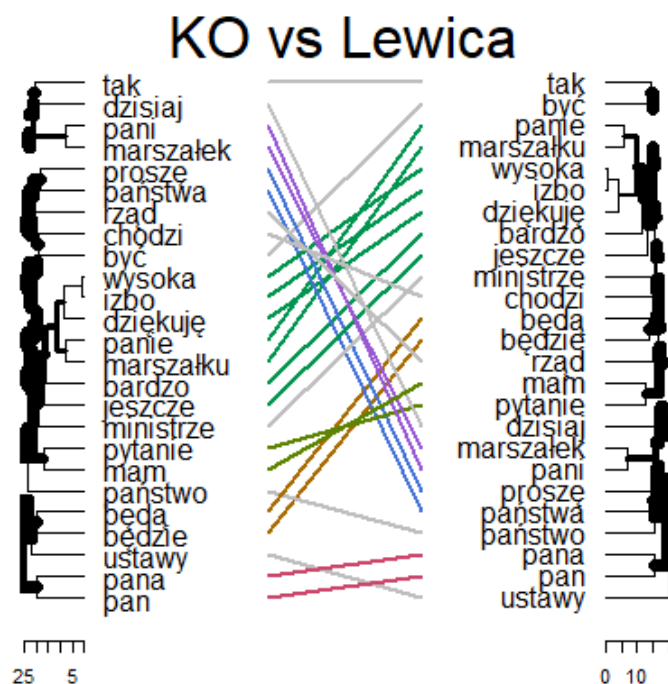


Źródło: opracowanie własne z wykorzystaniem biblioteki dendextend.

3.5.5. Koalicja Obywatelska (KO) vs Lewica

Porównując obydwie dendrogramy widoczne są podobieństwa w zakresie zwrotów typu: (pani, marszałek), (proszę, państwa), (wysoka, izbo, dziękuję, bardzo, panie, marszałku, bardzo, jeszcze), (mam, pytanie), (pana, pan), (będą, będzie).

Rysunek 77. Porównanie dendrogramów. Kolorami ujęto te same grupy klastrów. Po lewej: KO, po prawej: Lewica

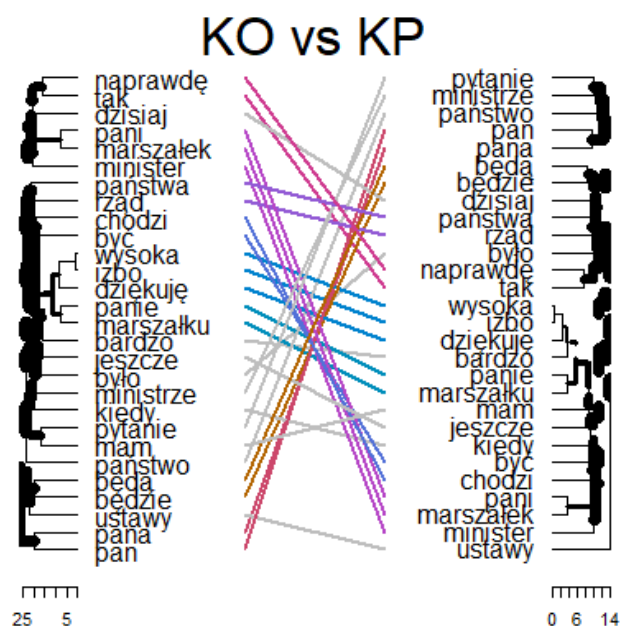


Źródło: opracowanie własne z wykorzystaniem biblioteki dendextend

3.5.6. Koalicja Obywatelska (KO) vs Koalicja Polska (KP)

Porównując obydwa dendrogramy widać podobieństwa, jeśli chodzi przede wszystkim o zwroty typu: (tak, naprawdę), (pani, marszałek, minister), (państwa, rząd), (chodzi, być), (wysoka, izba, dziękuję, bardzo, panie, marszałku), (pana, pan), (będą, będzie).

Rysunek 78. Porównanie dendrogramów. Kolorami ujęto te same grupy klastrów. Po lewej: KO, po prawej: Koalicja Polska (KP)

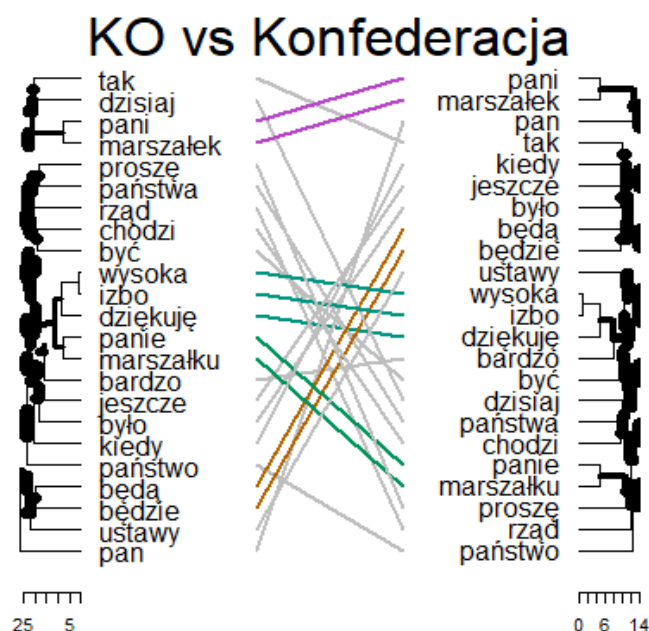


Źródło: opracowanie własne z wykorzystaniem biblioteki dendextend.

3.5.7. Koalicja Obywatelska (KO) vs Konfederacja

Porównując obydwa dendrogramy można dostrzec podobieństwa dotyczące zwrotów typu: (pani, marszałek), (wysoka, izbo, dziękuję), (panie, marszałku), (będą, będzie).

Rysunek 79. Porównanie dendrogramów. Kolorami ujęto te same grupy klastrów. Po lewej: KO, po prawej: Konfederacja

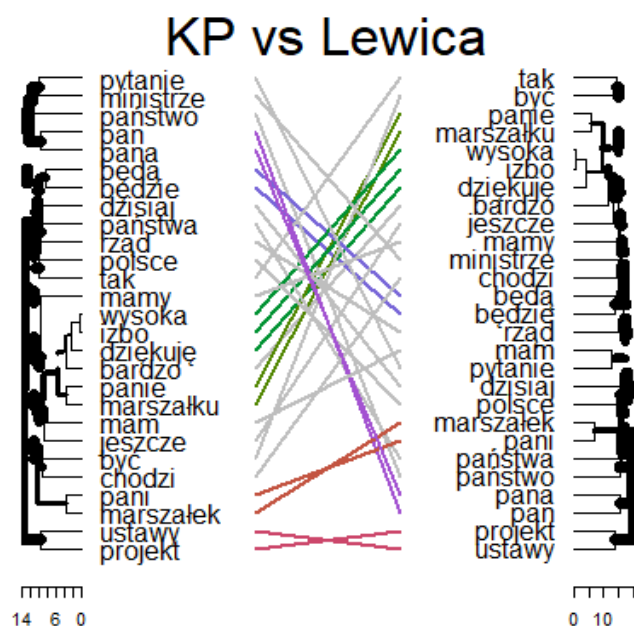


Źródło: opracowanie własne z wykorzystaniem biblioteki dendextend

3.5.8. Koalicja Polska (KP) vs Lewica

Porównując obydwie dendrogramy widać podobieństwa w przypadku zwrotów typu: (pan, pana), (pani, marszałek), (wysoka, izbo, dziękuję), (panie, marszałku), (projekt, ustawy), (będą, będzie).

Rysunek 80. Porównanie dendrogramów. Kolorami ujęto te same grupy klastrów. Po lewej: KP, po prawej: Lewica

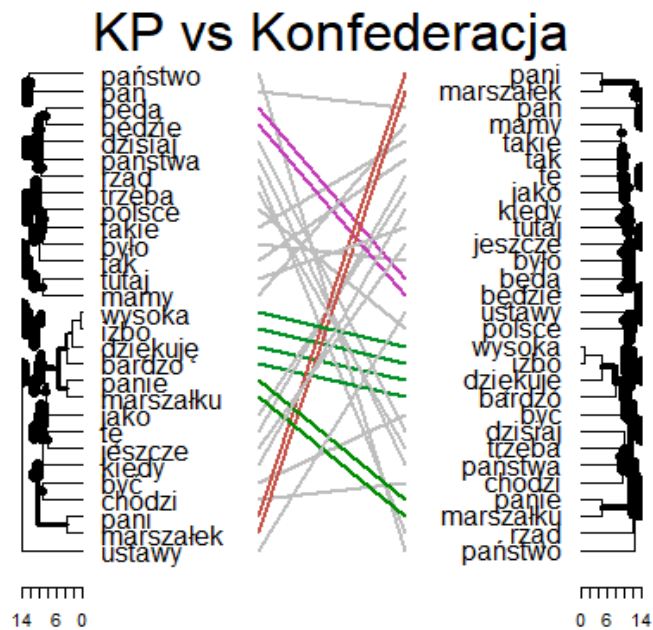


Źródło: opracowanie własne z wykorzystaniem biblioteki dendextend.

3.5.9. Koalicja Polska (KP) vs Konfederacja

Porównując obydwa dendrogramy widać podobieństwa, jeśli chodzi przede wszystkim o zwroty typu: (pani, marszałek), (wysoka, izbo, dziękuję, bardzo), (panie, marszałku), (będa, będzie).

Rysunek 81. Porównanie dendrogramów. Kolorami ujęto te same grupy klastrów. Po lewej: KP, po prawej: Konfederacja

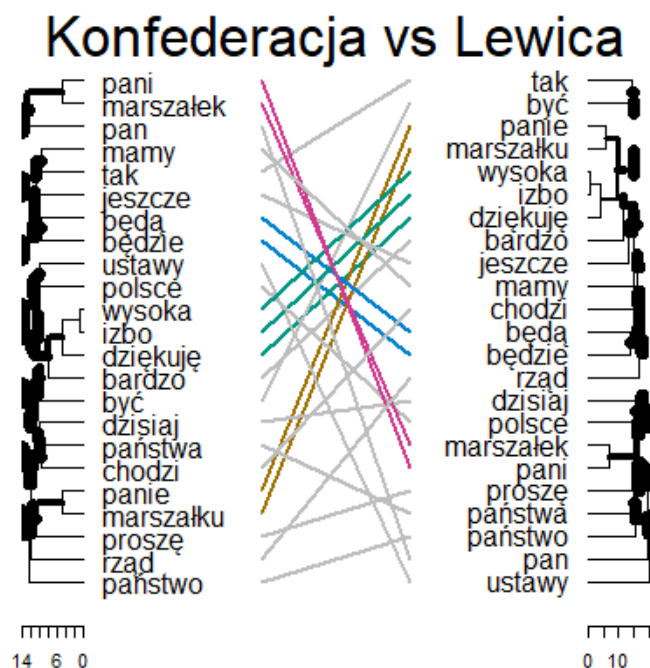


Źródło: opracowanie własne z wykorzystaniem biblioteki dendextend.

3.5.10. Konfederacja vs Lewica

W porównaniu obydwóch dendrogramów widoczne są podobieństwa przy zwrotach typu: (pani, marszałek), (wysoka, izbo, dziękuję), (panie, marszałku), (będą, będzie).

Rysunek 82. Porównanie dendrogramów. Kolorami ujęto te same grupy klastrów. Po lewej: Konfederacja, po prawej: Lewica



Źródło: opracowanie własne z wykorzystaniem biblioteki dendextend.

3.5.11. Porównanie indeksem Jaccarda

W początkowej fazie sprawdzono jaka jest zasobność indywidualnych słów każdej z partii. Ze względu na różną liczbę wystąpień wartość ta waha się od 10 812 do 24 234. Aby zoptymalizować i wyrównać słowniki użyto po 10 000 najczęściej używanych słów dla każdej z partii.

Tabela 13. Liczność użytych słów w badanym okresie

PiS	19844
KO	24234
Lewica	17190
KP	12483
Konfederacja	10812

Źródło: opracowanie własne

Wyniki przedstawiają się w następujący sposób:

Tabela 14. Indeks Jaccarda pomiędzy partiami

	PiS	KO	Lewica	KP	Konfederacja
PiS	1	0,56	0,55	0,55	0,44
KO	0,56	1	0,62	0,57	0,50
Lewica	0,55	0,62	1	0,55	0,48
KP	0,55	0,57	0,55	1	0,46
Konfederacja	0,44	0,50	0,48	0,46	1

Źródło: opracowanie własne

Tabela powyżej pokazuje wartości *Jaccard index*. Jest to naturalne, że przy tak dużej liczności słowników słowa będą się powtarzać dla każdej z partii, dlatego wartości będą uwzględniać jakąś część wspólną zbiorów. Z Tabeli 14. można odczytać, że najwyższą wartością podobieństwa cechuje się KO oraz Lewica, gdzie parametr podobieństwa jest na poziomie 0,62. Najniższą wartość prezentuje kombinacja PiS oraz Konfederacja (parametr *Jaccard index* równa się 0,44). Wartości sugerują, że Lewica oraz Koalicja Obywatelska używają najbardziej podobnych słowników słów (w grupie 10 000 najczęściej używanych słów), partia Konfederacja najbardziej odbiega od reszty partii w tym współczynniku. Tak jak w poprzedniej części w przypadku klastrów widać ogólne podobieństwa słów typu: (pani, marszałek), (panie, marszałku), (wysoka, izbo), (będą, będzie).

3.6. Analiza koszykowa zwrotów występujących razem

W tej części skupiono się na powiązaniach słów ze sobą. W tym celu wykorzystano technikę uczenia maszynowego zwaną regułami asocjacyjnymi (*ang. Association rules*). Jest to technika często używana w analizie koszykowej, ale ponadto stosuje się ją w obszarach bankowości, ubezpieczeń, medycyny, telekomunikacji. Pomaga znaleźć związki między elementami (produktami) i prawdopodobieństwami zaistnienia tych sytuacji.

Po analizie słów najczęściej występujących zdecydowano się usunąć słowa będące ogólnym przywitaniem z marszałkiem, posłami bądź Wysoką Izbą. Należą do nich: *bardzo, szanowni, dziękuję, pani, panie, marszałek, marszałku, ministrze, szanowna, szanowny, wysoka, izbo, panowie, posłowie, premierze*. Słowa te na dalszych etapach analizy zostały usunięte.

Do znalezienia połączeń pomiędzy elementami wykorzystano pakiet *R* o nazwie *arules*⁴⁶¹. Ten pakiet jest idealny do reprezentowania, analizowania i manipulowania danymi i wzorcami transakcji. Transakcje w tym przypadku oznaczają indywidualne wypowiedzi polityków. Do wizualizacji wykorzystano pakiet *arulesViz*⁴⁶². Metoda analizy koszykowej to technika, która wychwytyje siłę powiązań (asocjacji) między różnymi obiektami i potrafi wskazać wzorce występowania tych obiektów jednocześnie (współwystępowania). Często stosowana jest w wyszukiwaniu kombinacji produktów, które są często kupowane razem, co pozwala zidentyfikować wzorce i wskazać zestawy produktów kupowanych dla klientów – stąd wzięła nazwę analizy koszykowej⁴⁶³.

Aby osiągnąć ten efekt wykorzystano 3 parametry reguł asocjacyjnych: Wsparcie, Zaufanie, Podniesienie.

Wsparcie (*ang. Support*) pokazuje, jak często zestaw przedmiotów występuje we wszystkich transakcjach. Załóżmy, że mamy zestaw $X = \{\text{masło, chleb}\}$ i jego wsparcie równe 0,2. Oznacza to, że ten zestaw pozycji występuje w 1/5 wszystkich transakcji, które zawierają 20% z nich.

$$\text{Support}(\{A\} \rightarrow \{B\}) = \frac{\text{Number of transactions with both } A \text{ and } B}{\text{Total number of transactions}} = P(A \cap B)$$

Zaufanie (*ang. Confidence*) jest wskaźnikiem, który pokazuje prawdopodobieństwo wystąpienia reguły, która jest prawdziwa. Załóżmy, że mamy regułę $\{\text{masło, chleb}\} \Rightarrow \{\text{mleko}\}$ ma zaufanie $0,3 / 0,3 = 1$, co oznacza, że kupując masło i chleb, kupowaliśmy również mleko (prawdopodobieństwo 100%).

$$\text{Confidence}(\{A\} \rightarrow \{B\}) = \frac{\text{Number of transactions with both } A \text{ and } B}{\text{Total number of transactions with } A} = \frac{P(A \cap B)}{P(A)}$$

⁴⁶¹ Biblioteka *arules*, <https://www.rdocumentation.org/packages/arules/versions/1.7-2>, 28.12.2021.

⁴⁶² Biblioteka *arulesViz*, <http://mhahsler.github.io/arulesViz>, 28.12.2021.

⁴⁶³ Analiza koszykowa, <https://businessintelligence.pl/analiza-koszykowa>, 01.02.2022.

Podniesienie (*ang. Lift*) jest obliczeniem warunkowego prawdopodobieństwa wystąpienia $\{B\}$ danego $\{A\}$. Jest to pomiar prawdopodobieństwa wystąpienia, jeśli elementy są od siebie niezależne.

Jeśli wartość tego parametru jest > 1 , to pokazuje stopień, w jakim te dwa elementy są od siebie zależne – im wartość większa tym większa zależność.

Jeśli wartość podniesienia jest < 1 , pokazuje, że elementy są zastępowalne. To sugeruje, że jeden element ma negatywny wpływ na obecność innego elementu i odwrotnie.⁴⁶⁴

$$ExpectedConfidence (\{A\} \rightarrow \{B\}) = \frac{Number\ of\ transaction\ B}{Total\ number\ of\ transactions} = P(B)$$

$$Lift (\{A\} \rightarrow \{B\}) = \frac{Transactions\ containing\ A\ and\ B}{ExpectedConfidence} = \frac{P(A \cap B)}{P(A) \cdot P(B)}$$

Na potrzeby analizy posłużono się widokami ze względu na *Lift* i *Confidence*.

3.6.1. Prawo i Sprawiedliwość (PiS)

Widok ze względu na *Lift* (minimum 2 elementy)

Jeśli chodzi o parametr *Lift* największymi wartościami charakteryzują się połączenia pomiędzy: $\{platformy\} \Rightarrow \{obywatelskiej\}$, $\{obywatelskiej\} \Rightarrow \{platformy\}$, $\{czytania, projekt\} \Rightarrow \{pierwszego\}$.

Tabela 15. Widok ze względu na *Lift* (min. 2 elementy) – PiS

rank	rules	support	confidence	lift	count
1	$\{platformy\} \Rightarrow \{obywatelskiej\}$	0,05	0,84	11,6	78
2	$\{obywatelskiej\} \Rightarrow \{platformy\}$	0,05	0,74	11,6	78
3	$\{czytania, projekt\} \Rightarrow \{pierwszego\}$	0,05	0,82	11,3	73
4	$\{komisji, projekcie, sprawozdanie, ustawy\} \Rightarrow \{rządowym\}$	0,05	0,72	10,4	77
5	$\{projekcie, sprawozdanie, ustawy\} \Rightarrow \{rządowym\}$	0,05	0,72	10,3	79

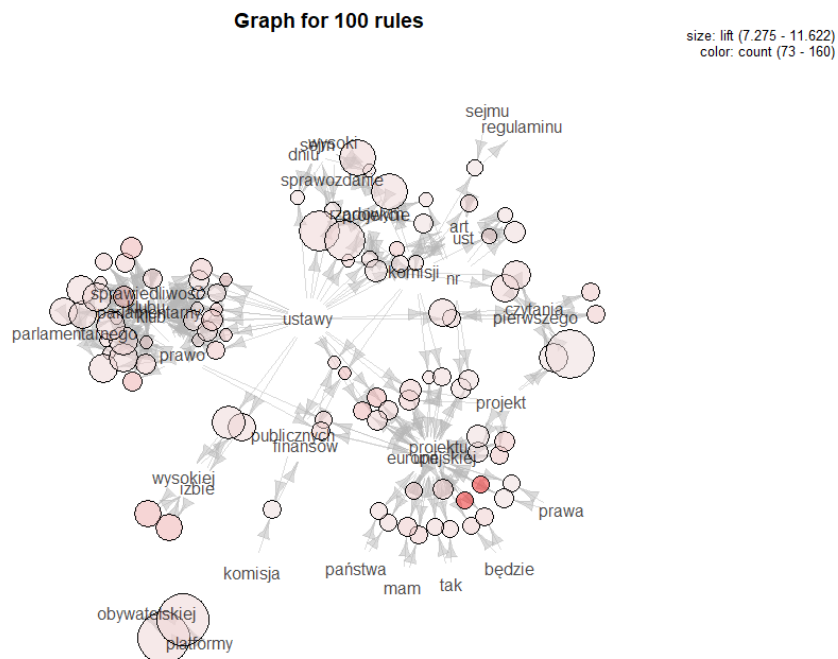
⁴⁶⁴ Association rules, <https://towardsdatascience.com/association-rules-2-aa9a77241654>, 28.12.2021; Association Mining with R, <http://r-statistics.co/Association-Mining-With-R.html>, 28.12.2021

6	{komisji, projekcie, sprawozdanie} => {rządowym}	0,05	0,69	9,9	77
7	{projekcie, sprawozdanie} => {rządowym}	0,05	0,69	9,9	79
8	{izbie, ustawy} => {wysokiej}	0,05	0,93	9,5	77
9	{pierwszego, projekt} => {czytania}	0,05	0,95	9,1	73
10	{klub, parlamentarnego, sprawiedliwość} => {parlamentarny}	0,05	0,67	9,1	78
11	{klub, parlamentarnego, prawo} => {parlamentarny}	0,05	0,67	9,1	78
12	{klub, klubu, parlamentarnego, sprawiedliwość} => {parlamentarny}	0,05	0,67	9,1	78
13	{klub, klubu, parlamentarnego, prawo} => {parlamentarny}	0,05	0,67	9,1	78
14	{klub, parlamentarnego, prawo, sprawiedliwość} => {parlamentarny}	0,05	0,67	9,1	78
15	{klub, klubu, parlamentarnego, prawo, sprawiedliwość} => {parlamentarny}	0,05	0,67	9,1	78
16	{komisji, pierwszego} => {czytania}	0,05	0,94	9,0	79
17	{pierwszego, ustawy} => {czytania}	0,05	0,94	9,0	78
18	{czytania, komisji} => {pierwszego}	0,05	0,65	9,0	79
19	{klub, parlamentarnego} => {parlamentarny}	0,05	0,66	8,9	78
20	{klub, klubu, parlamentarnego} => {parlamentarny}	0,05	0,66	8,9	78

Opracowanie własne z wykorzystaniem biblioteki *arules*

Rysunek 83. Graf. Widok ze względu na *Lift* (min. 2 elementy) – PiS

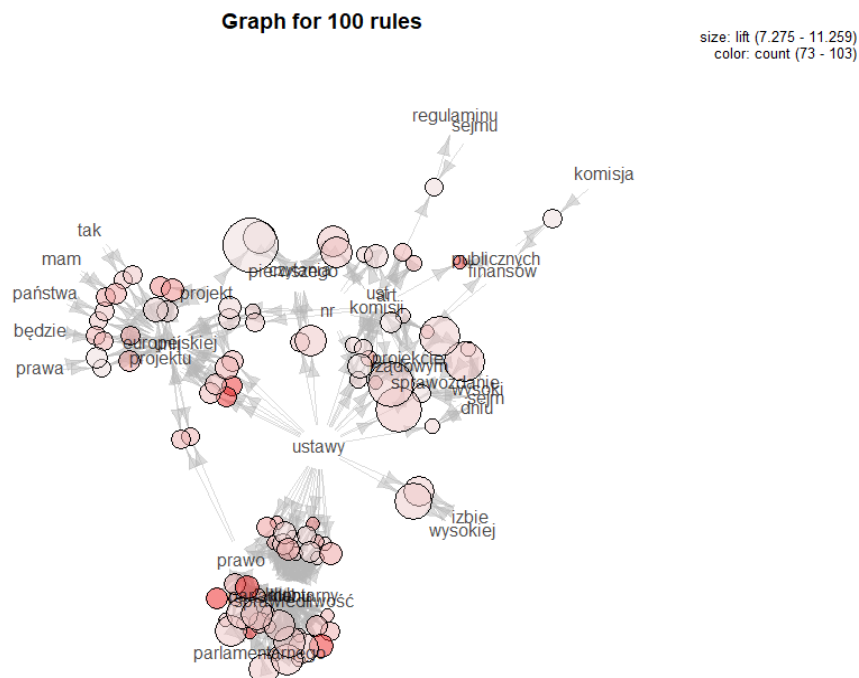
Na grafie poniżej widzimy kilka tematycznych grup, do największych należą *sprawozdania projektu* oraz połączenia słów *platformy obywatelskiej*.



Opracowanie własne z wykorzystaniem biblioteki *arulesViz*

Rysunek 84. Graf. Widok ze względu na *Lift* (min. 3 elementy) – PiS

Do największych grup ze względu na *Lift* należą obiekty: *projekt*, *pierwszego czytania* oraz *sprawozdanie*, *projekcie*, *rządowym*.



Opracowanie własne z wykorzystaniem biblioteki *arulesViz*

Widok ze względu na *Confidence* (minimum 2 elementy)

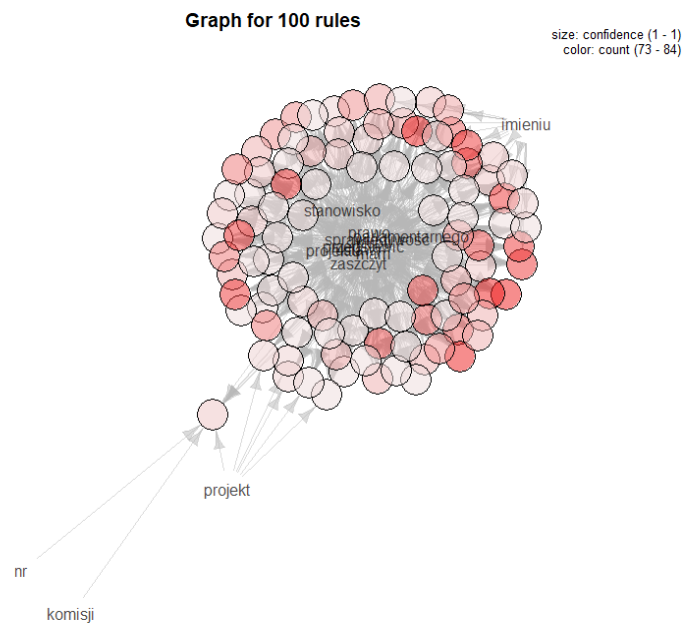
Jeśli chodzi o parametr *Confidence* największymi wartościami (1) charakteryzują się połączenia w znacznej liczbie, do najczęściej występujących należą: {klubu, sprawiedliwość} => {prawo}, {przedstawić, zaszczyt} => {mam}, {projektu, zmianie} => {ustawy}.

Tabela 16. Widok ze względu na *Confidence* (min. 2 elementy) – PiS.

rank	rules	support	confidence	lift	count
1	{klubu, sprawiedliwość} => {prawo}	0,16	1,00	3,6	233
2	{przedstawić, zaszczyt} => {mam}	0,16	1,00	2,7	231
3	{projektu, zmianie} => {ustawy}	0,15	1,00	2,3	219
4	{sprawiedliwość, ustawy} => {prawo}	0,15	1,00	3,6	217
5	{przede} => {wszystkim}	0,14	1,00	5,2	209
6	{nr, zmianie} => {ustawy}	0,14	1,00	2,3	208
7	{przedstawić, ustawy, zaszczyt} => {mam}	0,14	1,00	2,7	202
8	{parlamentarnego, prawo} => {klubu}	0,13	1,00	5,4	194
9	{parlamentarnego, sprawiedliwość} => {klubu}	0,13	1,00	5,4	193
10	{parlamentarnego, prawo, sprawiedliwość} => {klubu}	0,13	1,00	5,4	193
11	{parlamentarnego, sprawiedliwość} => {prawo}	0,13	1,00	3,6	193
12	{klubu, parlamentarnego, sprawiedliwość} => {prawo}	0,13	1,00	3,6	193
13	{niektórych, ustaw} => {ustawy}	0,12	1,00	2,3	177
14	{ustaw, zmianie} => {ustawy}	0,12	1,00	2,3	177
15	{projektu, sprawiedliwość} => {prawo}	0,12	1,00	3,6	173
16	{niektórych, ustaw, zmianie} => {ustawy}	0,12	1,00	2,3	169
17	{klub, sprawiedliwość} => {prawo}	0,11	1,00	3,6	166
18	{projektu, sprawiedliwość, ustawy} => {prawo}	0,11	1,00	3,6	163
19	{parlamentarnego, ustawy} => {klubu}	0,11	1,00	5,4	162
20	{komisji, nr, zmianie} => {ustawy}	0,11	1,00	2,3	162

Opracowanie własne z wykorzystaniem biblioteki *arules*

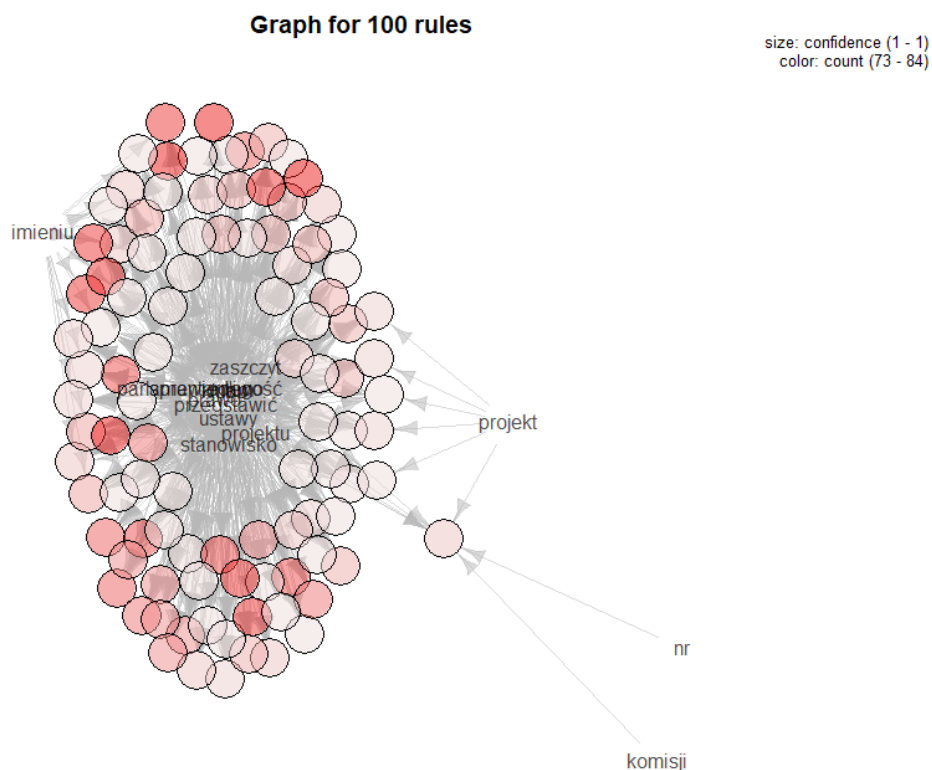
Rysunek 85. Graf. Widok ze względu na *Confidence* (min. 2 elementy) – PiS.



Źródło: Opracowanie własne z wykorzystaniem biblioteki arulesViz

Słowa mające najwyższy parametr *Confidence* dotyczą zwrotów przedstawienia przez Prawo i Sprawiedliwość projektu ustawy wraz ze zwrotem: *mam zaszczyt*.

Rysunek 86. Graf. Widok ze względu na *Confidence* (min. 3 elementy) – PiS.



Źródło: Opracowanie własne z wykorzystaniem biblioteki *arulesViz*

Słowa tworzące 3 elementowe powiązania mające najwyższy parametr *Confidence* dotyczą podobnie jak wyżej zwrotów: przedstawienia przez partię (Prawo i Sprawiedliwość) projektu ustawy i podanie numeru komisji.

3.6.2. Koalicja Obywatelska (KO)

Widok ze względu na *Lift* (minimum 2 elementy)

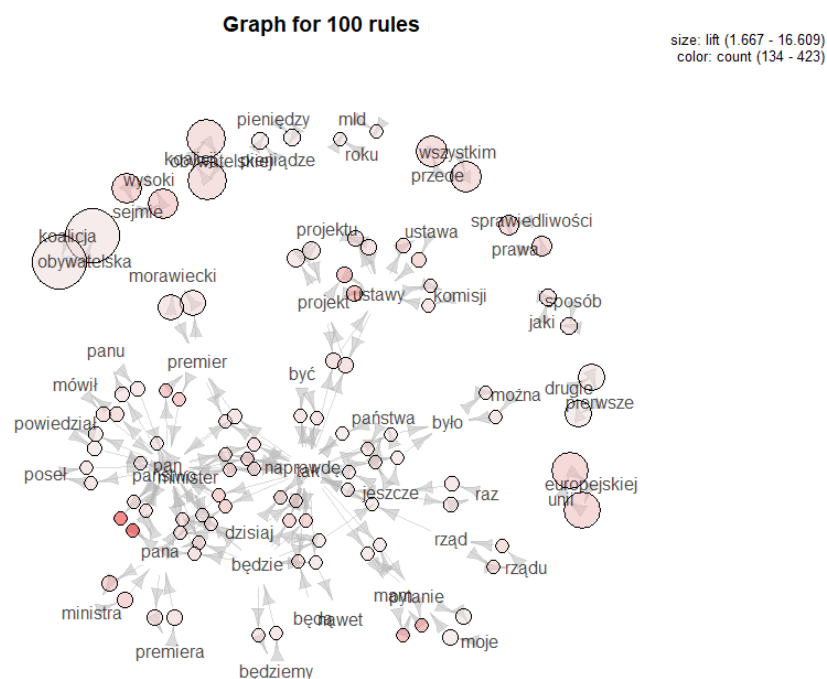
Jeśli chodzi o parametr *Lift* największymi wartościami charakteryzują się połączenia pomiędzy: $\{koalicja\} \Rightarrow \{obywatelska\}$, $\{obywatelska\} \Rightarrow \{koalicja\}$, $\{koalicji\} \Rightarrow \{obywatelskiej\}$

Tabela 17. Widok ze względu na *Lift* (min. 2 elementy) – KO

rank	rules	support	confidence	lift	count
1	{koalicja} => {obywatelska}	0,05	0,97	16,61	139
2	{obywatelska} => {koalicja}	0,05	0,90	16,61	139
3	{koalicji} => {obywatelskiej}	0,07	0,89	11,01	182
4	{obywatelskiej} => {koalicji}	0,07	0,85	11,01	182
5	{unii} => {europejskiej}	0,08	0,95	10,19	215
6	{europejskiej} => {unii}	0,08	0,86	10,19	215
7	{przede} => {wszystkim}	0,07	0,98	8,12	195
8	{wszystkim} => {przede}	0,07	0,61	8,12	195
9	{wysoki} => {sejmie}	0,08	0,95	7,81	223
10	{sejmie} => {wysoki}	0,08	0,69	7,81	223
11	{drugie} => {pierwsze}	0,05	0,67	6,52	136
12	{pierwsze} => {drugie}	0,05	0,50	6,52	136
13	{morawiecki} => {premier}	0,06	0,88	6,29	153
14	{premier} => {morawiecki}	0,06	0,41	6,29	153
15	{sprawiedliwości} => {prawa}	0,07	0,72	4,38	198
16	{prawa} => {sprawiedliwości}	0,07	0,45	4,38	198
17	{projektu} => {projekt}	0,06	0,56	3,26	149
18	{projekt} => {projektu}	0,06	0,33	3,26	149
19	{jaki} => {sposób}	0,06	0,53	3,14	166
20	{sposób} => {jaki}	0,06	0,37	3,14	166

Opracowanie własne z wykorzystaniem biblioteki arules

Rysunek 87. Graf. Widok ze względu na *Lift* (min. 2 elementy) – KO

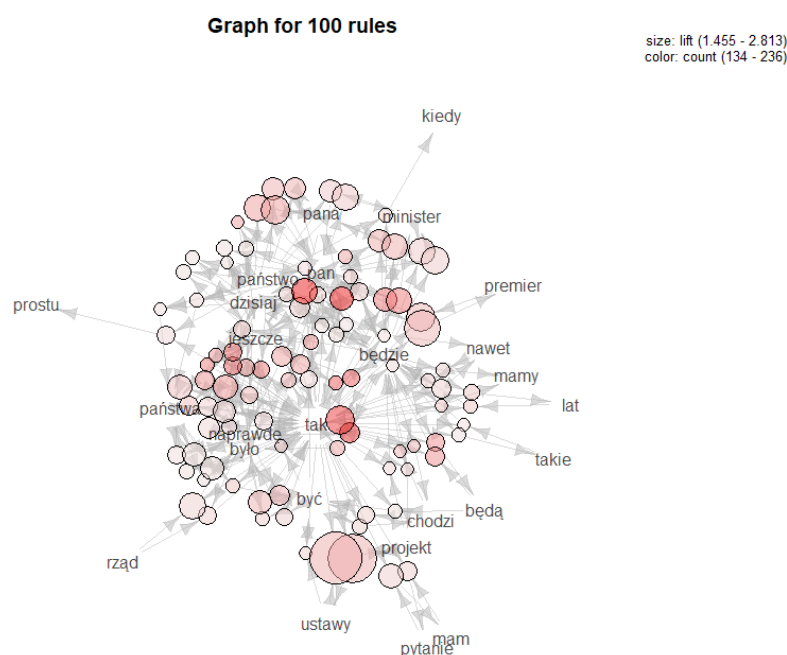


Źródło: Opracowanie własne z wykorzystaniem biblioteki arulesViz

Widzimy że ze względu na parametr *Lift* najsilniejsze powiązania dotyczą kombinacji: *koalicja*, *obywatelska* oraz *unii*, *europejskiej*⁴⁶⁵.

⁴⁶⁵ Wszystkie słowa w zbiorze, ujednolicone do małych liter po przygotowaniu danych.

Rysunek 88. Graf. Widok ze względu na *Lift* (min. 3 elementy) – KO



Źródło: Opracowanie własne z wykorzystaniem biblioteki *arulesViz*

W widoku na *Lift* najmocniejsze powiązania dotyczą wzmianek projektu ustawy oraz pytań związanych z tym tematem.

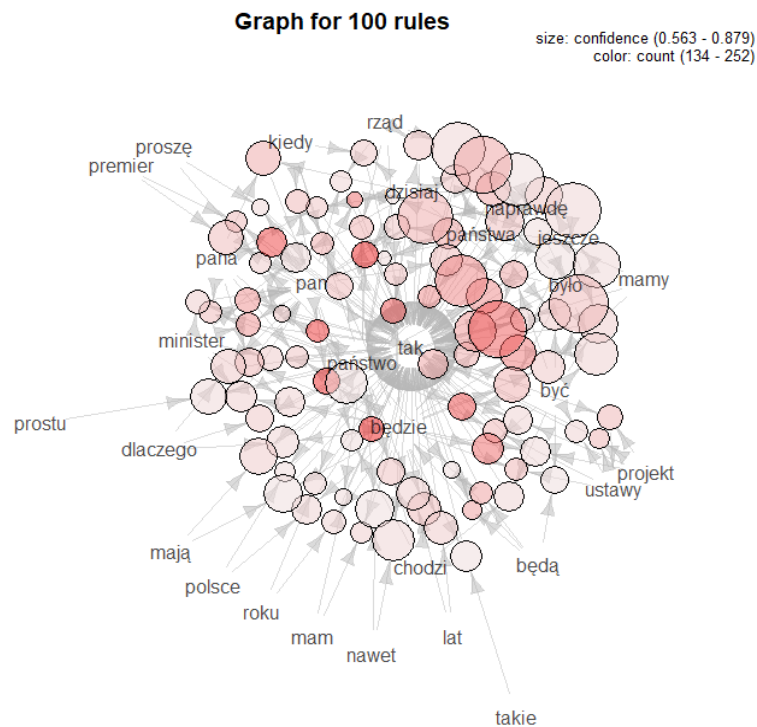
Widok ze względu na *Confidence* (minimum 2 elementy)

Jeśli chodzi o parametr *Confidence* największymi wartościami charakteryzują się połączenia: {przede} => {wszystkim}, {koalicja} => {obywatelska}, {unii} => {europejskiej}

Tabela 18. Widok ze względu na *Confidence* (min. 2 elementy) – KO

rank	rules	support	confidence	lift	count
1	{przede} => {wszystkim}	0,07	0,98	8,12	195
2	{koalicja} => {obywatelska}	0,05	0,97	16,61	139
3	{unii} => {europejskiej}	0,08	0,95	10,19	215
4	{wysoki} => {sejmie}	0,08	0,95	7,81	223
5	{obywatelska} => {koalicja}	0,05	0,90	16,61	139
6	{koalicji} => {obywatelskiej}	0,07	0,89	11,01	182
7	{morawiecki} => {premier}	0,06	0,88	6,29	153
8	{być, naprawdę} => {tak}	0,06	0,88	1,73	167
9	{będzie, naprawdę} => {tak}	0,08	0,87	1,72	224
10	{dzisiaj, naprawdę} => {tak}	0,07	0,87	1,70	186

Rysunek 90. Graf. Widok ze względu na *Confidence* (min. 3 elementy) – KO



Źródło: Opracowanie własne z wykorzystaniem biblioteki *arulesViz*.

Parametr *Confidence* ukazuje najsilniejsze powiązania dla 3 elementów dla różnych obiektów (*państwo*, *państwa*, *pan*, *być*) gdzie centralnym punktem jest słowo *tak*.

3.6.3. Lewica

Widok ze względu na *Lift* (minimum 2 elementy)

Jeśli chodzi o parametr *Lift* największymi wartościami charakteryzują się połączenia pomiędzy: $\{\text{unii}\} \Rightarrow \{\text{europejskiej}\}$, $\{\text{europejskiej}\} \Rightarrow \{\text{unii}\}$, $\{\text{polek}\} \Rightarrow \{\text{polaków}\}$

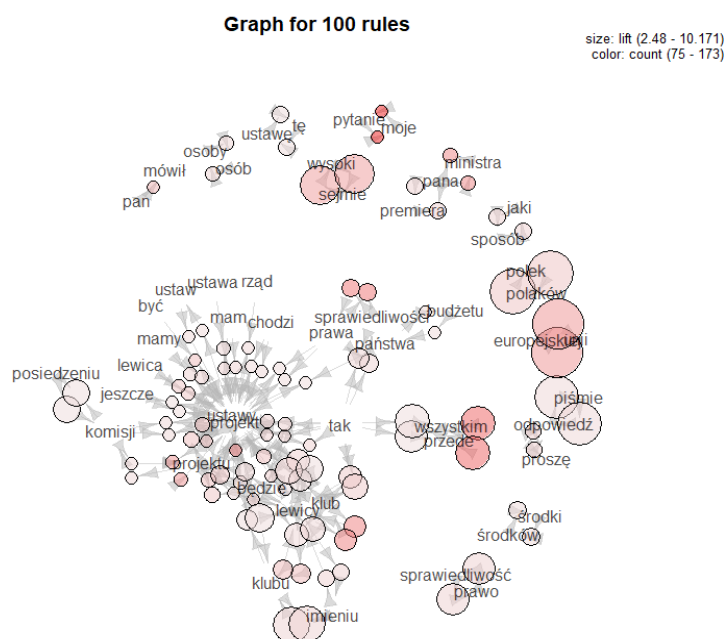
Tabela 19. Widok ze względu na *Lift* (min. 2 elementy) – Lewica

rank	rules	support	confidence	lift	count
1	$\{\text{unii}\} \Rightarrow \{\text{europejskiej}\}$	0,08	0,93	10,17	126
2	$\{\text{europejskiej}\} \Rightarrow \{\text{unii}\}$	0,08	0,92	10,17	126
3	$\{\text{polek}\} \Rightarrow \{\text{polaków}\}$	0,06	0,98	9,14	94
4	$\{\text{polaków}\} \Rightarrow \{\text{polek}\}$	0,06	0,59	9,14	94
5	$\{\text{piśmie}\} \Rightarrow \{\text{odpowiedź}\}$	0,05	0,77	8,69	79
6	$\{\text{odpowiedź}\} \Rightarrow \{\text{piśmie}\}$	0,05	0,59	8,69	79
7	$\{\text{wysoki}\} \Rightarrow \{\text{sejmie}\}$	0,08	0,92	7,88	123
8	$\{\text{sejmie}\} \Rightarrow \{\text{wysoki}\}$	0,08	0,71	7,88	123

9	{imieniu} => {klubu}	0,05	0,64	7,30	82
10	{klubu} => {imieniu}	0,05	0,63	7,30	82
11	{przede, tak} => {wszystkim}	0,05	0,99	6,79	76
12	{przede} => {wszystkim}	0,10	0,99	6,79	146
13	{wszystkim} => {przede}	0,10	0,67	6,79	146
14	{sprawiedliwość} => {prawo}	0,06	0,92	6,41	84
15	{prawo} => {sprawiedliwość}	0,06	0,39	6,41	84
16	{tak, wszystkim} => {przede}	0,05	0,63	6,39	76
17	{lewicy, ustawy} => {klubu}	0,05	0,51	5,77	76
18	{lewicy, projekt} => {klub}	0,05	0,57	5,43	76
19	{posiedzeniu} => {komisji}	0,05	0,82	5,40	75
20	{komisji} => {posiedzeniu}	0,05	0,33	5,40	75

Opracowanie własne z wykorzystaniem biblioteki arules.

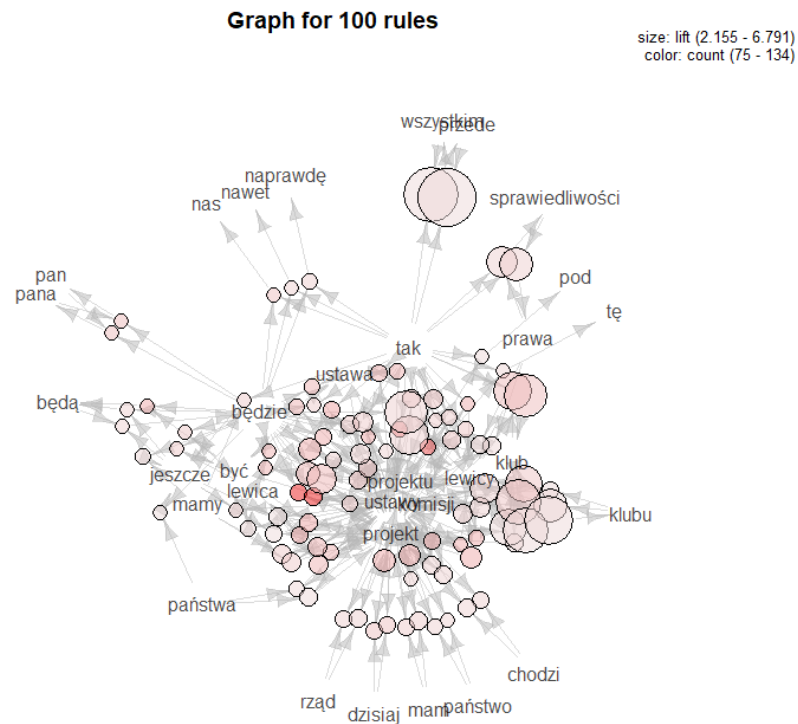
Rysunek 91. Graf. Widok ze względu na *Lift* (min. 2 elementy) – Lewica



Źródło: Opracowanie własne z wykorzystaniem biblioteki arulesViz.

Widok ze względu na parametr *Lift* pokazuje kilka punktów o największej powierzchni symbolizujący ww. parametr jak: *unii europejskiej*, *polec*, *polaków*, *odpowiedź piśmie*, *imieniu klubu* oraz *wysoki sejmie*.

Rysunek 92. Graf. Widok ze względu na *Lift* (min. 3 elementy) – Lewica



Źródło: Opracowanie własne z wykorzystaniem biblioteki *arulesViz*

W przypadku 3 elementowych widać największe obiekty na grafie to *tak*, *przede*, *wszystkim* oraz *lewicy*, *projekt*, *klubu*.

Widok ze względu na *Confidence* (minimum 2 elementy)

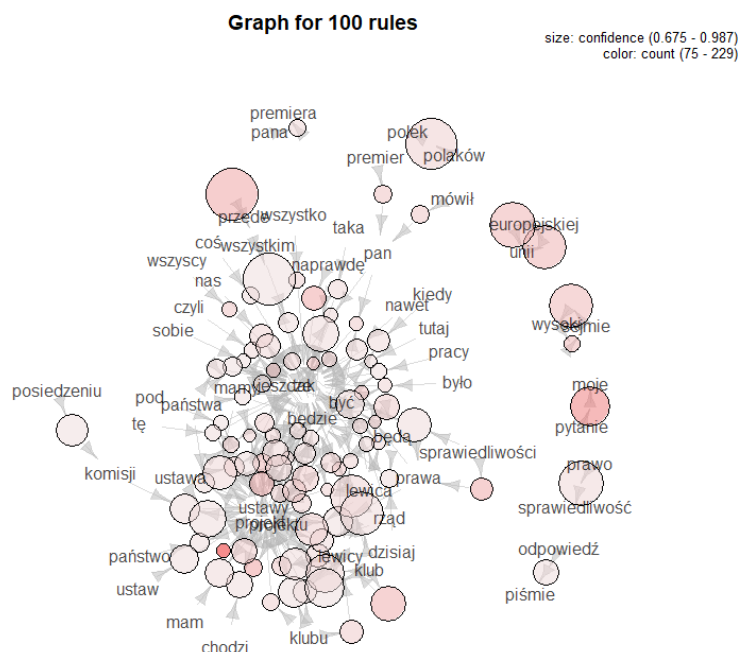
Jeśli chodzi o parametr *Confidence* największymi wartościami charakteryzują się połączenia: $\{\text{przede}, \text{tak}\} \Rightarrow \{\text{wszystkim}\}$, $\{\text{przede}\} \Rightarrow \{\text{wszystkim}\}$, $\{\text{polek}\} \Rightarrow \{\text{polaków}\}$

Tabela 20. Widok ze względu na *Confidence* (min. 2 elementy) – Lewica

rank	rules	support	confidence	lift	count
1	{przede, tak} => {wszystkim}	0,05	0,99	6,79	76
2	{przede} => {wszystkim}	0,10	0,99	6,79	146
3	{polek} => {polaków}	0,06	0,98	9,14	94
4	{unii} => {europejskiej}	0,08	0,93	10,17	126
5	{sprawiedliwość} => {prawo}	0,06	0,92	6,41	84
6	{europejskiej} => {unii}	0,08	0,92	10,17	126
7	{będzie, klub} => {lewicy}	0,05	0,92	4,75	80
8	{wysoki} => {sejmie}	0,08	0,92	7,88	123
9	{klub, tak} => {lewicy}	0,06	0,91	4,69	89
10	{klub, projekt} => {lewicy}	0,05	0,88	4,57	76
11	{moje} => {pytanie}	0,12	0,88	2,55	173
12	{klub, ustawy} => {lewicy}	0,06	0,88	4,53	85
13	{projekt, ustawa} => {ustawy}	0,05	0,87	2,79	79
14	{będzie, naprawdę} => {tak}	0,05	0,86	2,03	80
15	{klub} => {lewicy}	0,09	0,85	4,41	134
16	{mamy, projekt} => {ustawy}	0,06	0,85	2,73	83
17	{sprawiedliwości, tak} => {prawa}	0,05	0,84	4,12	81
18	{lewica, projekt} => {ustawy}	0,06	0,84	2,70	93
19	{lewicy, projektu} => {ustawy}	0,05	0,82	2,65	80
20	{posiedzeniu} => {komisji}	0,05	0,82	5,40	75

Opracowanie własne z wykorzystaniem biblioteki arules

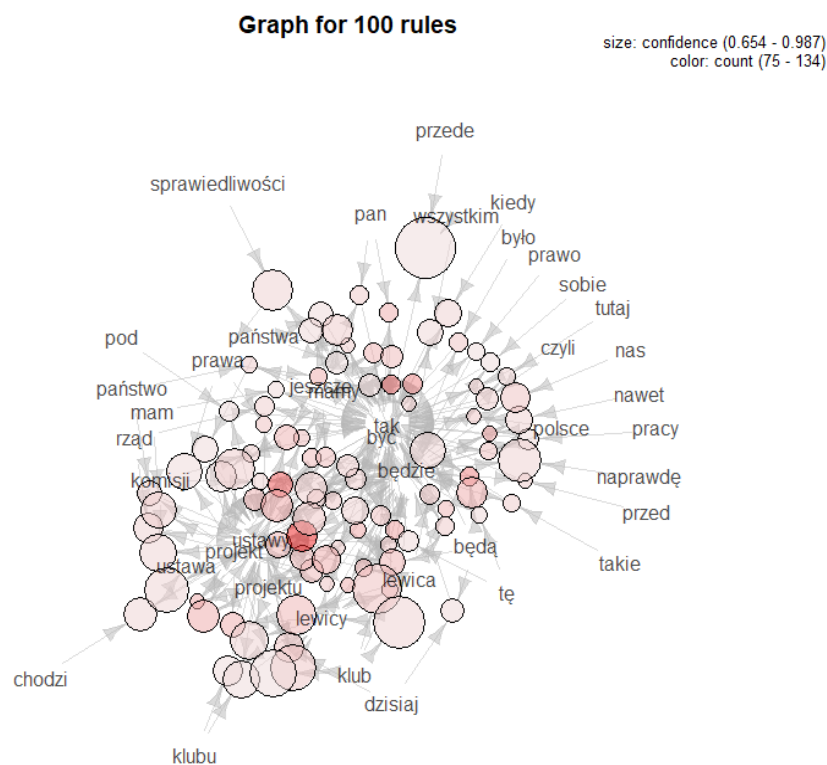
Rysunek 93. Widok ze względu na *Confidence* (min. 2 elementy) – Lewica



Źródło: Opracowanie własne z wykorzystaniem biblioteki arulesViz

Ze względu na *Confidence*, największe wartości prezentują zwroty *przede wszystkim*, *polek*, *polaków*, *unii europejskiej*, *prawo*, *sprawiedliwość* oraz zwrotu *wysoki sejmie*.

Rysunek 94. Widok ze względu na *Confidence* (min. 3 elementy) – Lewica



Źródło: Opracowanie własne z wykorzystaniem biblioteki *arulesViz*

Zwroty 3 elementowe największe wartości prezentują *tak*, *przed*, *wszystkim* oraz *będzie*, *klub*, *lewicy*.

3.6.4. Koalicja Polska (KP)

Widok ze względu na *Lift* (minimum 2 elementy)

Jeśli chodzi o parametr *Lift* największymi wartościami charakteryzują się połączenia pomiędzy: {polskiego, stronnictwa} => {ludowego}, {stronnictwa} => {ludowego}, {ludowego} => {stronnictwa}

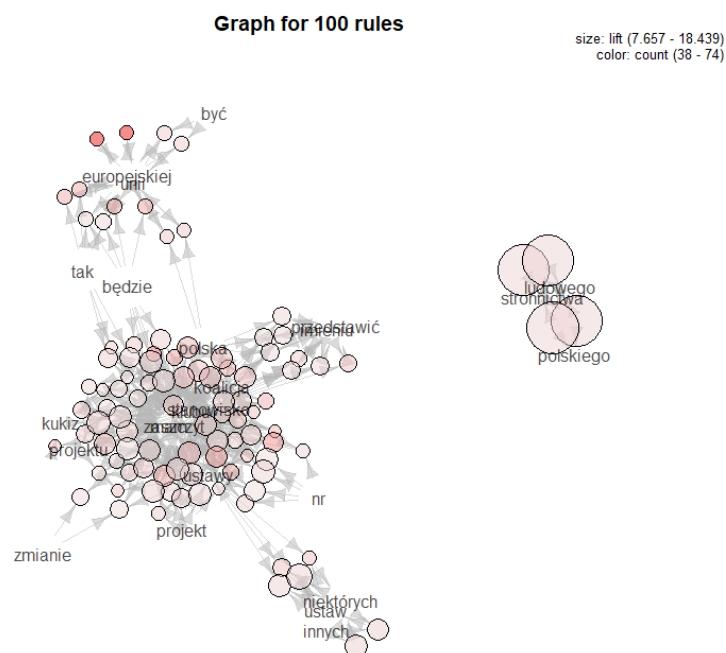
Tabela 21. Widok ze względu na *Lift* (min. 2 elementy) – KP

rank	rules	support	confidence	lift	count
1	{polskiego, stronnictwa} => {ludowego}	0,05	1,00	18,44	40
2	{stronnictwa} => {ludowego}	0,05	0,98	17,99	40
3	{ludowego} => {stronnictwa}	0,05	0,98	17,99	40
4	{ludowego, polskiego} => {stronnictwa}	0,05	0,98	17,99	40

5	{innych, niektórych, ustawy} => {ustaw}	0,05	0,95	11,08	40
6	{mam, nr, stanowisko} => {zaszczyt}	0,05	0,95	10,75	40
7	{kukiz, mam, stanowisko} => {zaszczyt}	0,06	0,93	10,53	42
8	{klubu, kukiz, mam, stanowisko} => {zaszczyt}	0,05	0,93	10,50	40
9	{klubu, mam, projekt, stanowisko} => {zaszczyt}	0,05	0,93	10,50	40
10	{klubu, mam, projekt, stanowisko, ustawy} => {zaszczyt}	0,05	0,93	10,48	39
11	{klubu, mam, nr} => {zaszczyt}	0,05	0,93	10,46	38
12	{kukiz, mam, stanowisko, ustawy} => {zaszczyt}	0,05	0,93	10,46	38
13	{mam, projekt, stanowisko} => {zaszczyt}	0,05	0,91	10,28	41
14	{mam, projekt, stanowisko, ustawy} => {zaszczyt}	0,05	0,91	10,26	40
15	{innych, ustaw} => {niektórych}	0,05	0,98	10,25	41
16	{innych, ustaw, ustawy} => {niektórych}	0,05	0,98	10,24	40
17	{będzie, klubu, mam, stanowisko} => {zaszczyt}	0,05	0,91	10,23	39
18	{będzie, mam, stanowisko, ustawy} => {zaszczyt}	0,05	0,91	10,23	39
19	{klubu, mam, stanowisko, ustawy} => {zaszczyt}	0,06	0,91	10,22	48
20	{klubu, mam, polska, stanowisko} => {zaszczyt}	0,06	0,90	10,20	47

Opracowanie własne z wykorzystaniem biblioteki *arules*

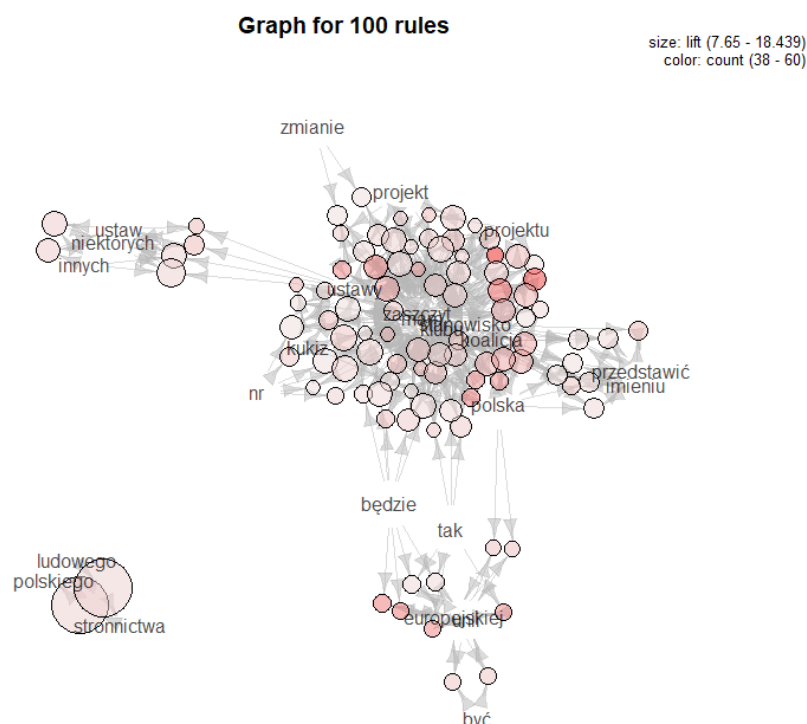
Rysunek 95. Widok ze względu na *Lift* (min. 2 elementy) – KP



Źródło: Opracowanie własne z wykorzystaniem biblioteki *arulesViz*

Nieporównywalnie duże wartości w stosunku do innych powiązań dotyczy nazwy partii: stronnictwa ludowego.

Rysunek 96. Widok ze względu na *Lift* (min. 3 elementy) – KP



Źródło: Opracowanie własne z wykorzystaniem biblioteki *arulesViz*

Dla minimum 3 elementowych powiązań również dotyczą nazwy partii – *polskiego stronnictwa ludowego*.

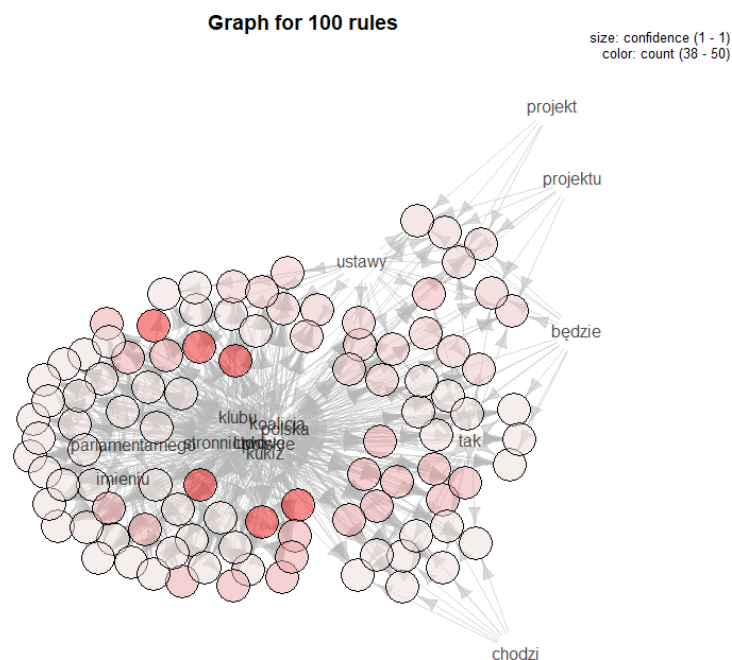
Widok ze względu na *Confidence* (minimum 2 elementy)

Jeśli chodzi o parametr *Confidence* największymi wartościami (1) charakteryzują się połączenia w znacznej liczbie, do najczęściej występujących należą: $\{koalicja\} \Rightarrow \{polska\}$, $\{koalicja, tak\} \Rightarrow \{polska\}$, $\{klubu, koalicja\} \Rightarrow \{polska\}$

Tabela 22. Widok ze względu na *Confidence* (min. 2 elementy) – KP

rank	rules	support	confidence	lift	count
1	$\{koalicja\} \Rightarrow \{polska\}$	0,26	1,00	2,94	199
2	$\{koalicja, tak\} \Rightarrow \{polska\}$	0,19	1,00	2,95	143
3	$\{klubu, koalicja\} \Rightarrow \{polska\}$	0,18	1,00	2,95	136
4	$\{będzie, koalicja\} \Rightarrow \{polska\}$	0,17	1,00	2,95	131
5	$\{koalicja, ustawy\} \Rightarrow \{polska\}$	0,16	1,00	2,95	123
6	$\{koalicja, kukiz\} \Rightarrow \{polska\}$	0,15	1,00	2,95	117
7	$\{koalicja, polskie\} \Rightarrow \{polska\}$	0,15	1,00	2,95	116
8	$\{stronnictwo\} \Rightarrow \{ludowe\}$	0,15	1,00	6,69	113

Rysunek 98. Widok ze względu na *Confidence* (min. 3 elementy) – KP



Źródło: Opracowanie własne z wykorzystaniem biblioteki arulesViz

Podobnie jak w poprzednim przypadku nie ma znaczących odstępstw i zwroty, które są najsilniejsze dotyczą nazw partii oraz projektu ustawy.

3.6.5. Konfederacja

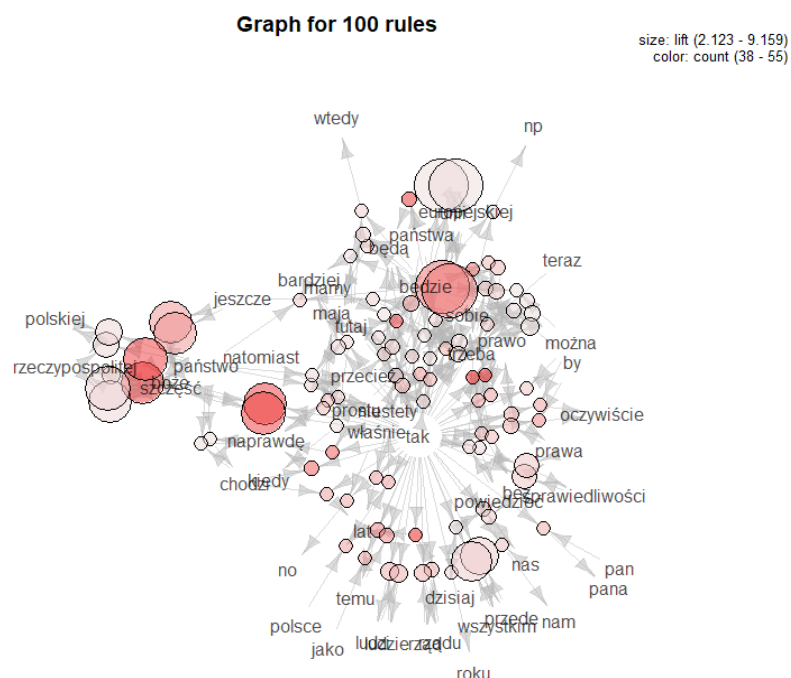
Widok ze względu na *Lift* (minimum 2 elementy)

Jeśli chodzi o parametr *Lift* największymi wartościami charakteryzują się połączenia pomiędzy: $\{facto\} \Rightarrow \{de\}$, $\{de\} \Rightarrow \{facto\}$, $\{europejskiej, tak\} \Rightarrow \{unii\}$

Tabela 23. Widok ze względu na *Lift* (min. 2 elementy) – KP

rank	rules	support	confidence	lift	count
1	{facto} => {de}	0,05	1,00	17,20	41
2	{de} => {facto}	0,05	0,93	17,20	41
3	{europejskiej, tak} => {unii}	0,06	0,91	9,16	49
4	{europejskiej} => {unii}	0,09	0,89	9,03	68
5	{unii} => {europejskiej}	0,09	0,91	9,03	68
6	{będzie, unii} => {europejskiej}	0,05	0,90	9,01	38

Rysunek 100. Widok ze względu na *Lift* (min. 3 elementy) – Konfederacja



Źródło: Opracowanie własne z wykorzystaniem biblioteki *arulesViz*

W przypadku związków trzejelementowych najsilniejsze z nich to *tak* lub *będzie*, *unii*, *europejskiej*, oraz zwrotu *szczęść boże*⁴⁶⁶ *rzeczpospolitej*.

Widok ze względu na *Confidence* (minimum 2 elementy)

Jeśli chodzi o parametr *Confidence* największymi wartościami (1) charakteryzują się połączenia w znacznej liczbie, do najczęściej występujących należą: {*szczęść*} => {*boże*}, {*państwo*, *szczęść*} => {*boże*}, {*szczęść*, *tak*} => {*boże*}.

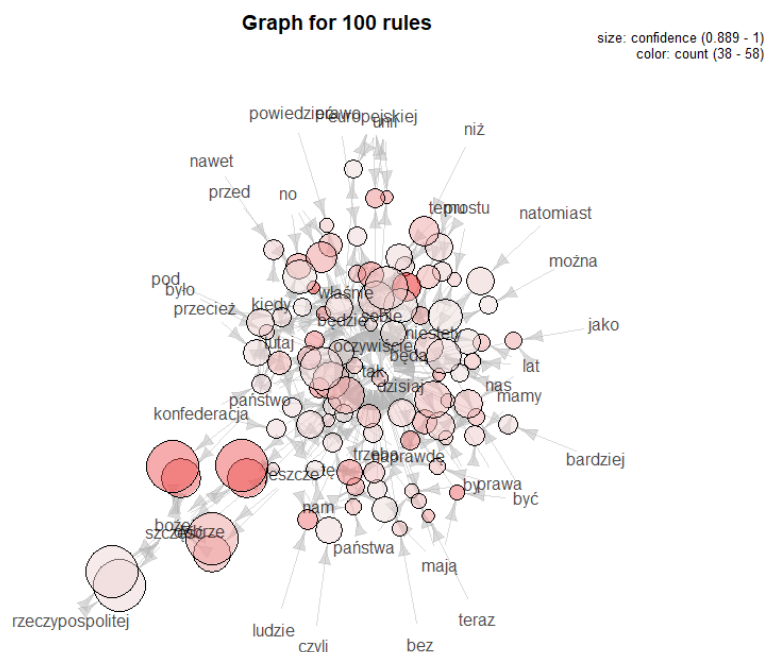
Tabela 24. Widok ze względu na *Confidence* (min. 2 elementy) – KP

rank	rules	support	confidence	lift	count
1	{facto} => {de}	0,05	1,00	17,20	41
2	{boże, rzeczpospolitej} => {szczęść}	0,05	1,00	7,42	39
3	{szczęść} => {boże}	0,13	1,00	7,07	102

⁴⁶⁶ Jest to charakterystyczne powitanie jednego z najbardziej aktywnych posłów – Grzegorza Brauna.

Ze względu na parametr *Confidence* i związki dwuelementowe to dotyczą one zwrotu *de facto* oraz *szcześć boże* z kombinacją słowa *rzeczypospolitej*.

Rysunek 102. Widok ze względu na *Confidence* (min. 3 elementy) – Konfederacja



Źródło: Opracowanie własne z wykorzystaniem biblioteki arulesViz

W przypadku trzejelementowych powiązań na parametr *Confidence* podobnie jak w poprzednim przypadku zwroty *szczęść boże Rzeczypospolitej*, oraz *szczęść boże konfederacja* prezentują najwyższe wartości w stosunku do innych.

3.6.6. Interpretacja wyników

Każda partia charakteryzuje się swoimi rodzajami połączeń i posiada swoje wzorce, jeśli chodzi o użycie słów w różnych kombinacjach, ale biorąc pod uwagę badanie wspólnej części wszystkich partii z wykorzystaniem analizy koszykowej i 20 najczęstszych połączeń, ponadto biorąc pod uwagę parametry jak *Confidence* i *Lift*, można stwierdzić, że istnieją pewne części wspólne. 4 partie często używały zwrotu *przede wszystkim*; 3 partie wspominały o *unii europejskiej*; oraz również posłowie z dwóch partii często stosowali zwrot *wysoki sejmie*.

- {przede} => {wszystkim} – Partie: PiS, KO, Lewica, Konfederacja
- {europejskiej} => {unii} – Partie: KO, Lewica, Konfederacja

- {unii} => {europejskiej} – Partie: KO, Lewica, Konfederacja
- {wysoki} => {sejmie} – Partie: KO, Lewica

Każda partia ma w swoim słowniku powiązania słów dotyczący ich nazw. Partia Koalicja Polska oprócz powtarzających się kombinacji nazwy partii nie mają w tej analizie charakterystycznych wzorców oprócz odwołania do Unii Europejskiej oraz projektu ustawy. Lewica również odwołuje się do Unii Europejskiej i nazwy partii, ale widać wzorzec w postaci zwrotu *wysoki sejmie* oraz *polek* i *polaków*. Ciekawą przypadłość widać w przypadku Koalicji Obywatelskiej i PiS: KO stosuje wobec swojej partii nazwę właściwą koalicyjną (*koalicja obywatelska*), ale PiS nawiązuje do nazwy: *platforma obywatelska*.

3.7. Analiza ngramów

W tej części z wykorzystaniem biblioteki *tidytext*⁴⁶⁷ skupiono się na analizie słów najczęściej współwystępujących w wypowiedziach polityków i przedstawiono je jako zestawienie elementów dwuczłściowych (*bigramy*), trzyczłściowych (*trigramy*), oraz czteroczłściowych. Wybrano najczęściej występujące i za pośrednictwem pakietu *ggplot2* stworzono wykresy prezentujące wspomniane dane z podziałem na poszczególne partie. Analiza miała na celu sprawdzenie wzorców słowników n-elementowych kombinacji słów wykorzystywanych przez polityków z poszczególnych formacji.

3.7.1. Prawo i Sprawiedliwość (PiS)

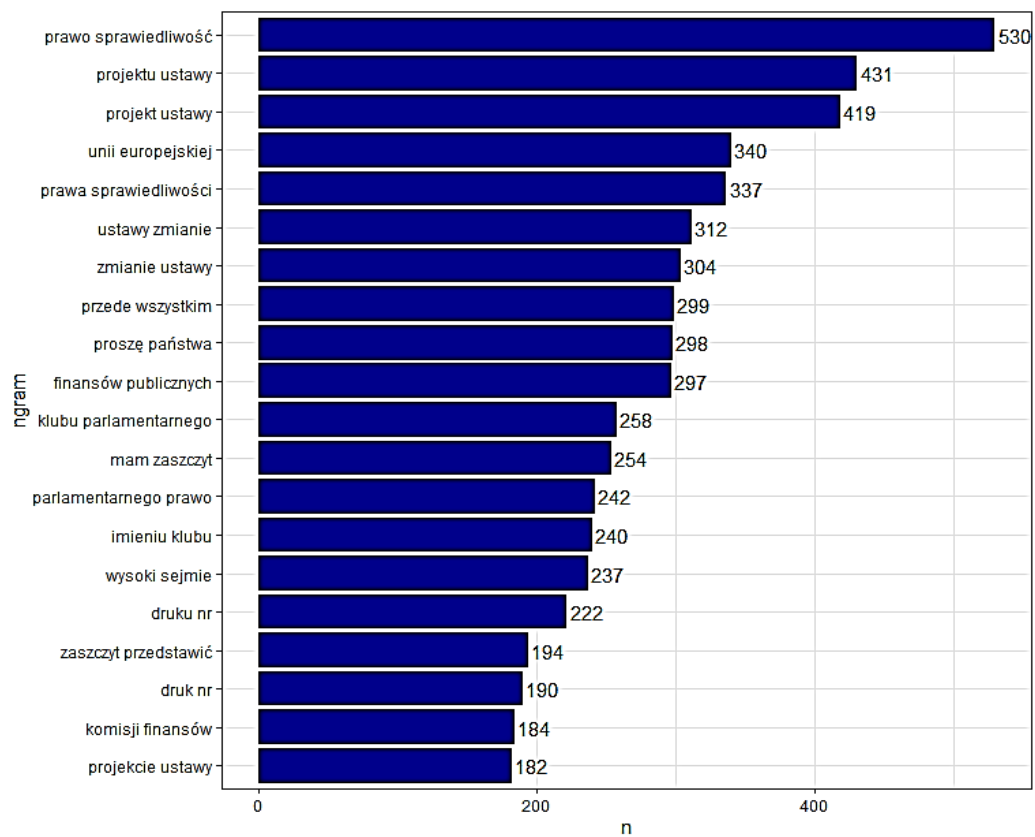
Do najczęściej występujących dwu-wyrazowych należą: *prawo sprawiedliwość*, *projekt ustawy*, *projekt ustawy*.

Do najczęściej występujących trzy-wyrazowych należą: *ustawy zmianie ustawy*, *parlamentarnego prawo sprawiedliwość*, *klubu parlamentarnego prawo*.

Do najczęściej występujących cztero-wyrazowych należą: *klubu parlamentarnego prawo sprawiedliwość*, *imieniu klubu parlamentarnego prawo*, *projektu ustawy zmianie ustawy*.

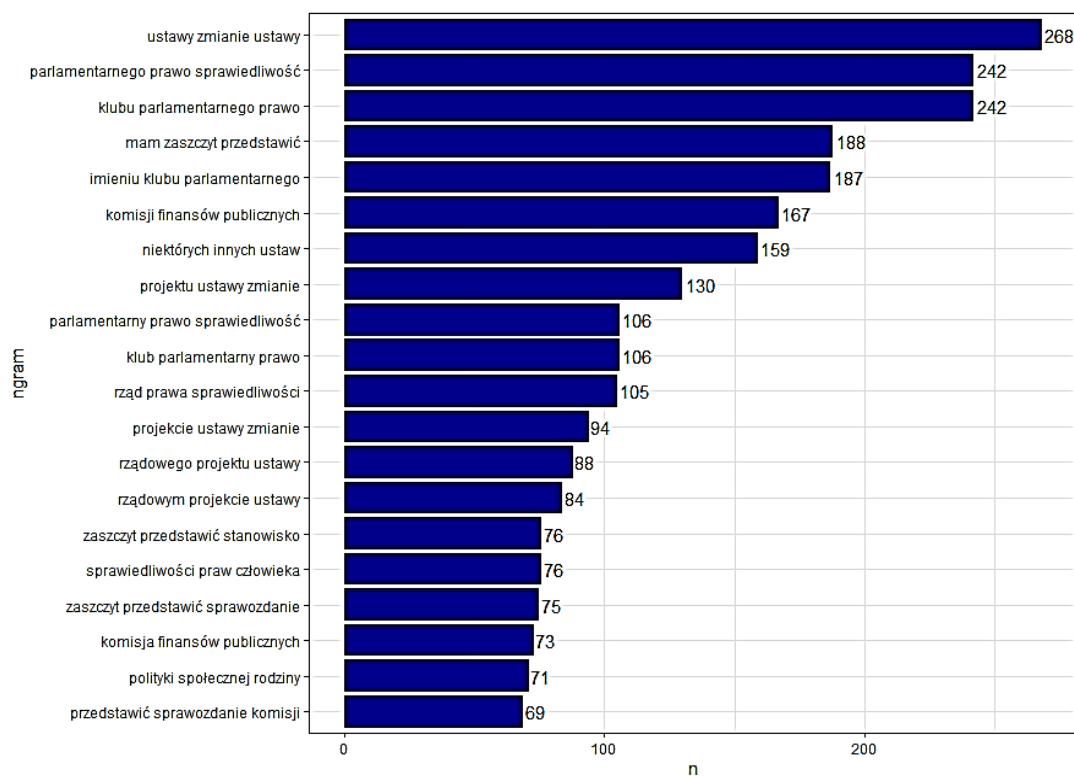
⁴⁶⁷Biblioteka *tidytext*, <https://www.rdocumentation.org/packages/tidytext/versions/0.3.2>, 29.12.2021.

Rysunek 103. Najczęściej występujące wyrazy 2-częściowe. Partia: PiS



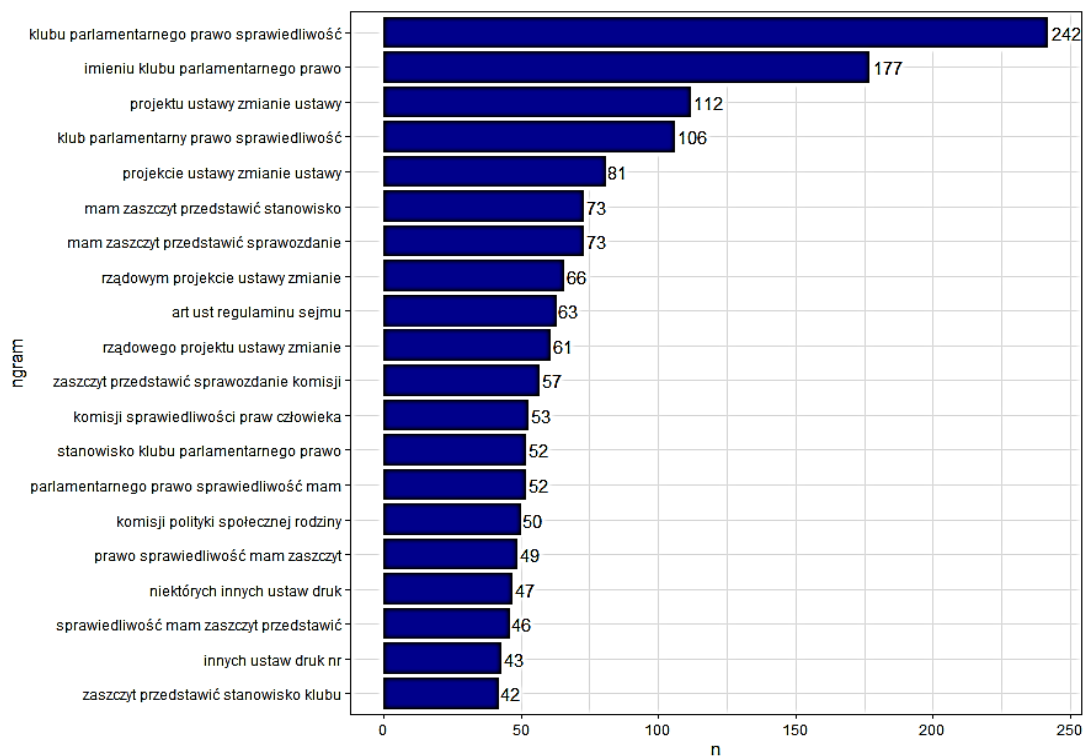
Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 104. Najczęściej występujące wyrazy 3-częściowe. Partia: PiS



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 105. Najczęściej występujące wyrazy 4-częściowe. Partia: PiS



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

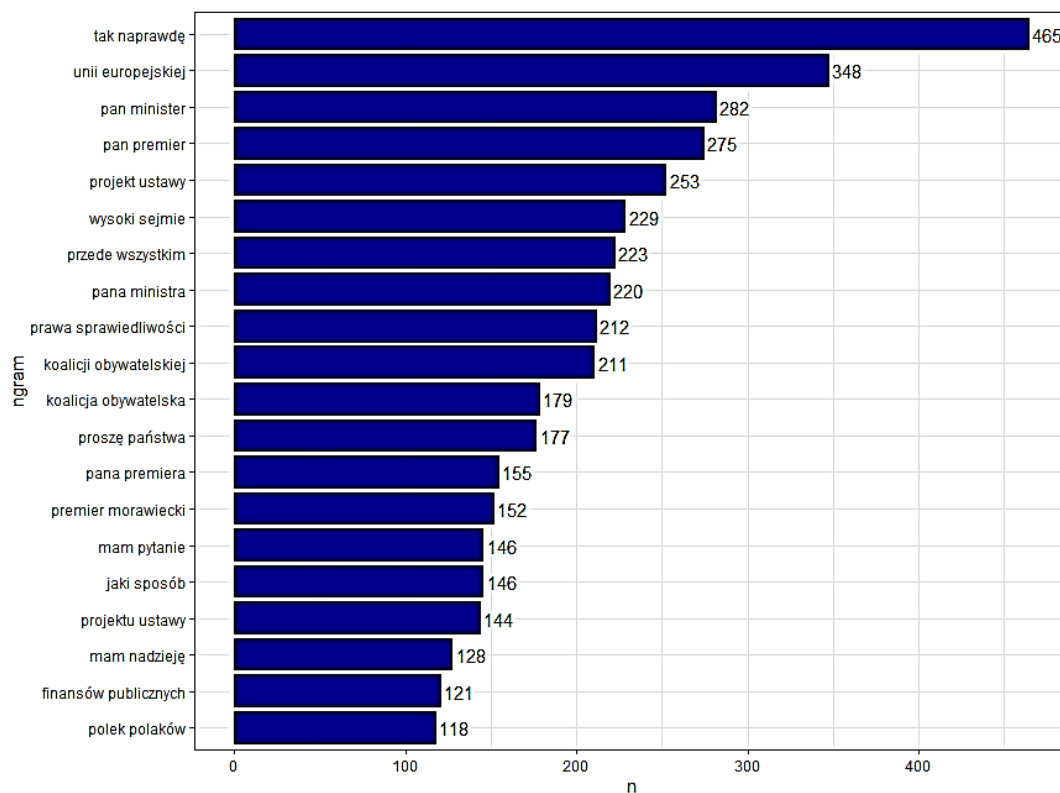
3.7.2. Koalicja Obywatelska (KO)

Do najczęściej występujących dwu-wyrazowych należą: *tak naprawdę, unii europejskiej, pan minister.*

Do najczęściej występujących trzy-wyrazowych należą: *pan premier morawiecki, klub koalicji obywatelskiej, klubu koalicji obywatelskiej.*

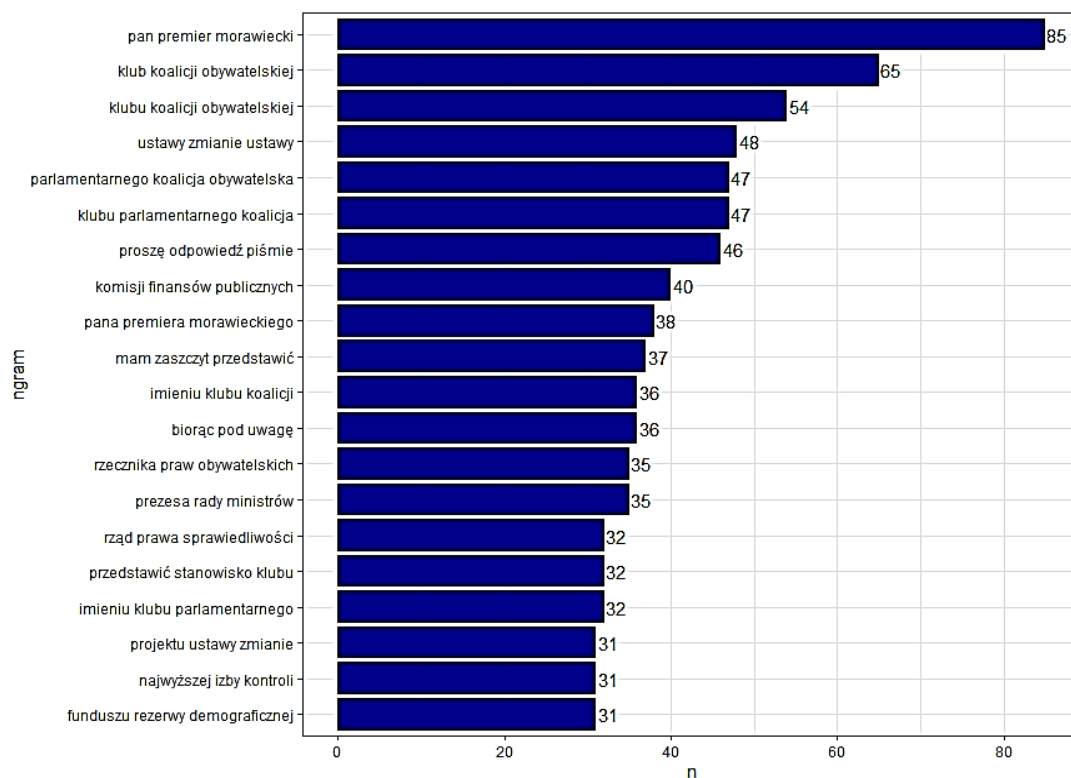
Do najczęściej występujących cztero-wyrazowych należą: *klubu parlamentarnego koalicja obywatelska, imieniu klubu koalicji obywatelskiej, klub parlamentarny koalicja obywatelska.*

Rysunek 106. Najczęściej występujące wyrazy 2-częściowe. Partia: KO



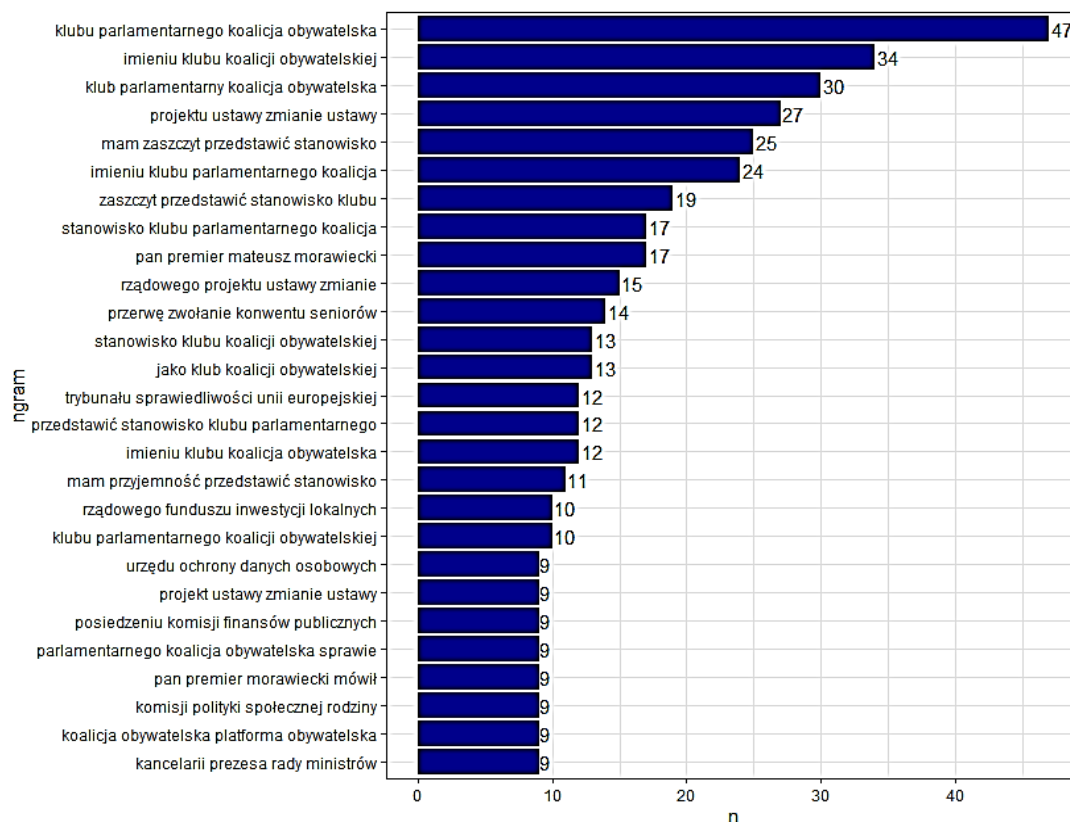
Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 107. Najczęściej występujące wyrazy 3-częściowe. Partia: KO



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 108. Najczęściej występujące wyrazy 4-częściowe. Partia: KO



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

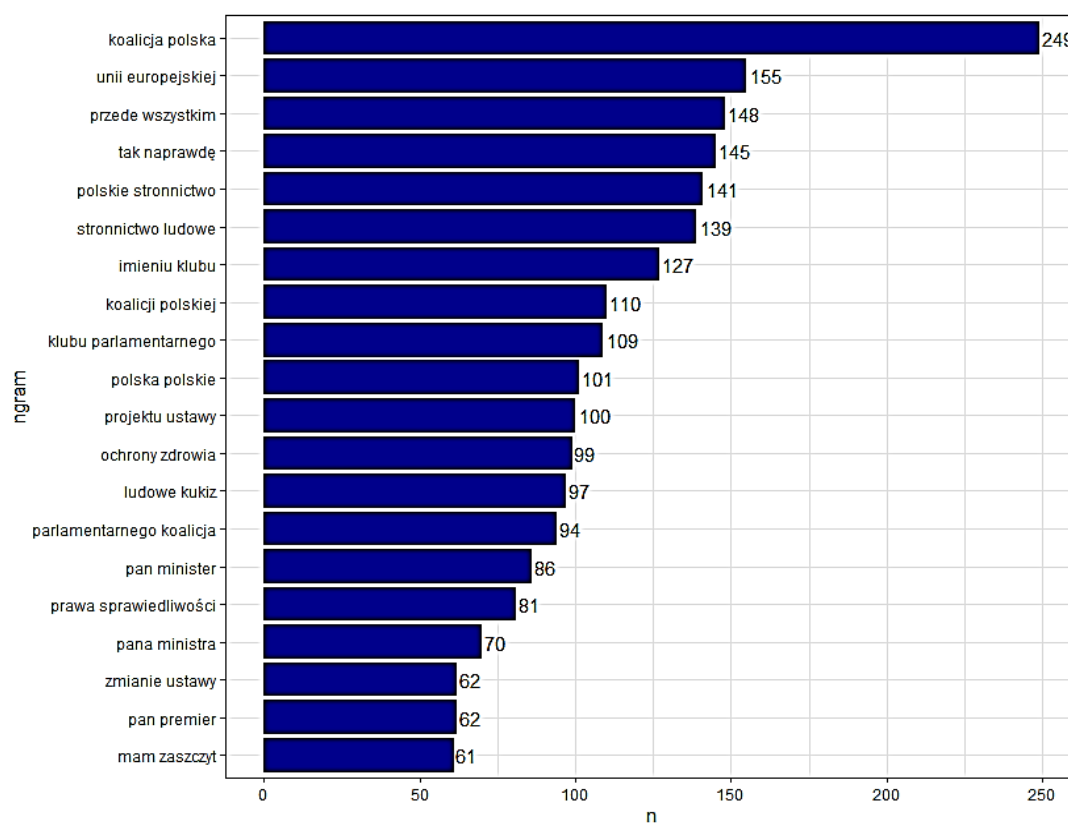
3.7.3. Koalicja Polska (KP)

Do najczęściej występujących dwu-wyrazowych należą: *koalicja polska*, *unii europejskiej*, *przede wszystkim*.

Do najczęściej występujących trzy-wyrazowych należą: *polskie stronnictwo ludowe*, *polska polskie stronnictwo*, *koalicja polska polskie*.

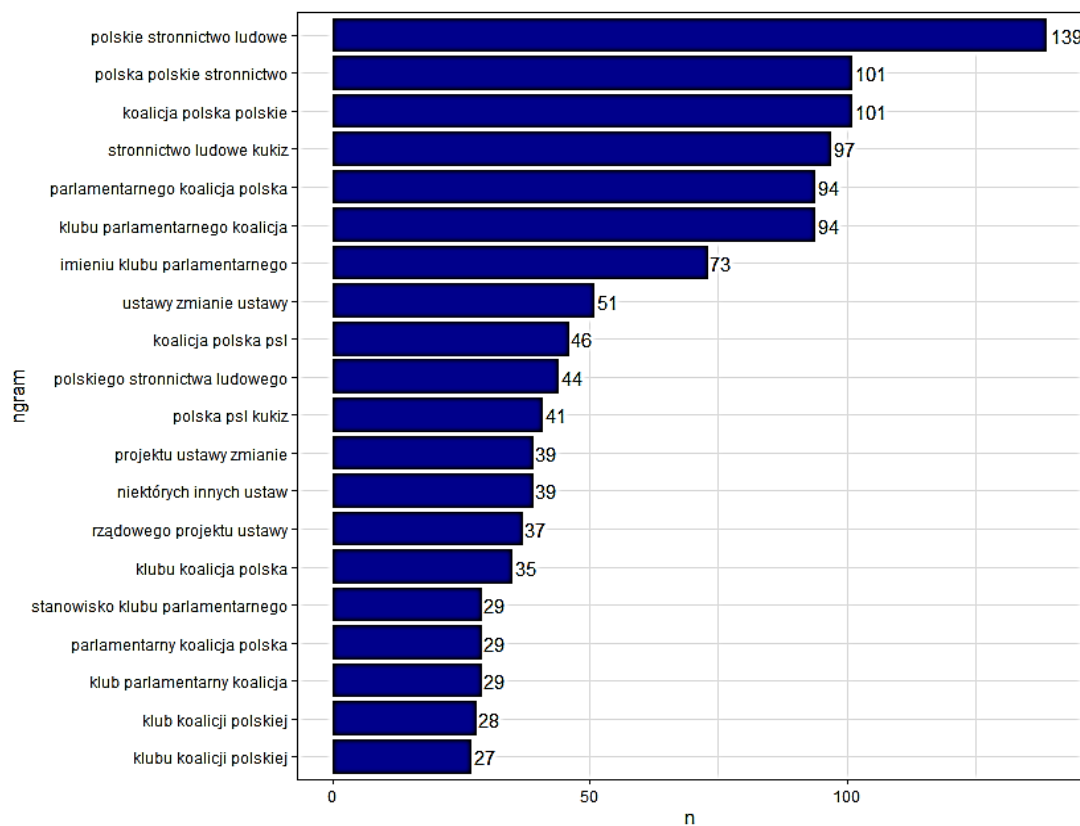
Do najczęściej występujących cztero-wyrazowych należą: *koalicja polska polskie stronnictwo*, *polska polskie stronnictwo ludowe*, *polskie stronnictwo ludowe kukiz*.

Rysunek 109. Najczęściej występujące wyrazy 2-częściowe. Partia: KP



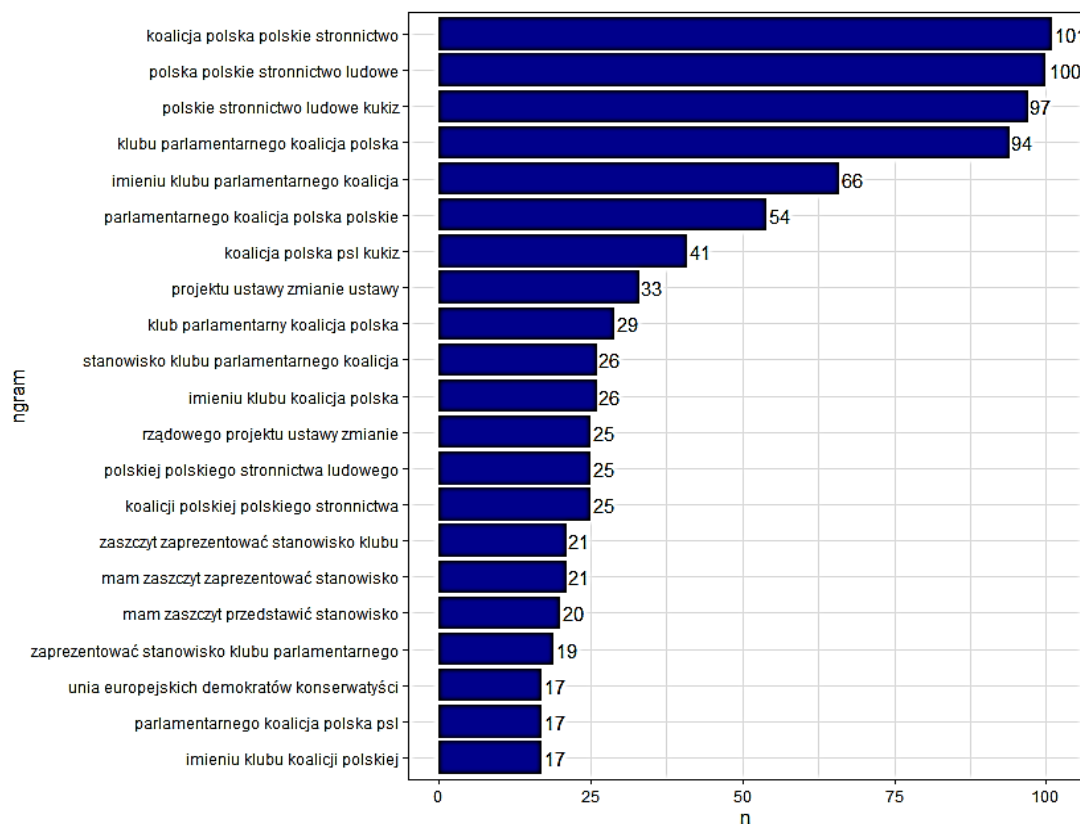
Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 110. Najczęściej występujące wyrazy 3-częściowe. Partia: KP



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 111. Najczęściej występujące wyrazy 4-częściowe. Partia: KP



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

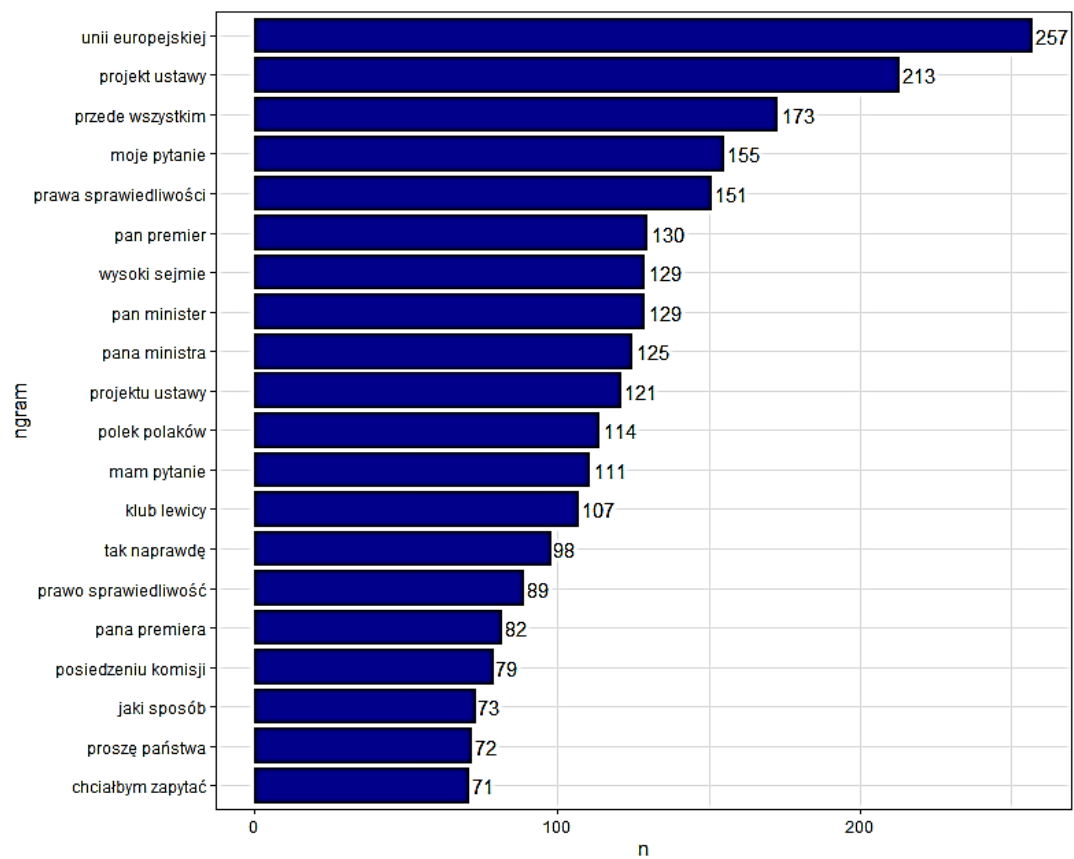
3.7.4. Lewica

Do najczęściej występujących dwu-wyrazowych należą: *unii europejskiej, projekt ustawy, przede wszystkim*.

Do najczęściej występujących trzy-wyrazowych należą: *ustawy zmianie ustawy, proszę odpowiedź piśmie, moje pytanie takie*.

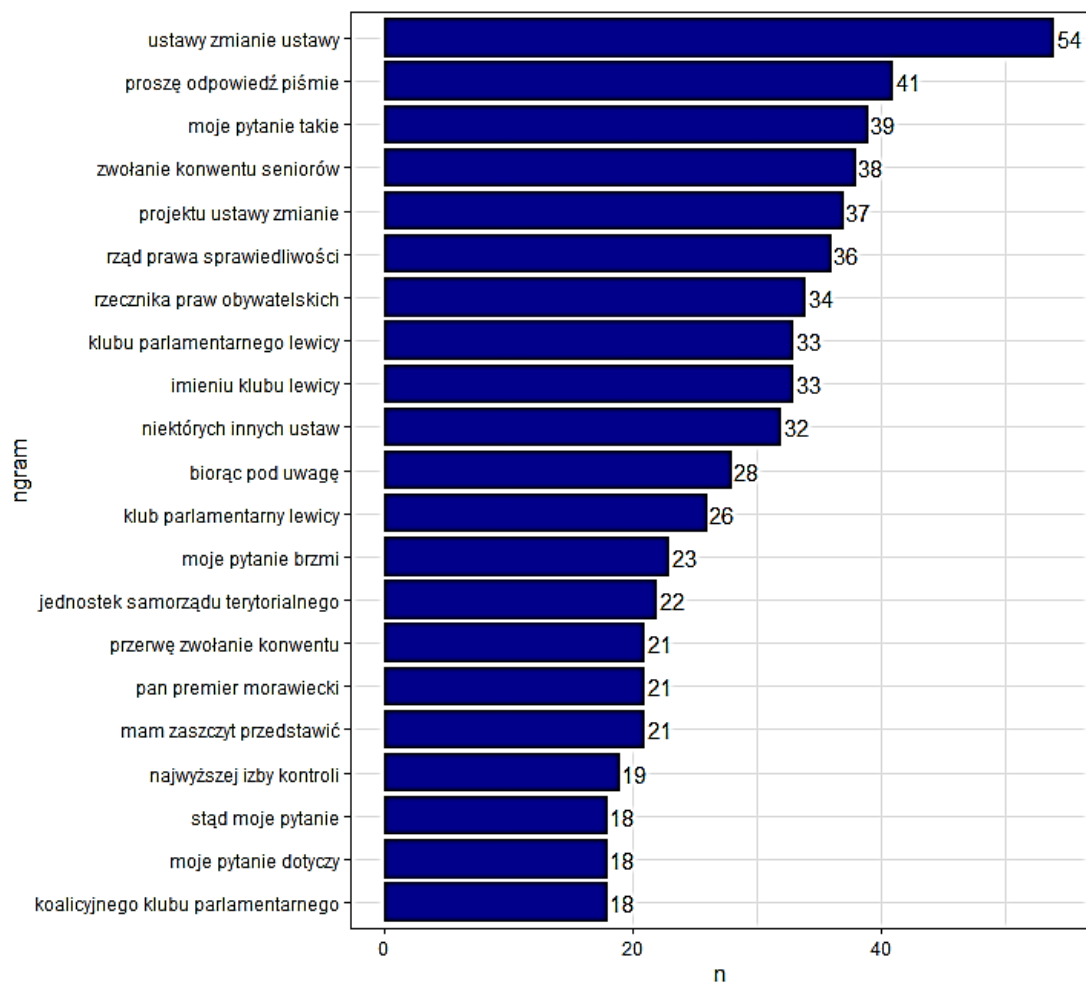
Do najczęściej występujących cztero-wyrazowych należą: *projektu ustawy zmianie ustawy, przerwę zwołanie konwentu seniorów, komisji sprawiedliwości praw człowieka*.

Rysunek 112. Najczęściej występujące wyrazy 2-częściowe. Partia: Lewica



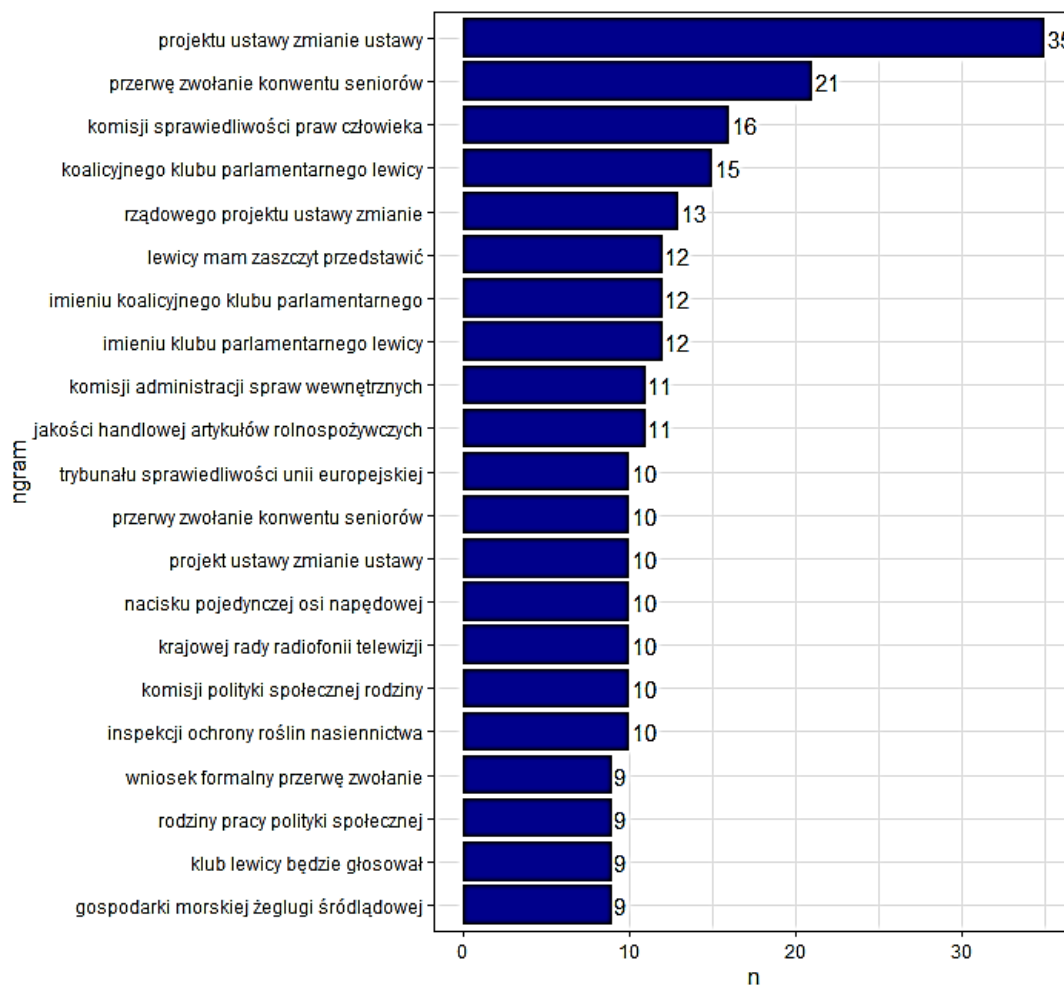
Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 113. Najczęściej występujące wyrazy 3-częściowe. Partia: Lewica



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 114. Najczęściej występujące wyrazy 4-częściowe. Partia: Lewica



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

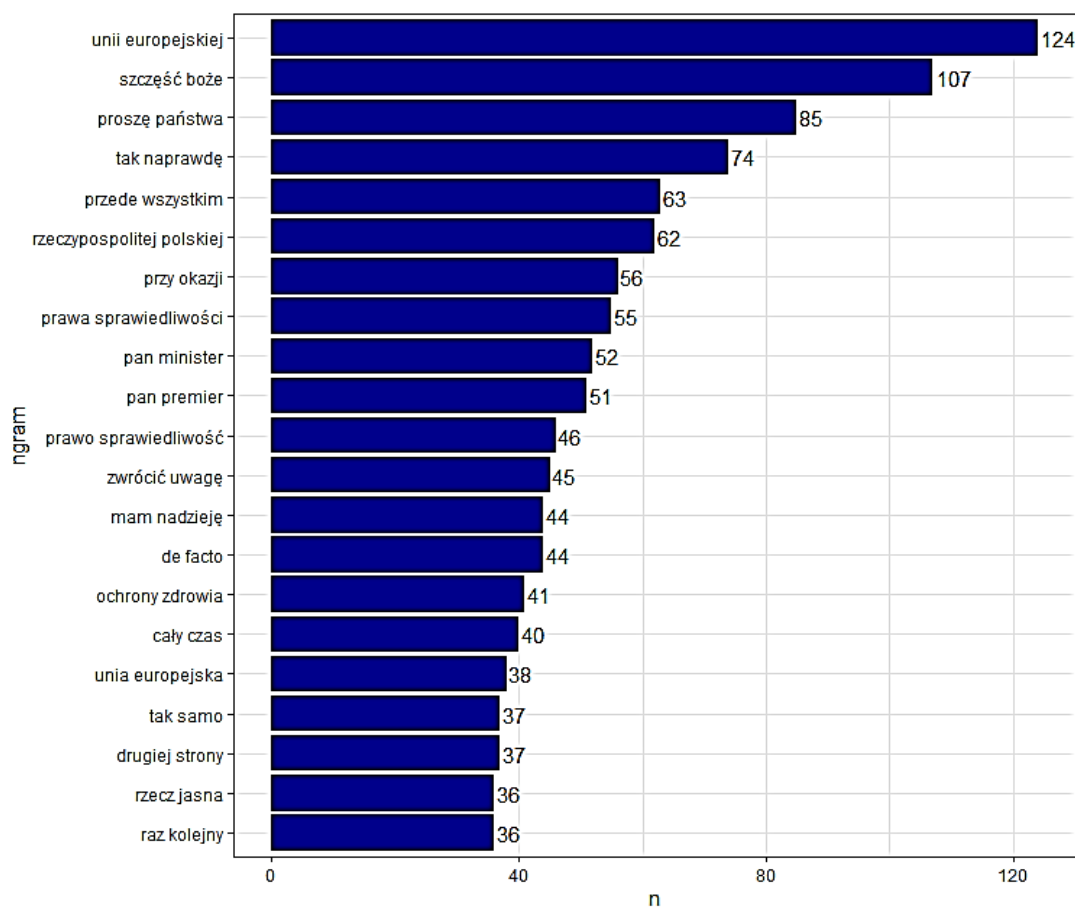
3.7.5. Konfederacja

Do najczęściej występujących dwu-wyrazowych należą: *unii europejskiej, szczęść boże, proszę państwa.*

Do najczęściej występujących trzy-wyrazowych należą: *zwołania konwentu seniorów, celem zwołania konwentu, sejm rzeczpospolitej polskiej.*

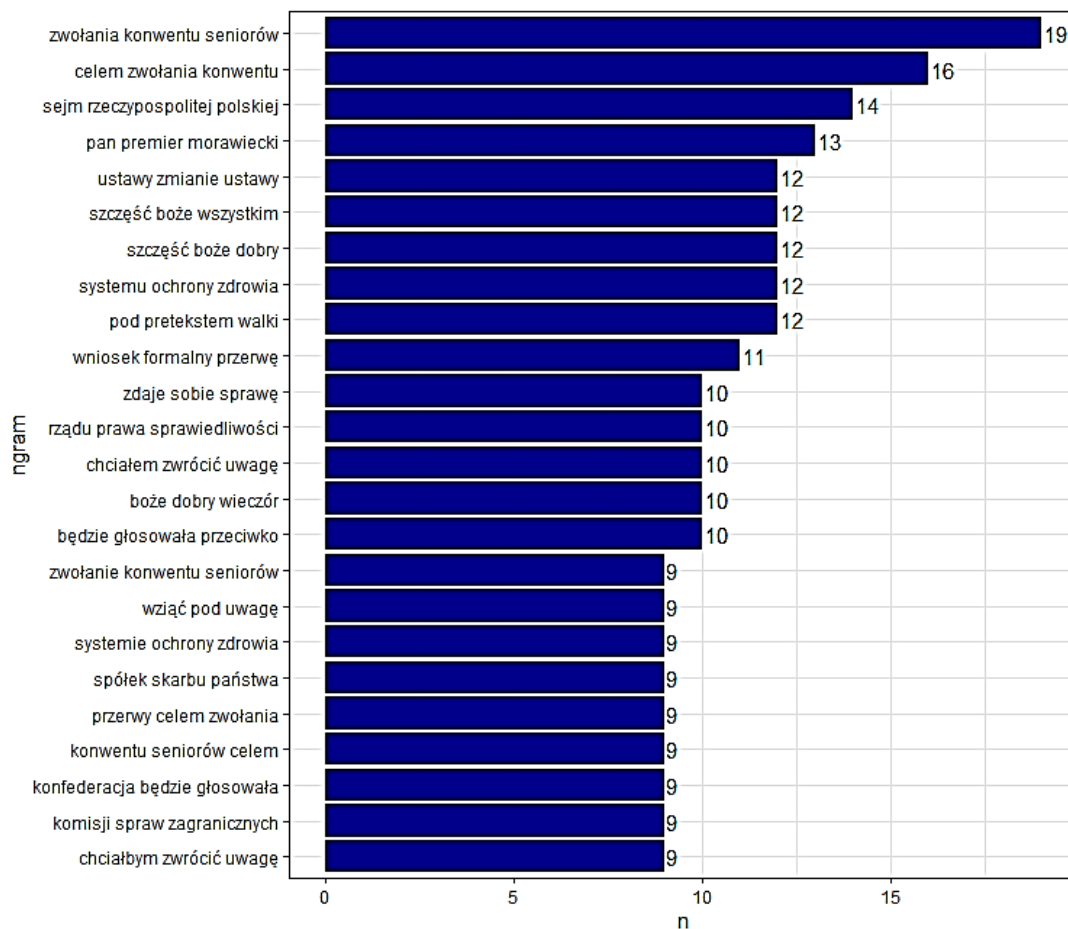
Do najczęściej występujących cztero-wyrazowych należą: *celem zwołania konwentu seniorów, szczęść boże dobry wieczór, przerwy celem zwołania konwentu.*

Rysunek 115. Najczęściej występujące wyrazy 2-częściowe. Partia: Konfederacja



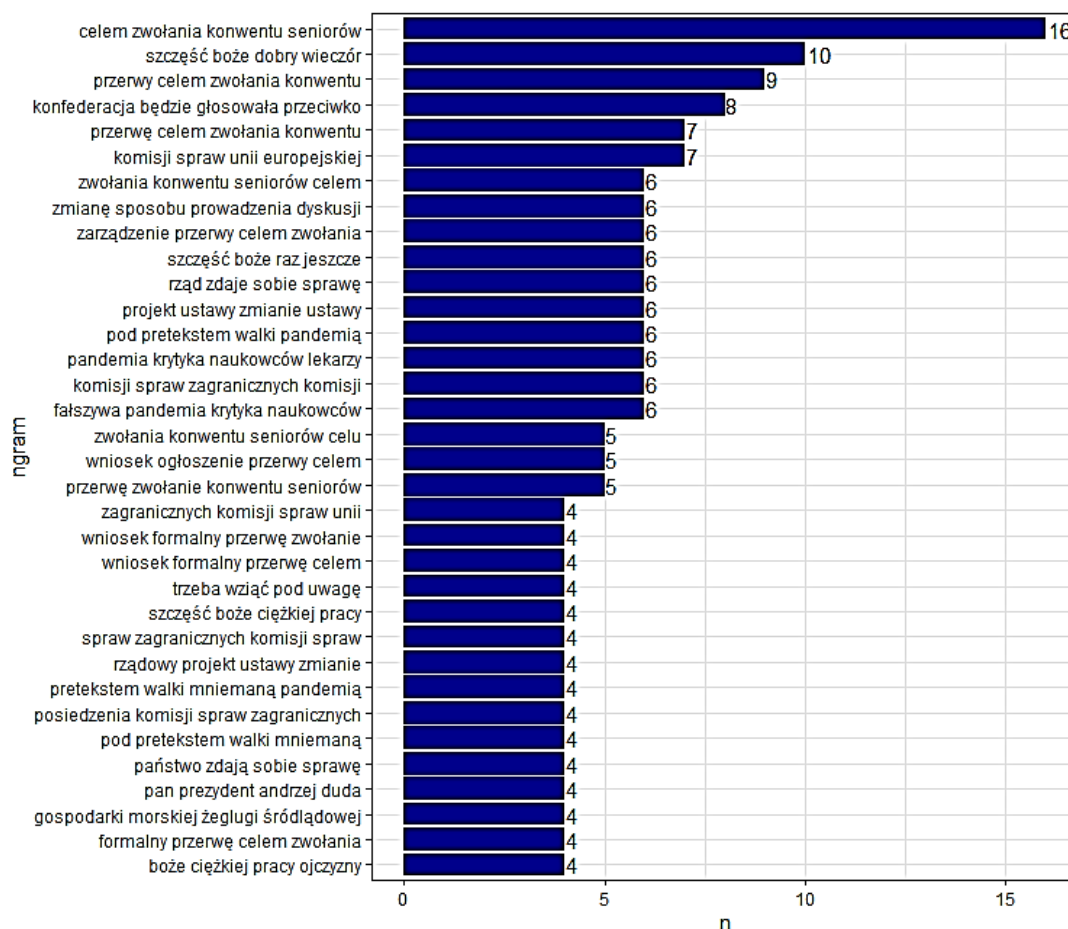
Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 116. Najczęściej występujące wyrazy 3-częściowe. Partia: Konfederacja



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 117. Najczęściej występujące wyrazy 4-częściowe. Partia: Konfederacja



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

3.7.6. Podsumowanie

Analiza ngramów pokazuje, że każda z grup ma swoje indywidualne zwroty n-częściowe i warto analizować każdą z osobna. Istnieją pewne charakterystyki, które są wspólne, jak wymienienie nazwy partii dla każdej z grup, lub używanie ogólnych zwrotów, jak *przede wszystkim*. Każda z partii w najczęściej używanych posiada zwrot dotyczący Unii Europejskiej. Natomiast istnieją zwroty, które charakteryzują tylko niektóre partie jak zwrot *polek [i] polaków* charakteryzujący KO i Lewicę. *Szczęść boże* czy *de facto* używany często przez posłów Konfederacji.

3.8. Analiza sentymentów

Głównym celem tej części był pomiar emocji biorąc pod uwagę użyte słowa używane przez posłów na mównicy sejmowej. Aby dokonać tego pomiaru posłużono się biblioteką

słów dostępną z serwisu *SłowoSieć – Politechniki Wrocławskiej* (plwordnet.pwr.wroc.pl)⁴⁶⁸. Dodatkowo, aby wzmocnić pomiar i uzyskać nacechowanie podstawowych słów, jak również ich odmian posłużono się biblioteką *PoliMorfologik*⁴⁶⁹. Złączenie baz z nacechowaniem (*SłowoSieć*) oraz odmianami słów (*PoliMorfologik*). Złączenie baz słów umożliwiło pomiar nie tylko na podstawowe brzmienie, ale również odmiany np. pies - psem, psu .itd. Wybrano z bazy słowa które mają nacechowanie negatywne (-m, -s) oraz nacechowanie pozytywne (+m, +s) oraz wszystkie te, które prezentują określone emocje. Większość słów ma przypisanych kilka emocji lub rodzajów nacechowania w zależności od kontekstu, dlatego niektóre słowa mogą mieć kilka emocji, więc te mogą być uwzględnione w kilku statystykach.

3.8.1. Wyniki ogólne

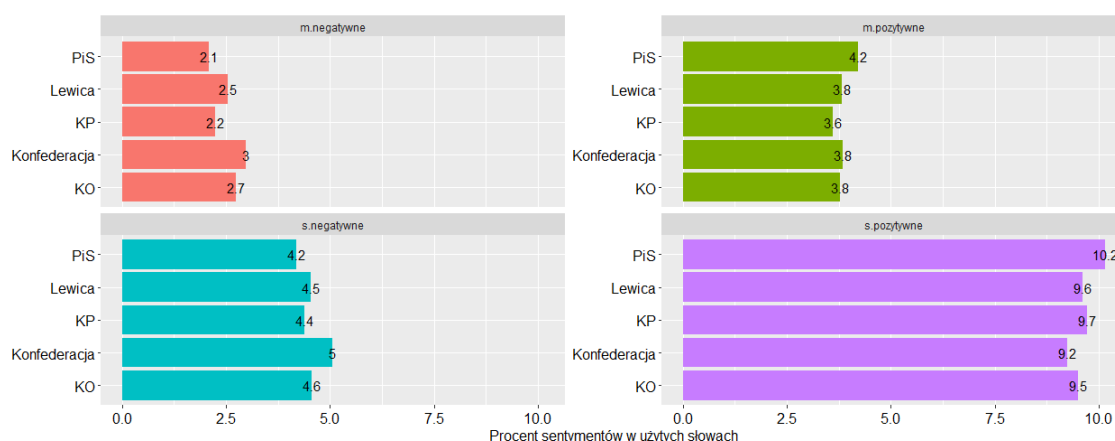
- Słów o nacechowaniu mocno negatywnym najczęściej notuje Konfederacja (3%), najmniej PiS (2,1%);
- Słów o nacechowaniu słabo negatywnym najczęściej notuje Konfederacja (5%), najmniej PiS (4,2%);
- Słów o nacechowaniu mocno pozytywnym najczęściej notuje PiS (4,2%), najmniej KP (3,6%);
- Słów o nacechowaniu słabo pozytywnym najczęściej notuje PiS (10,2%), najmniej Konfederacja (9,2%).

Dokładne wyniki przedstawia Rysunek 118.

⁴⁶⁸ A. Dziob, M. Piasecki, E. Rudnicka, *plWordNet 4.1—a Linguistically Motivated, Corpus-based Bilingual Resource*. In: Fellbaum, C., Vossen, P., Rudnicka, E., Maziarz, M., & Piasecki, M. (Eds.), *Proceedings of the 10th Global WordNet Conference: July 23-27, 2019, Wrocław (Poland)*. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, pp. 353--362. (pdf,Bibtex); M. Piasecki, S. Szpakowicz, B. Broda. *A Wordnet from the Ground Up*. Wrocław, Oficyna Wydawnicza Politechniki Wrocławskiej, 2009. (pdf,Bibtex)

⁴⁶⁹ M. Miłkowski, Biblioteka odmian słów, <https://github.com/morfologik/polimorfologik/blob/master/src/LICENSE.txt>, 28.12.2021; Copyright (c) \$copyright_date, Marcin Miłkowski All rights reserved.

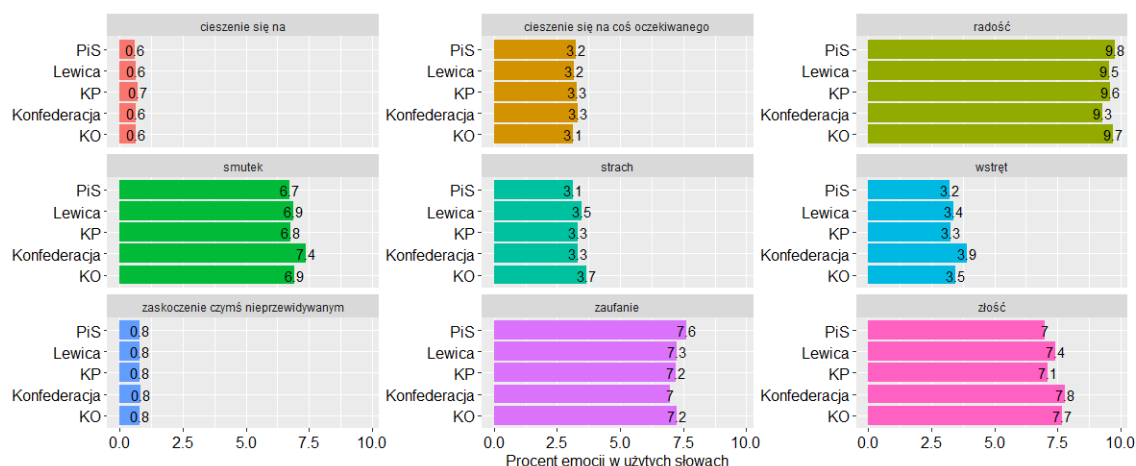
Rysunek 118. Procent sentymentów w użytych słowach z podziałem na partie



Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Najbardziej liczne emocje to radość, smutek, zaufanie oraz złość. Jeśli chodzi o radość najwyższe wartości notuje PiS (9,8%). W przypadku emocji smutku najwyższe wartości procentowe notuje Konfederacja (7,4%). Jeśli chodzi o zaufanie najwyższą wartość uzyskuje PiS (7,6%). Najwyższą wartość jeśli chodzi o złość notuje Konfederacja (7,8%).

Rysunek 119. Procent emocji w użytych słowach z podziałem na partie



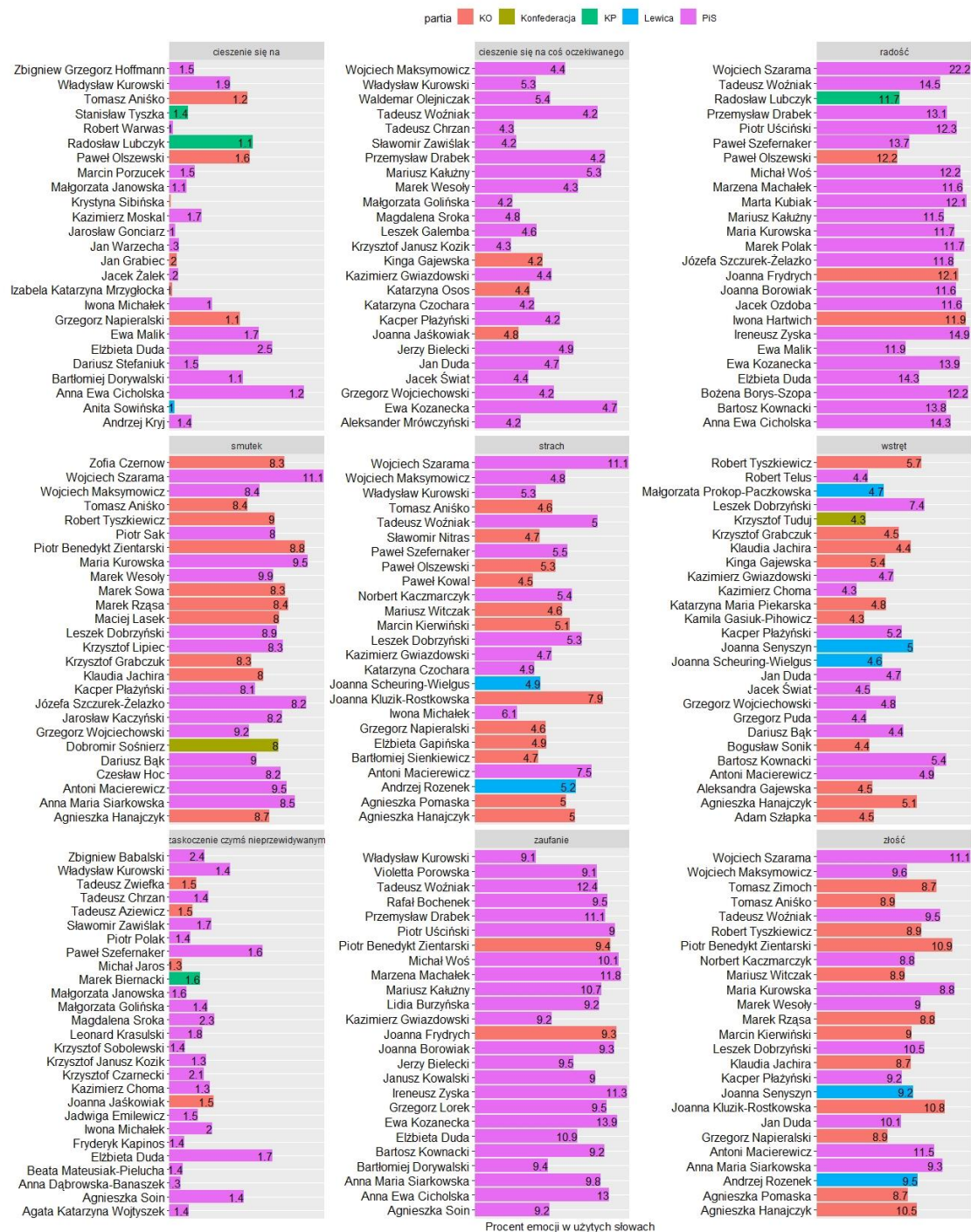
Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Wykres *Procent emocji w użytych słowach z podziałem na posłów* prezentuje 25 najwyższych wartości wśród posłów wszystkich formacji. Rozpatrując emocje pozytywne (radość, zaufanie, cieszenie się na coś oczekiwanego) widać większą licznosc posłów PiS

w stosunku do innych partii. W przypadku emocji negatywnych (smutek, złość, strach, wstręt) proporcje ulegają zmianie i widzimy większą w stosunku do emocji pozytywnych liczbę posłów z partii opozycyjnych.

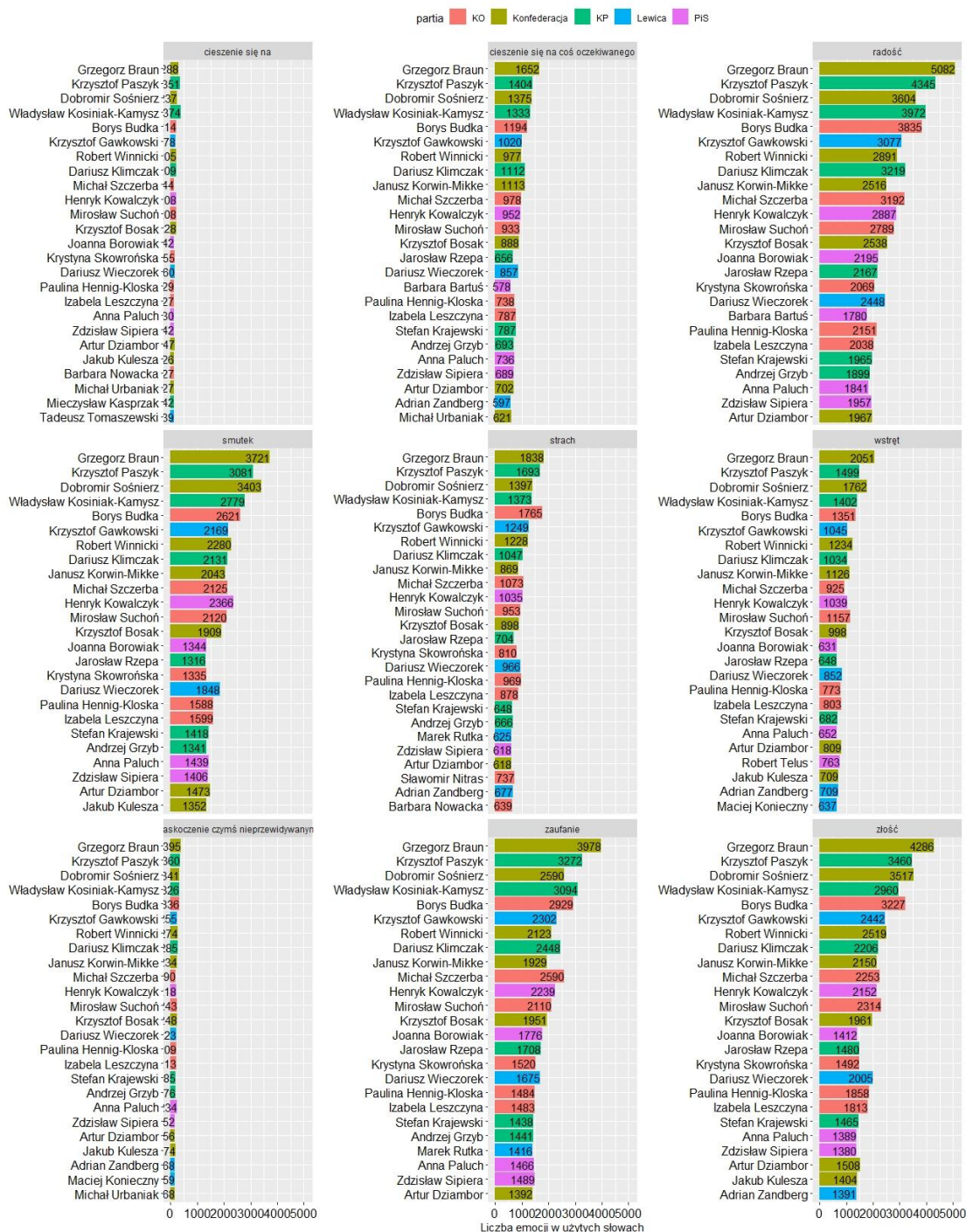
Liczba emocji w użytych słowach z podziałem na posłów pokazuje największe liczbowe wartości w użytych słowach pod względem emocji w wypowiedziach. Wartości są zależne od ogólnej liczby wypowiedzi posłów. Z tego powodu zwykle osoby mające najwyższą pozycję na liście mają jedną z największych wartości jeśli chodzi o liczbę wypowiedzi ogółem.

Rysunek 120. Procent emocji w użytych słowach z podziałem na posłów



Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Rysunek 121. Liczba emocji w użytych słowach z podziałem na posłów

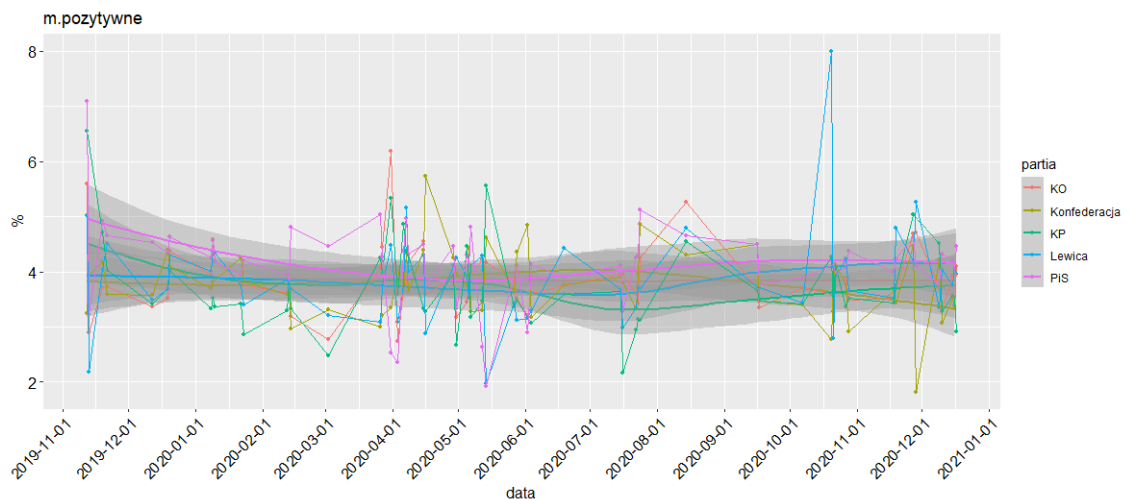


Źródło: opracowanie własne z użyciem biblioteki ggplot2.

3.8.2. Wyniki procentowe nacechowania w jednostce czasu

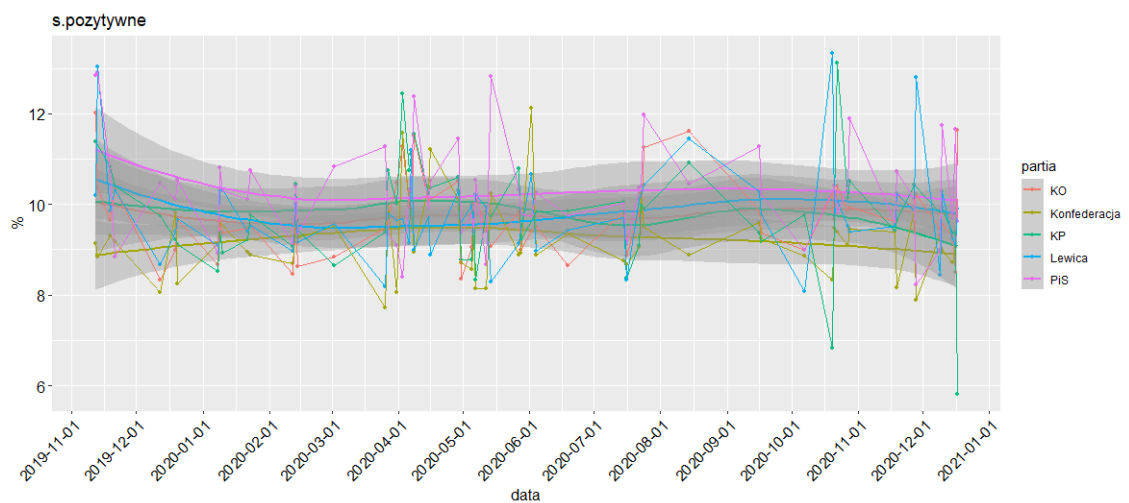
W przypadku słów o nacechowaniu mocno i słabo pozytywnym w procentowym ujęciu partia PiS przoduje w badanym okresie. W przypadku słów mocno i słabo negatywnym najwyższe wartości w badanym okresie prezentują partie: Konfederacja i Lewica.

Rysunek 122. Procent emocji mocno pozytywnych w użytych słowach z podziałem na partie w czasie



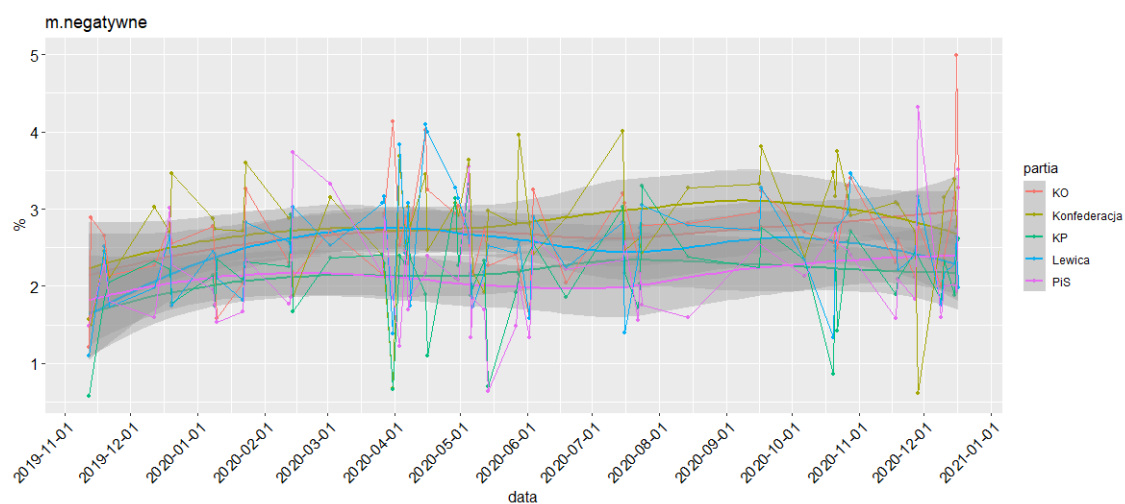
Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Rysunek 123. Procent emocji słabo pozytywnych w użytych słowach z podziałem na partie w czasie



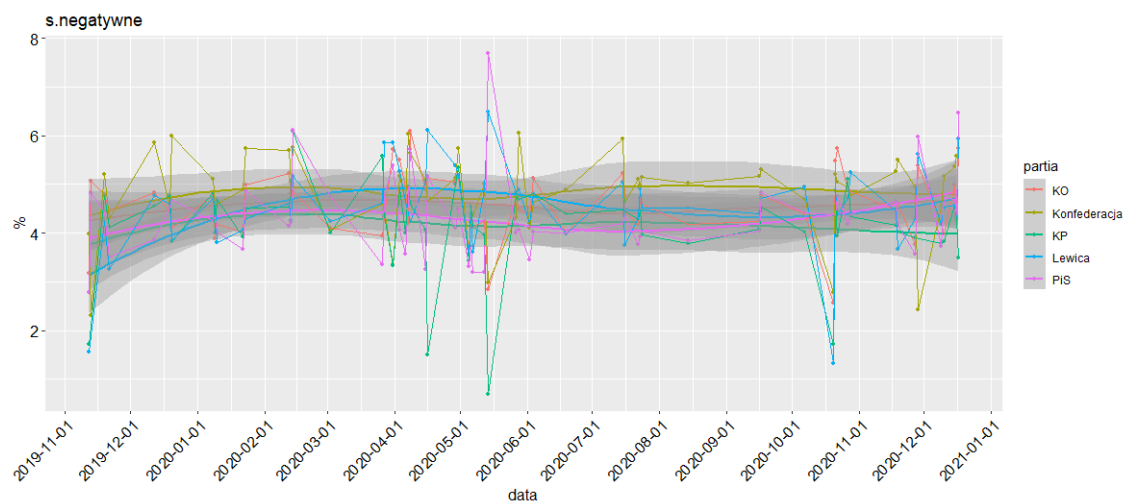
Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Rysunek 124. Procent emocji mocno negatywnych w użytych słowach z podziałem na partie w czasie



Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Rysunek 125. Procent emocji słabo negatywnych w użytych słowach z podziałem na partie w czasie



Źródło: opracowanie własne z użyciem biblioteki ggplot2.

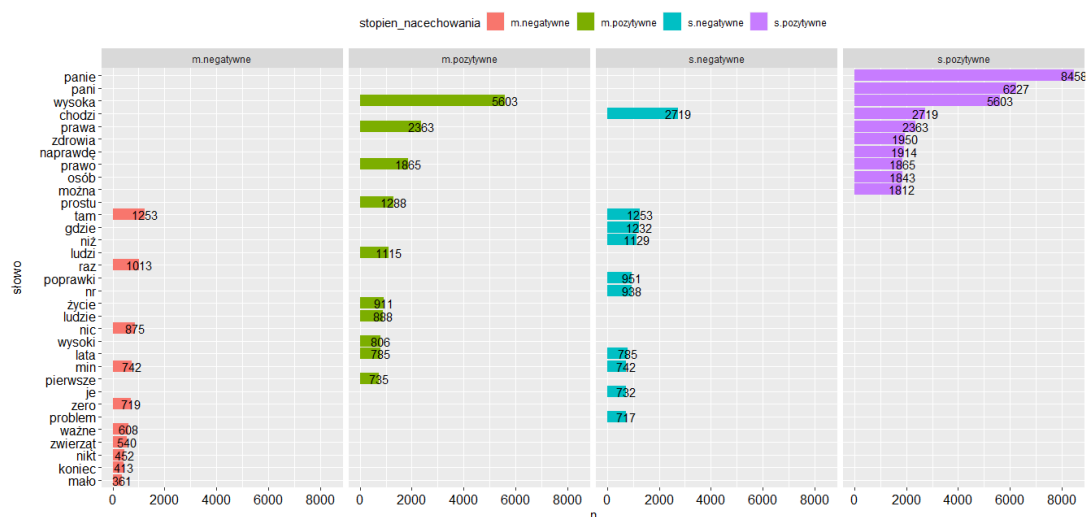
3.8.3. Najczęściej występujące słowa w wypowiedziach mające nacechowanie

Do najczęściej używanych słów mających nacechowanie mocno pozytywne należą (*wysoka, prawa, prawo*), w przypadku słów słabo pozytywnych są to słowa (*panie, pani, wysoka*).

Do najczęściej używanych słów mających nacechowanie mocno negatywne należą (*tam, raz, nic*), w przypadku słów słabo pozytywnych są to słowa (*chodzi, tam, gdzie*).

Należy pamiętać, że słowa wypowiadane w różnych kontekstach mogą mieć różne znaczenia, dlatego niektóre ze słów mogą pojawić się w kilku statystykach.

Rysunek 126. Najczęściej wypowiadane słowa mające nacechowanie



Źródło: opracowanie własne z użyciem biblioteki ggplot2.

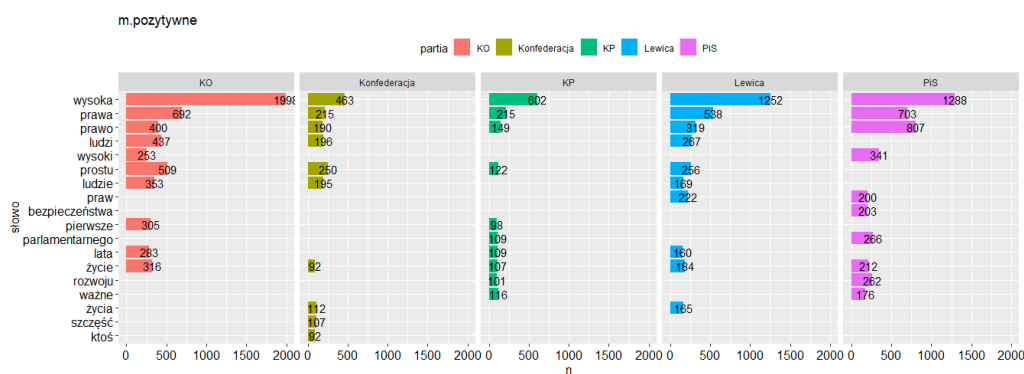
W przypadku słów mocno pozytywnych o największej liczności należą: *wysoka* (najwyższa wartość dla PiS), *prawa* i *prawo* (najwyższe wartości dla PiS).

W przypadku słów słabo pozytywnych o największej liczności należą: *panie* (najwyższa wartość dla KO), *pani* (najwyższa wartość dla KO), *wysoka* (najwyższa wartość dla KO).

W przypadku słów mocno negatywnych o największej liczności należą: *zero* (najwyższa wartość dla KO), *tam* (najwyższa wartość dla KO), *raz* (najwyższa wartość dla KO).

W przypadku słów słabo negatywnych o największej liczności należą: *nr* (najwyższa wartość dla PiS), *chodzi* (najwyższa wartość dla KO), *poprawek* (najwyższa wartość dla PiS).

Rysunek 127. Najczęściej wypowiedzane słowa mające nacechowanie mocno pozytywne.



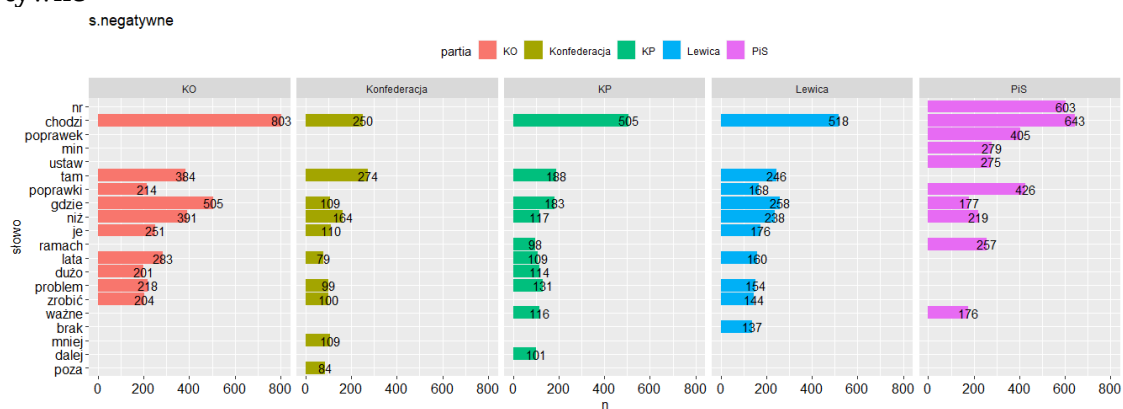
Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Rysunek 128. Najczęściej wypowiedzane słowa mające nacechowanie słabo pozytywne



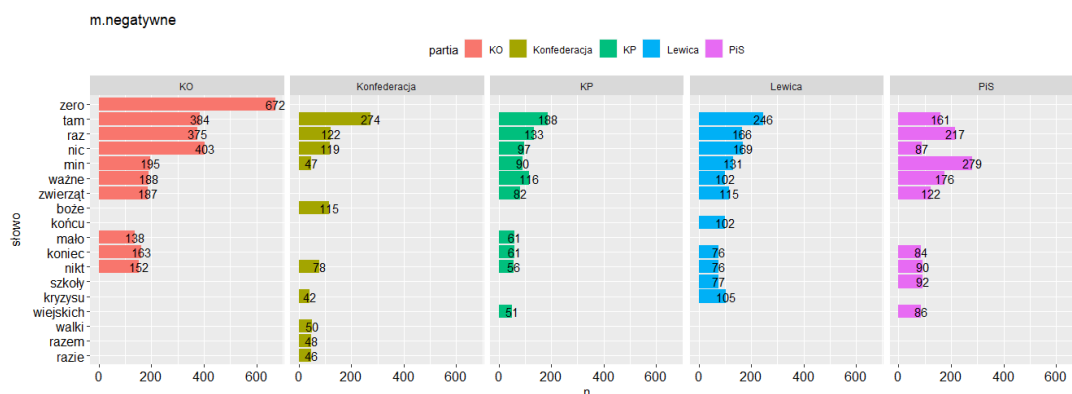
Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Rysunek 129. Najczęściej wypowiedzane słowa mające nacechowanie słabo negatywne



Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Rysunek 130. Najczęściej wypowiedzane słowa mające nacechowanie mocno negatywne



Źródło: opracowanie własne z użyciem biblioteki ggplot2.

3.8.4. Procentowy udział rodzaju nacechowania w wypowiedziach poszczególnych posłów

PiS

Posłowie z partii PiS, którzy mają najwyższe wartości procentowe negatywnych (słabo negatywnych + mocno negatywnych) sentymentów w swoich wypowiedziach to: Leszek Dobrzyński [16,2%], Dariusz Bąk [12,5%], Jan Duda [11,2%].

Posłowie z partii PiS, którzy mają najwyższe wartości procentowe pozytywnych (słabo pozytywnych + mocno pozytywnych) sentymentów w swoich wypowiedziach to: Tadeusz Woźniak [27,1%], Anna Ewa Cicholska [25,1%], Ewa Kozanecka [25%].

Tabela 25. Posłowie z udziałem sentymentów w wypowiedziach. Partia: PiS

Posłowie	m.neg	s.neg	neu- tralny	s.poz	m.poz
Adam Śnieżek	1,4	4,5	80,2	8,9	5,0
Agata Katarzyna Wojtysek	1,2	4,9	80,3	9,9	3,7
Agnieszka Górka	2,5	3,9	80,6	9,2	3,9
Agnieszka Soin	2,0	2,2	78,5	12,6	4,7
Aleksander Mrówczyński	2,1	4,3	77,5	11,2	4,9
Andrzej Gawron	2,4	4,9	78,1	11,2	3,4
Andrzej Kosztowniak	1,5	3,8	79,8	10,5	4,4
Andrzej Kryj	1,5	2,8	85,5	6,8	3,4
Andrzej Szlachta	1,9	4,6	79,1	10,1	4,2
Anita Czerwińska	1,6	3,2	79,2	11,3	4,7
Anna Dąbrowska-Banaszek	2,5	4,2	83,9	6,4	3,0

Anna Ewa Cicholska	1,7	4,7	68,5	18,0	7,1
Anna Gembicka	1,4	3,9	79,8	10,9	4,0
Anna Krupka	1,8	2,7	83,6	6,4	5,5
Anna Kwiecień	1,8	3,7	80,3	10,1	4,1
Anna Maria Siarkowska	3,9	4,8	74,4	12,1	4,8
Anna Milczanowska	1,8	3,3	79,2	10,1	5,6
Anna Paluch	1,5	4,9	81,9	8,5	3,2
Antoni Macierewicz	5,6	5,2	78,0	8,6	2,6
Arkadiusz Czartoryski	2,0	3,4	79,7	10,2	4,8
Arkadiusz Mularczyk	2,0	3,0	80,5	9,9	4,6
Artur Szałabawka	1,7	3,8	82,5	8,8	3,3
Barbara Bartuś	2,0	4,1	80,5	9,9	3,6
Barbara Dziuk	2,2	4,0	78,9	10,9	4,0
Bartłomiej Dorywalski	2,0	2,6	76,8	11,9	6,6
Bartosz Kownacki	0,6	2,4	74,4	14,9	7,7
Beata Mateusiak-Pielucha	0,4	1,3	85,9	9,4	3,0
Beata Strzałka	3,8	5,8	76,2	10,1	4,2
Bolesław Piecha	2,1	4,5	80,1	9,5	3,8
Bożena Borys-Szopa	1,4	3,3	81,1	10,3	4,0
Czesław Hoc	3,6	5,2	73,5	12,4	5,2
Daniel Milewski	1,5	3,1	80,4	10,7	4,3
Danuta Nowicka	2,6	4,5	80,0	9,7	3,2
Dariusz Bąk	5,2	7,3	72,7	9,3	5,4
Dariusz Olszewski	3,0	4,6	79,2	8,6	4,7
Dariusz Stefaniuk	0,3	2,7	86,8	8,4	1,9
Dominik Tarczyński	0,0	3,9	84,2	6,6	5,3
Dominika Chorosińska	1,7	3,8	79,2	11,3	4,0
Edward Siarka	2,7	3,6	79,7	9,9	4,0
Elżbieta Duda	0,0	2,4	78,6	15,5	3,6
Elżbieta Płonka	2,6	4,3	78,0	11,0	4,2
Ewa Kozanecka	0,0	1,9	73,1	16,7	8,3
Ewa Malik	2,4	5,9	74,8	13,9	3,0
Ewa Szymańska	2,0	5,2	81,2	8,4	3,2
Fryderyk Kapinos	2,7	3,1	83,0	7,8	3,3
Gabriela Masłowska	1,8	4,0	80,0	10,3	3,9
Grzegorz Lorek	2,8	4,8	75,3	12,7	4,4
Grzegorz Matusiak	1,5	3,2	81,5	10,6	3,2
Grzegorz Piechowiak	0,8	1,6	84,6	9,8	3,3
Grzegorz Puda	3,7	5,5	77,8	10,3	2,8
Grzegorz Wojciechowski	2,0	6,7	78,2	9,0	4,2
Henryk Kowalczyk	2,0	5,0	80,2	9,0	3,8
Ireneusz Zyska	0,8	3,1	73,3	14,5	8,4
Iwona Arent	1,2	3,7	80,7	9,9	4,6
Iwona Michałek	6,1	4,6	74,8	11,5	3,1
Jacek Kurzepa	1,3	3,4	81,1	9,4	4,8
Jacek Ozdoba	2,0	3,8	79,6	11,7	3,0

Jacek Świat	3,8	3,8	76,6	10,2	5,7
Jacek Żalek	3,1	2,1	79,3	11,7	3,8
Jadwiga Emilewicz	0,9	2,6	82,1	9,8	4,5
Jan Duda	6,1	5,1	75,5	11,2	2,0
Jan Kanthak	1,7	4,4	78,7	10,6	4,6
Jan Mosiński	1,7	3,8	78,7	10,1	5,7
Jan Warzecha	1,7	4,3	77,7	12,2	4,2
Janusz Kowalski	2,3	2,9	78,1	11,4	5,2
Janusz Śniadek	1,9	2,2	84,0	7,8	4,0
Jarosław Gonciarz	2,5	5,0	75,8	12,3	4,4
Jarosław Kaczyński	2,8	4,1	77,1	10,4	5,6
Jarosław Krajewski	1,6	3,2	80,9	10,0	4,3
Jarosław Zieliński	1,7	2,9	82,5	9,5	3,5
Jerzy Bielecki	1,1	3,1	75,1	14,1	6,6
Jerzy Małecki	1,5	3,8	78,6	11,8	4,3
Jerzy Materna	1,3	4,7	79,1	10,1	4,8
Jerzy Paul	2,7	3,9	79,0	10,1	4,4
Jerzy Polaczek	1,6	4,0	80,8	9,9	3,7
Joanna Borowiak	2,4	4,2	76,0	12,4	5,1
Joanna Lichocka	2,7	3,8	80,1	9,2	4,1
Józefa Szczurek-Żelazko	2,3	5,3	75,9	12,2	4,2
Kacper Płażyński	3,6	5,7	74,4	11,6	4,8
Kamil Bortniczuk	1,5	4,2	80,3	10,5	3,5
Katarzyna Czochara	0,8	5,8	79,7	11,6	2,1
Katarzyna Sójka	0,7	2,1	83,9	10,1	3,1
Kazimierz Choma	2,4	5,0	81,6	6,2	4,7
Kazimierz Gołojuch	0,8	1,9	83,3	10,7	3,4
Kazimierz Gwiazdowski	3,7	3,7	76,1	12,1	4,5
Kazimierz Moskal	1,9	1,9	81,5	11,3	3,4
Kazimierz Smoliński	1,5	3,7	82,7	8,7	3,5
Krzysztof Czarnecki	1,0	3,3	81,8	11,0	2,9
Krzysztof Janusz Kozik	0,5	2,5	80,7	11,4	5,0
Krzysztof Lipiec	2,5	5,6	76,2	10,9	4,9
Krzysztof Sobolewski	2,2	3,6	79,9	10,1	4,2
Krzysztof Szulowski	2,8	3,9	80,7	9,1	3,6
Krzysztof Tchórzewski	1,3	4,7	79,7	10,1	4,1
Leonard Krasulski	1,9	1,9	81,8	11,8	2,6
Leszek Dobrzyński	8,1	8,1	68,5	8,4	6,9
Leszek Galebba	0,8	3,4	82,9	10,2	2,7
Lidia Burzyńska	1,9	3,1	79,3	10,7	5,1
Łukasz Schreiber	2,6	4,4	78,9	9,7	4,4
Maciej Górski	2,3	6,7	77,5	9,5	4,0
Magdalena Sroka	0,7	4,6	81,6	9,4	3,7
Małgorzata Golińska	1,9	4,2	81,4	8,8	3,7
Małgorzata Janowska	1,2	2,4	85,1	8,2	3,2
Marcin Horała	2,6	4,7	78,5	9,6	4,6

Marcin Porzucek	2,0	4,2	79,7	9,9	4,1
Marek Ast	1,1	3,3	80,2	10,6	4,8
Marek Matuszewski	1,5	4,1	79,4	11,4	3,7
Marek Polak	2,8	4,7	76,4	12,1	4,0
Marek Suski	2,7	4,2	77,9	10,8	4,4
Marek Wesoły	1,3	6,6	78,3	11,8	2,0
Maria Kurowska	1,5	6,7	78,4	10,3	3,1
Mariusz Kałużny	1,9	4,1	75,2	11,9	6,9
Mariusz Trepka	1,4	3,6	78,4	12,8	3,8
Marta Kubiak	0,0	3,1	81,3	10,9	4,7
Marzena Machałek	0,7	2,0	76,8	13,3	7,2
Michał Jach	0,3	2,5	86,7	7,9	2,5
Michał Woś	3,6	1,8	74,7	14,2	5,7
Michał Wypij	0,8	2,3	83,8	7,3	5,8
Mirosława Stachowiak-Ró- żeczka	2,4	4,4	80,1	9,5	3,6
Norbert Kaczmarczyk	2,2	5,9	79,0	9,9	2,9
Patryk Wicher	2,0	3,6	76,9	12,2	5,3
Paweł Hreniak	1,6	4,2	78,7	11,1	4,3
Paweł Lisiecki	1,8	6,3	81,0	7,6	3,3
Paweł Rychlik	2,7	3,7	79,1	10,3	4,2
Paweł Szefernaker	1,5	3,6	73,0	14,6	7,3
Piotr Babinetz	0,8	4,3	82,3	9,0	3,5
Piotr Kaleta	2,1	3,1	79,7	10,3	4,8
Piotr Król	2,1	4,4	79,1	10,2	4,3
Piotr Polak	1,1	3,7	81,9	9,8	3,6
Piotr Sak	3,1	5,5	79,3	9,1	3,0
Piotr Uruski	2,3	4,5	80,5	9,0	3,8
Piotr Uściński	0,5	3,1	78,0	12,6	5,8
Przemysław Czarnecki	1,8	4,7	79,4	9,0	5,1
Przemysław Czarnek	2,1	3,8	81,4	9,4	3,4
Przemysław Drabek	1,9	2,3	75,5	14,2	6,1
Radosław Fogiel	2,1	3,7	80,2	10,0	3,9
Rafał Bochenek	2,3	4,6	76,7	12,5	3,8
Robert Gontarz	1,4	5,4	78,1	10,8	4,3
Robert Telus	3,4	4,5	78,7	9,0	4,5
Robert Warwas	1,3	3,5	78,0	12,8	4,4
Ryszard Bartosik	2,7	3,9	78,9	10,3	4,3
Ryszard Terlecki	2,2	3,7	79,7	10,4	4,0
Sławomir Skwarek	1,5	3,6	78,5	11,0	5,3
Sławomir Zawiślak	0,0	1,7	85,1	9,1	4,1
Sylwester Tułajew	2,1	3,5	79,9	9,9	4,6
Szymon Szynkowski vel Sęk	1,4	2,3	84,2	8,3	3,8
Tadeusz Chrzan	1,8	3,5	79,8	11,0	3,9
Tadeusz Cymański	2,4	4,9	78,1	10,0	4,6
Tadeusz Woźniak	4,4	6,2	62,3	16,8	10,3

Teresa Pamuła	1,5	4,4	81,0	9,4	3,7
Teresa Wargocka	2,3	4,6	77,4	10,6	5,0
Tomasz Latos	2,1	3,9	78,2	12,2	3,7
Tomasz Rzymkowski	0,4	4,6	80,4	10,2	4,4
Tomasz Zieliński	2,0	3,5	80,2	9,3	5,1
Urszula Rusecka	2,3	3,9	79,1	10,4	4,3
Violetta Porowska	2,7	4,4	76,4	11,9	4,6
Waldemar Andzel	3,3	4,5	78,2	9,6	4,5
Waldemar Olejniczak	0,6	2,4	85,6	9,0	2,4
Wiesław Janczyk	1,8	4,0	80,4	9,5	4,3
Wiesław Krajewski	1,3	4,6	78,1	12,0	3,9
Witold Czarnecki	1,7	4,1	78,7	11,6	3,9
Władysław Kurowski	0,7	3,4	78,2	14,3	3,4
Włodzimierz Tomaszewski	1,8	3,5	88,5	4,9	1,3
Wojciech Maksymowicz	4,0	6,3	75,3	9,8	4,6
Wojciech Szarama	0,0	0,0	85,7	14,3	0,0
Wojciech Zubowski	2,2	3,8	82,4	8,3	3,3
Zbigniew Babalski	0,0	4,4	82,9	10,1	2,5
Zbigniew Dolata	1,7	4,3	77,9	11,2	4,9
Zbigniew Girzyński	1,8	4,9	79,5	9,3	4,5
Zbigniew Rau	1,4	2,9	83,7	7,2	4,8
Zdzisław Sipiera	1,3	3,4	84,5	8,4	2,4

Źródło: opracowanie własne.

KO

Posłowie z partii KO, którzy mają najwyższe wartości procentowe negatywnych (słabo negatywnych + mocno negatywnych) sentymentów w swoich wypowiedziach to: Tomasz Aniśko [11%], Agnieszka Hanajczyk [10,4%], Marek Rząsa [10%].

Posłowie z partii KO, którzy mają najwyższe wartości procentowe pozytywnych (słabo pozytywnych + mocno pozytywnych) sentymentów w swoich wypowiedziach to: Piotr Benedykt Zientarski [18,7%], Grzegorz Napieralski [18,2%], Mateusz Bochenek [16,7%].

Tabela 26. Posłowie z udziałem sentymentów w wypowiedziach. Partia: KO

Posłowie	m.neg	s.neg	neutralny	s.poz	m.poz
Adam Szłapka	3,9	4,4	77,9	10,1	3,7
Agnieszka Hanajczyk	4,0	6,4	76,9	9,9	2,8
Agnieszka Pomaska	2,0	4,4	81,9	8,8	2,8
Aleksander Miszalski	2,2	3,8	81,6	8,8	3,5
Aleksandra Gajewska	4,8	4,0	78,4	9,5	3,3
Andrzej Szewiński	2,8	4,1	80,9	8,7	3,5

Anna Wasilewska	2,5	4,2	81,7	8,5	3,1
Arkadiusz Marchewka	2,6	4,4	79,1	9,5	4,3
Arkadiusz Myrcha	2,6	4,6	81,3	8,3	3,2
Barbara Dolniak	1,8	3,7	82,7	8,3	3,5
Barbara Nowacka	3,6	5,4	76,0	10,4	4,5
Bartłomiej Sienkiewicz	3,2	4,5	79,2	9,2	3,9
Bogusław Sonik	4,3	5,1	79,2	8,0	3,4
Borys Budka	3,1	4,6	78,6	9,5	4,2
Cezary Grabarczyk	2,2	4,9	80,2	9,1	3,6
Cezary Tomczyk	2,6	4,2	80,9	8,8	3,5
Czesław Mroczek	3,3	3,9	79,2	9,6	4,0
Dariusz Joński	2,6	4,9	79,7	8,8	4,0
Dariusz Rosati	2,5	5,0	80,0	8,7	3,9
Dorota Niedziela	2,5	3,6	82,7	8,2	3,1
Elżbieta Gapińska	2,2	2,9	82,7	7,4	4,9
Elżbieta Gelert	1,0	3,9	85,5	8,1	1,6
Eugeniusz Czykwin	1,9	2,5	82,2	8,8	4,6
Ewa Kołodziej	2,7	4,6	78,2	9,9	4,7
Franciszek Sterczewski	3,2	4,3	77,5	10,5	4,5
Gabriela Lenartowicz	2,2	4,1	81,3	8,7	3,7
Grzegorz Napieralski	2,6	4,4	74,8	13,3	4,9
Grzegorz Rusiecki	1,8	4,3	78,7	10,5	4,6
Grzegorz Schetyna	2,0	4,6	80,5	8,6	4,4
Henryka Krzywonos-Strycharska	2,9	5,1	76,8	11,6	3,6
Ireneusz Raś	1,5	3,6	81,9	9,3	3,7
Iwona Hartwich	2,6	4,7	77,2	11,1	4,5
Iwona Maria Kozłowska	2,2	4,2	77,3	11,8	4,5
Iwona Śledzińska-Katarasińska	1,7	3,5	82,6	9,0	3,1
Izabela Katarzyna Mrzygłocka	1,3	6,0	76,0	11,3	5,3
Izabela Leszczyna	2,8	4,9	80,5	8,1	3,7
Jacek Protas	2,5	3,9	79,7	9,3	4,7
Jagna Marczułajtis-Walczak	1,9	3,3	81,9	10,1	2,9
Jakub Rutnicki	2,5	3,8	78,8	10,8	4,2
Jan Grabiec	1,7	3,1	80,3	10,6	4,2
Janusz Cichoń	2,0	5,1	80,4	9,6	2,9
Jarosław Urbaniak	2,9	5,1	80,5	8,1	3,5
Jerzy Hardie-Douglas	2,6	5,5	80,1	8,9	2,9
Joanna Fabisiak	2,2	5,1	79,2	9,2	4,3
Joanna Frydrych	3,8	4,4	75,4	11,7	4,7
Joanna Jaśkowiak	1,6	5,7	76,1	11,5	5,1
Joanna Kluzik-Rostkowska	3,3	4,6	80,8	9,4	2,0
Joanna Mucha	1,9	2,7	85,4	7,0	3,0
Kamila Gasiuk-Pihowicz	3,9	5,3	76,0	9,8	5,0
Katarzyna Lubnauer	3,3	4,8	80,2	8,2	3,6
Katarzyna Maria Piekarska	3,9	4,8	77,8	10,4	3,2
Katarzyna Osos	1,8	4,1	78,8	11,0	4,3

Kazimierz Plocke	2,9	4,2	80,0	8,9	4,0
Killion Munyama	0,5	3,3	83,5	8,8	3,8
Kinga Gajewska	3,7	6,3	74,6	10,3	5,1
Klaudia Jachira	4,0	5,8	77,0	9,1	4,0
Konrad Frysztak	1,7	3,6	82,2	9,5	3,0
Krystyna Sibińska	2,3	5,8	77,9	10,5	3,5
Krystyna Skowrońska	1,8	4,5	81,5	9,2	3,0
Krystyna Szumilas	2,6	4,8	81,4	7,6	3,5
Krzysztof Gadowski	1,5	3,6	82,8	8,6	3,5
Krzysztof Grabczuk	3,0	5,4	77,4	9,7	4,5
Krzysztof Mieszkowski	3,2	3,4	77,9	9,2	6,3
Krzysztof Piątkowski	2,3	4,1	80,8	9,0	3,8
Krzysztof Truskolaski	2,0	4,6	79,7	9,6	4,1
Maciej Lasek	3,8	5,2	76,7	10,3	4,0
Małgorzata Chmiel	2,9	5,3	79,2	8,9	3,7
Małgorzata Kidawa-Błońska	2,6	4,6	78,4	10,2	4,2
Małgorzata Niemczyk	2,5	5,4	78,2	9,7	4,2
Małgorzata Pępek	2,6	4,6	79,3	9,7	3,8
Małgorzata Tracz	4,3	5,1	75,4	10,9	4,3
Marcin Kierwiński	2,8	4,1	80,1	10,3	2,7
Marek Hok	1,5	3,6	82,7	9,6	2,6
Marek Krząkała	2,1	3,0	80,9	9,3	4,7
Marek Rząsa	4,4	5,6	73,9	11,0	5,0
Marek Sowa	3,6	5,5	78,4	8,7	3,8
Maria Małgorzata Janyska	3,1	4,0	81,1	8,6	3,2
Mariusz Witczak	3,4	4,1	78,8	9,8	3,9
Marta Golbik	2,1	4,0	83,3	8,1	2,5
Marta Weisło	3,3	5,0	78,5	9,6	3,6
Marzena Okła-Drewnowicz	3,0	5,1	79,2	9,5	3,2
Mateusz Bochenek	2,6	4,2	76,5	11,1	5,5
Michał Gramatyka	2,1	4,0	79,7	10,4	3,8
Michał Jaros	2,2	4,6	80,6	9,3	3,2
Michał Krawczyk	2,5	4,6	81,8	7,8	3,3
Michał Szerba	2,4	4,1	78,4	11,1	4,0
Mirosław Suchoń	3,1	4,8	78,4	9,9	3,8
Mirosława Nykiel	3,2	4,6	78,7	9,3	4,2
Monika Rosa	3,5	5,0	77,1	10,8	3,6
Monika Wielichowska	3,4	4,7	77,5	10,0	4,4
Paulina Hennig-Kloska	2,7	5,0	79,5	9,9	2,9
Paweł Kowal	2,0	4,5	80,9	9,3	3,2
Paweł Olszewski	1,4	2,1	81,4	10,7	4,3
Paweł Papke	2,7	4,3	78,5	10,4	4,0
Paweł Poncyłjusz	1,4	3,2	83,0	9,0	3,4
Paweł Zalewski	2,9	3,9	77,6	11,0	4,6
Piotr Adamowicz	2,6	2,9	80,0	10,3	4,2
Piotr Benedykt Zientarski	2,4	5,6	73,4	13,5	5,2

Piotr Borys	3,1	4,8	77,6	10,5	4,0
Rafał Grupański	3,2	3,6	78,7	10,1	4,4
Rajmund Miller	2,4	3,8	82,9	8,0	3,0
Robert Kropiwnicki	2,0	3,8	79,9	9,9	4,4
Robert Tyszkiewicz	6,3	3,7	78,9	8,2	3,0
Ryszard Wilczyński	3,0	5,0	80,6	7,9	3,6
Sławomir Jan Piechota	1,8	4,8	80,3	9,8	3,3
Sławomir Neumann	2,0	5,5	80,4	9,8	2,3
Sławomir Nitras	2,5	3,8	81,3	9,2	3,3
Tadeusz Aziewicz	2,8	5,0	77,4	10,7	4,1
Tadeusz Zwiefka	0,7	3,0	85,2	8,2	3,0
Tomasz Aniśko	3,9	7,1	78,5	7,0	3,5
Tomasz Głogowski	2,7	4,1	78,6	10,5	4,1
Tomasz Kostuś	3,3	5,0	76,7	10,6	4,4
Tomasz Lenz	2,8	3,4	82,1	9,0	2,8
Tomasz Olichwer	2,1	3,9	80,4	10,1	3,5
Tomasz Piotr Nowak	1,8	4,4	80,9	9,0	3,9
Tomasz Szymański	2,6	5,5	77,5	10,4	3,9
Tomasz Zimoch	2,7	4,4	79,1	9,7	4,0
Urszula Augustyn	2,1	4,0	80,0	10,1	3,9
Urszula Zielińska	3,1	4,3	79,3	9,4	4,0
Waldemar Sługocki	3,0	4,9	79,4	8,8	3,9
Waldy Dzikowski	1,3	3,0	80,6	10,6	4,6
Witold Zembaczyński	2,0	2,7	81,7	9,2	4,5
Wojciech Król	3,4	5,0	76,2	9,9	5,6
Wojciech Saługa	2,4	4,5	80,4	9,5	3,2
Zbigniew Konwiński	1,7	5,4	81,1	8,9	2,8
Zofia Czernow	3,0	5,6	78,4	9,1	3,8

Źródło: opracowanie własne.

Lewica

Posłowie z partii Lewica, którzy mają najwyższe wartości procentowe negatywnych (słabo negatywnych + mocno negatywnych) sentymentów w swoich wypowiedziach to: Joanna Senyszyn [10,1%], Małgorzata Prokop-Paczkowska [9,5%], Anita Kucharska-Dziedzic [9,1%].

Posłowie z partii Lewica, którzy mają najwyższe wartości procentowe pozytywnych (słabo pozytywnych + mocno pozytywnych) sentymentów w swoich wypowiedziach to: Maciej Konieczny [16,3%], Bogusław Wontor [15,8%], Beata Maciejewska [15,7%].

Tabela 27. Posłowie z udziałem sentymentów w wypowiedziach. Partia: Lewica

Posłowie	m.neg	s.neg	neutralny	s.poz	m.poz
Adrian Zandberg	2,9	5,3	78,3	9,7	2,9
Agnieszka Dziemianowicz-Bąk	3,5	5,1	76,5	10,2	3,5
Andrzej Rozenek	3,3	4,5	78,4	10,3	3,3
Andrzej Szejna	2,5	4,2	78,7	10,2	2,5
Anita Kucharska-Dziedzic	3,6	5,6	77,0	10,0	3,6
Anita Sowińska	2,3	4,8	80,0	9,5	2,3
Anna Maria Żukowska	2,0	4,1	81,1	9,1	2,0
Beata Maciejewska	3,2	3,9	77,3	10,7	3,2
Bogusław Wontor	1,4	3,2	79,7	11,8	1,4
Daria Gosek-Popiołek	3,1	5,2	78,9	9,4	3,1
Dariusz Wiczorek	1,8	4,1	83,2	7,9	1,8
Jacek Czerniak	2,1	3,5	80,0	11,2	2,1
Jan Szopiński	1,8	3,7	80,8	9,8	1,8
Joanna Scheuring-Wielgus	3,4	4,7	78,8	8,9	3,4
Joanna Senyszyn	4,2	5,8	76,3	8,8	4,2
Karolina Pawliczak	3,0	5,4	76,9	9,8	3,0
Katarzyna Kotula	2,8	4,5	79,1	10,3	2,8
Katarzyna Kretkowska	2,3	4,6	81,7	8,4	2,3
Katarzyna Ueberhan	2,8	4,8	78,0	10,1	2,8
Krzysztof Gawkowski	3,0	4,8	78,0	9,8	3,0
Krzysztof Śmiszek	2,7	4,2	78,6	10,0	2,7
Maciej Gdula	1,6	3,6	79,9	10,5	1,6
Maciej Konieczny	3,1	5,3	75,3	11,2	3,1
Maciej Kopiec	2,3	4,3	78,1	9,7	2,3
Magdalena Biejat	2,9	5,1	78,4	9,9	2,9
Małgorzata Prokop-Paczkowska	3,7	5,8	78,0	8,4	3,7
Małgorzata Sekuła-Szmajdzińska	1,6	4,3	79,5	10,5	1,6
Marcelina Zawisza	2,5	4,7	79,0	10,2	2,5
Marcin Kulasek	2,9	4,1	79,9	9,5	2,9
Marek Dyduch	2,0	5,3	81,4	8,1	2,0
Marek Rutka	2,4	4,1	79,5	10,1	2,4
Monika Falej	2,7	4,5	79,3	9,8	2,7
Monika Pawłowska	2,7	4,9	78,9	9,2	2,7
Paulina Matysiak	3,0	4,8	78,2	10,1	3,0
Paweł Krutul	2,5	4,5	78,8	10,6	2,5
Przemysław Koperski	2,3	4,4	79,8	9,9	2,3
Rafał Adamczyk	1,7	4,1	79,9	10,3	1,7
Robert Kwiatkowski	1,7	3,6	81,7	9,2	1,7
Robert Obaz	2,6	4,6	80,5	9,3	2,6
Romuald Ajchler	1,7	4,2	82,3	8,7	1,7
Tadeusz Tomaszewski	2,0	3,9	79,4	11,0	2,0

Tomasz Trela	2,6	4,5	81,0	8,4	2,6
Wanda Nowicka	3,2	4,8	77,4	10,5	3,2
Wiesław Buż	1,5	4,1	81,2	9,6	1,5
Wiesław Szczepański	1,6	3,8	83,6	8,5	1,6
Włodzimierz Czarzasty	2,6	5,2	78,0	10,0	2,6
Zdzisław Wolski	2,6	4,4	81,4	8,5	2,6

Źródło: opracowanie własne.

KP

Posłowie z projektu politycznego KP, którzy mają najwyższe wartości procentowe negatywnych (słabo negatywnych + mocno negatywnych) sentymentów w swoich wypowiedziach to: Jacek Protasiewicz [7,7%], Stanisław Żuk [7,6%], Mirosław Maliszewski [7,5%].

Posłowie z KP, którzy mają najwyższe wartości procenta pozytywnych (słabo pozytywnych + mocno pozytywnych) sentymentów w swoich wypowiedziach to: Radosław Lubczyk [16,9%], Stanisław Tyszka [14,9%], Władysław Teofil Bartoszewski [14,9%].

Tabela 28. Posłowie z udziałem sentymentów w wypowiedziach. Partia: KP

Posłowie	m.neg	s.neg	neutralny	s.poz	m.poz
Agnieszka Ścigaj	2,3	4,7	80,4	9,4	3,2
Andrzej Grzyb	1,8	4,1	82,3	8,7	3,2
Bożena Żelazowska	2,5	3,6	80,2	9,4	4,3
Czesław Siekierski	1,9	5,2	78,5	10,5	3,7
Dariusz Klimczak	1,8	4,1	80,5	10,1	3,4
Dariusz Kurzawa	1,9	4,6	80,2	9,3	4,0
Jacek Protasiewicz	2,9	4,8	79,5	9,1	3,7
Jacek Tomczak	1,8	4,5	82,0	8,7	3,0
Jarosław Rzepa	2,3	3,7	79,3	11,0	3,7
Jarosław Sachajko	2,1	5,2	78,9	10,3	3,6
Jolanta Fedak	2,5	4,5	80,6	8,9	3,5
Krzysztof Paszyk	2,2	4,6	79,4	10,2	3,6
Marek Biernacki	2,9	4,3	80,8	9,7	2,4
Marek Sawicki	2,6	4,4	82,6	7,3	3,1
Mieczysław Kasprzak	2,7	4,8	78,8	10,2	3,5
Mirosław Maliszewski	2,1	5,4	83,3	6,9	2,3
Paweł Szramka	2,2	4,9	78,2	10,8	4,0
Piotr Zgorzelski	2,2	3,9	81,0	9,2	3,7
Radosław Lubczyk	1,8	4,2	77,1	12,7	4,2
Stanisław Tyszka	2,0	4,3	78,8	10,4	4,5

Stanisław Żuk	2,8	4,8	78,4	10,7	3,3
Stefan Krajewski	2,2	4,6	80,2	9,6	3,5
Urszula Nowogórska	1,7	3,4	82,2	9,4	3,3
Urszula Paślawska	2,1	4,1	79,7	10,5	3,5
Władysław Kosiniak-Kamysz	2,6	4,3	79,0	9,9	4,3
Władysław Teofil Bartoszewski	2,4	3,9	78,8	9,6	5,2
Zbigniew Ziejewski	2,6	3,9	81,8	8,4	3,2

Źródło: opracowanie własne.

Konfederacja

Posłowie z koła Konfederacja, którzy mają najwyższe wartości procentowe negatywnych (słabo negatywnych + mocno negatywnych) sentymentów w swoich wypowiedziach to: Dobromir Sośnierz [9,4%], Janusz Korwin-Mikke [8,6%], Grzegorz Braun [8,5%].

Posłowie z koła Konfederacja, którzy mają najwyższe wartości procentowe pozytywnych (słabo pozytywnych + mocno pozytywnych) sentymentów w swoich wypowiedziach to: Krzysztof Tuduj [15,5%], Grzegorz Braun [14,5%], Michał Urbaniak [13,7%].

Tabela 29. Posłowie z udziałem sentymentów w wypowiedziach. Partia: Konfederacja

Posłowie	m.neg	s.neg	neutralny	s.poz	m.poz
Artur Dziambor	2,6	4,8	79,9	9,1	3,7
Dobromir Sośnierz	3,8	5,6	79,3	8,1	3,3
Grzegorz Braun	3,5	5,0	77,1	10,0	4,4
Jakub Kulesza	2,7	4,6	81,1	8,4	3,2
Janusz Korwin-Mikke	3,1	5,5	77,9	9,3	4,2
Konrad Berkowicz	2,7	5,2	78,5	9,5	4,0
Krystian Kamiński	2,1	4,9	80,1	9,5	3,4
Krzysztof Bosak	2,5	4,8	79,5	9,6	3,6
Krzysztof Tuduj	2,9	5,5	76,1	10,5	5,0
Michał Urbaniak	2,0	4,2	80,1	9,6	4,1
Robert Winnicki	2,8	5,0	79,3	9,1	3,8

Źródło: opracowanie własne.

3.8.5. Podsumowanie

Słów o nacechowaniu mocno i słabo negatywnym najczęściej notuje Konfederacja (8%), najmniej PiS (6,3%). Słów o nacechowaniu mocno pozytywnym najczęściej notuje PiS (4,2%), najmniej KP (3,6%). Słów o nacechowaniu słabo pozytywnym najczęściej notuje PiS (10,2%), najmniej Konfederacja (9,2%). Najbardziej liczne emocje w wypowiedzianych słowach przez polityków to radość, smutek, zaufanie oraz złość. Jeśli chodzi o emocję radości to najwyższe wartości pod względem liczności wystąpień notuje PiS (9,8%). W przypadku emocji smutku najwyższe wartości procentowe notuje Konfederacja (7,4%). Jeśli chodzi o zaufanie najwyższą wartość uzyskuje PiS (7,6%), zaś najwyższą wartość jeśli chodzi o złość notuje Konfederacja (7,8%).

Oczywiście dokładne nacechowanie zależy od kontekstu wypowiedzi np. słowo *nikt* lub słowo *zero* (klasyfikowane jako mocno negatywne) zależą od wypowiadanego kontekstu. Z pewnością będą zdarzać się przypadki, że słowo klasyfikowane jako negatywne lub pozytywne w użytym kontekście w całym zdaniu nie będzie miało takiego ładunku. Jest to pewne przybliżenie, co oznacza, że sytuacja ze złym sklasyfikowaniem słów (i wypowiedzi) może zdarzyć się jednakowo często wśród posłów z każdej z partii, stąd też przyjęte zostało założenie polegające na pomiarze częstości słów. Jest to obszar, który wymaga rozwinięcia, zbudowania odpowiedniej bazy i zastosowania kolejnych metod analitycznych ze wsparciem językoznawców do usprawnień w budowaniu klasyfikatorów z ujęciem kontekstów wypowiedzi.

Podsumowując, posłowie należący do partii rządzącej używają języka podkreślając pozytywne nacechowanie (również dotyczy to pozytywnych emocji jak radość, zaufanie), zaś w przypadku nacechowania negatywnego (w tym emocji negatywnych jak złość, strach, wstręt) zwykle wyższe wartości udziału użytych słów w swoich wypowiedziach prezentuje opozycja.

3.9. Dane z Twittera

W kolejnej części rozdziału 3. skupiono się na analizie danych pochodzących z serwisu Twitter. Analiza miała na celu sprawdzenie czy wypowiedzi osób wspominających o danej partii będą tworzyły podobne słowniki oraz czy będą niosły podobne nacechowanie czy emocje. Niektóre metody jak m.in. analiza liczności słów, ngramy, czy sentymenty powtarzają się z poprzednią częścią. Pomiar relacji z użyciem analizy koszykowej zastąpiono analizą korelacji. Ze względu na wypowiedzi internetowe niektóre słowa

mogą zawierać niecenzuralny charakter, dlatego w niektórych momentach celowo zastosowano znak gwiazdki *.

Dane zostały pobrane do analizy z narzędzia *bsi-sentiment*⁴⁷⁰ napisanego w języku *Python* i obejmują ten sam okres, jak w przypadku danych posłów z mównicy sejmowej (okres: 01.11.2019 do 31.12.2020).

Dane do analizy zostały odpowiednio przygotowane w języku *R*. Wykonane zostały standardowe przekształcenia występujące w analizie tekstu jak: ujednolicenie do małych liter, usuwanie łączników, nawiasów i innych symboli⁴⁷¹.

Następnie usunięto z wypowiedzi spójniki oraz słowa mające marginalne znaczenie takie jak: *i, to, na, z, jest, że, ze, się, sie, o, do, a, tym, czy, jak, tego, które, ktore, sa, są, ale, co, dla, za, po, tej, przez, po, bo, którzy, która, którzy, która*. Utrzymano mniejszy zakres słów do usunięcia niż w przypadku danych z Sejmu RP, a skupiono się na usunięciu innych symboli (dane były znacznie bardziej „brudne”). Ponadto dane zawierają skróty, slangi i wyrazy pisane czasami niechlujnie.

Aby dokonać tych przekształceń posłużono głównie bibliotekami jak: *tidytext, tm, dplyr*.

Tabela 30. Liczba wpisów zawierających nazwę partii w badanym okresie, a wynik wyborczy w 2019.

Partia	Liczba wpisów zawierających nazwę partii w badanym okresie	Wynik wyborczy 2019 [%]	Liczba mandatów
PiS	789 272	43,59	235
KO	4 727	27,4	134
Lewica	93 064	12,56	49
KP	723	8,55	30
Konfederacja	74 670	6,81	11

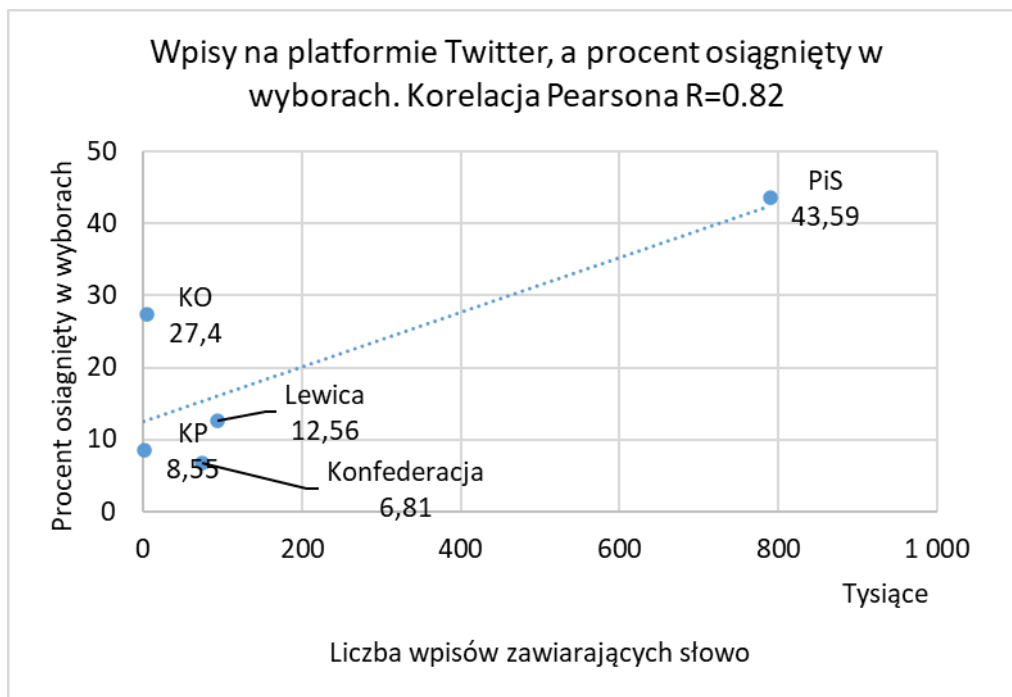
Źródło: opracowanie własne.

Korelacja liczby wpisów zawierających nazwę partii, a wynikiem wyborczym wynosi $R=0,82$ oraz korelacja liczby wpisów zawierających nazwę partii, a mandatami osiąga wynik $R=0,83$. Stanowi to informację, że w obydwu przypadkach mamy do czynienia z silną korelacją dodatnią wspomnianych partii, a osiągniętym wynikiem do Sejmu RP.

⁴⁷⁰ Bsi-sentiment...dz. cyt., 05.05.2021.

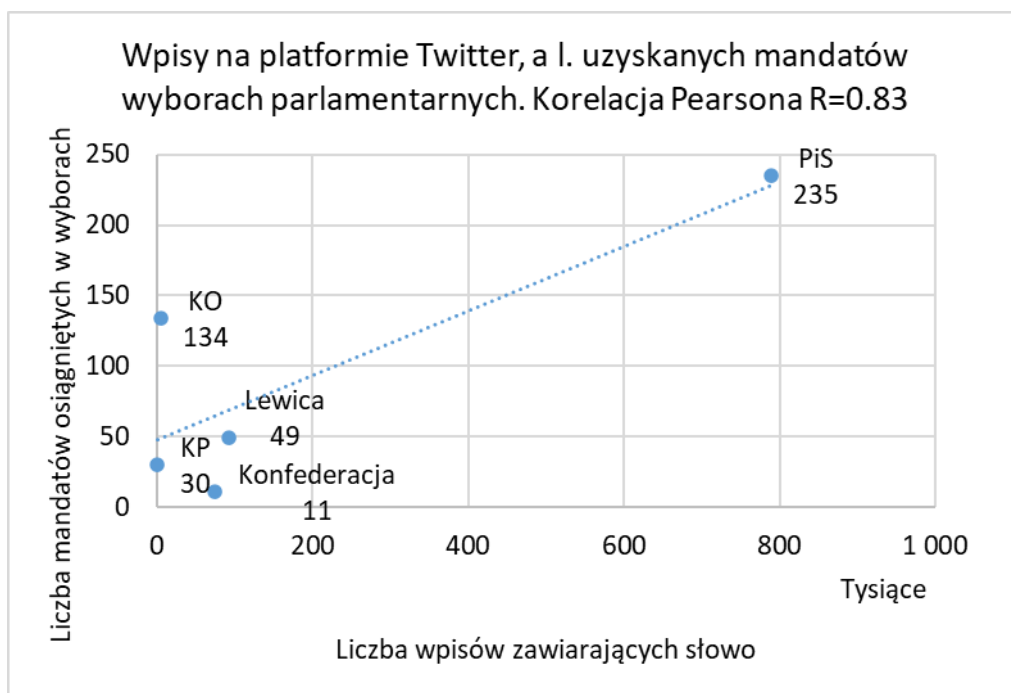
⁴⁷¹ Symboli takich jak: przecinek, cudzysłów, znak procenta, cyfry

Rysunek 131. Liczba wpisów zawierających nazwę partii, a wynik osiągnięty w wyborach do Sejmu RP w 2019. Korelacja Pearsona.



Źródło: opracowanie własne.

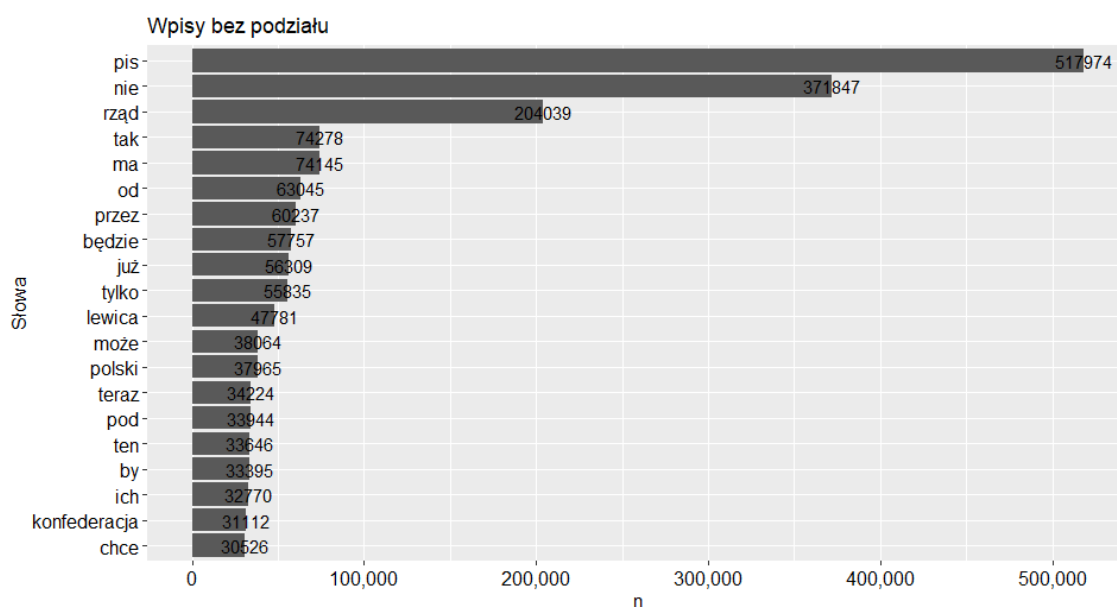
Rysunek 132. Liczba wpisów zawierających nazwę partii, a liczba uzyskanych mandatów do Sejmu RP w 2019. Korelacja Pearsona.



Źródło: opracowanie własne.

Do najczęściej wypowiadanych słów w tweetach należą: *pis* – 517974 razy, *nie* – 371847 razy oraz *rzqd* – 204039 razy.

Rysunek 133. Liczba najczęściej występujących słów w tweetach – bez podziału na partie.



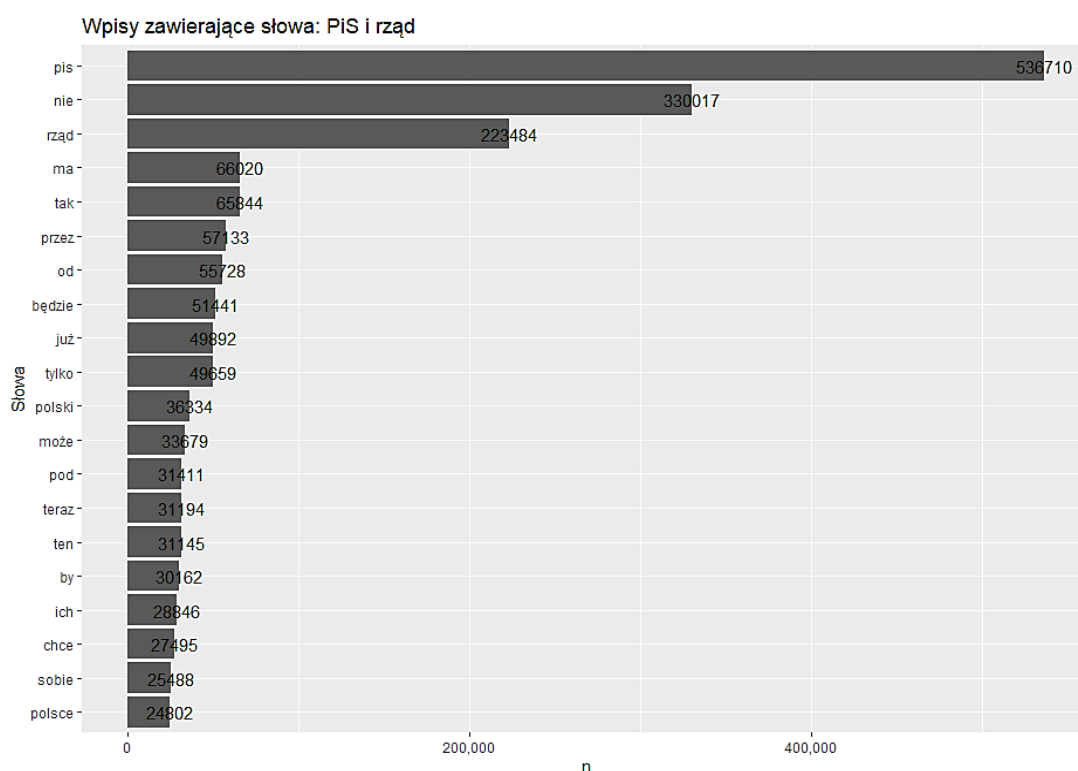
Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

3.9.1. Liczność słów w tweetach z podziałem na nazwę partii użytą w tweecie

Prawo i Sprawiedliwość (PiS)

Do najczęściej wypowiadanych słów w *tweetach* zawierających słowo PiS lub rząd należą: *pis* – 536710 razy, *nie* – 330017 razy oraz *rzqd* – 223484 razy.

Rysunek 134. Liczba najczęściej występujących słów w tweetach – PiS

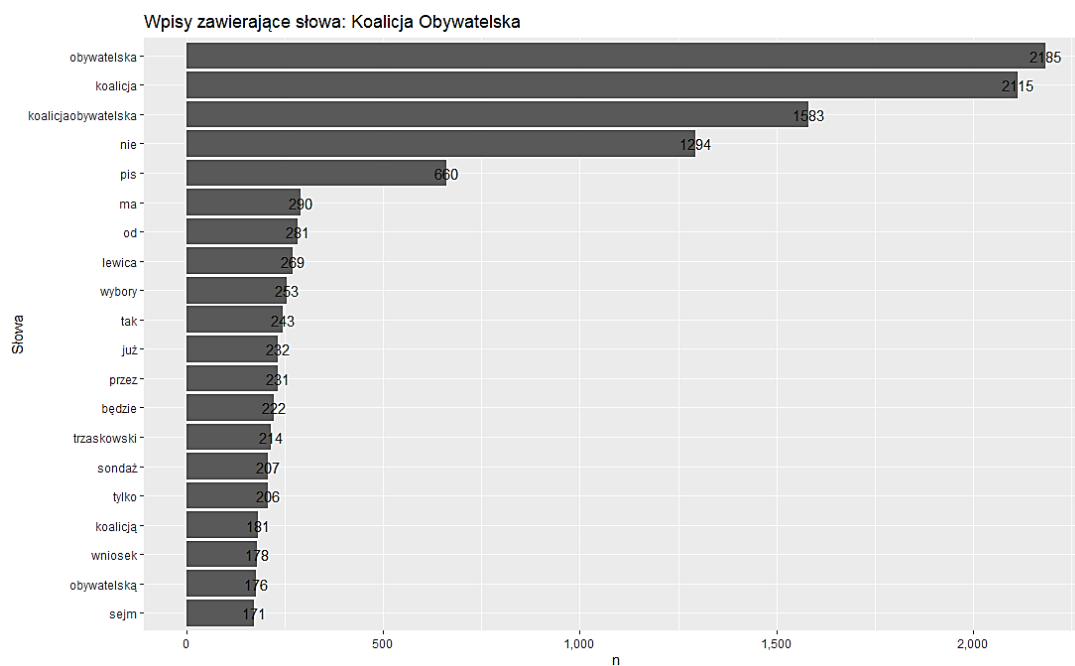


Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Koalicja Obywatelska (KO)

Do najczęściej wypowiedzianych słów w *tweetach* zawierających słowo Koalicja Obywatelska należą: *obywatelska* – 2185 razy, *koalicja* – 2115 razy oraz *koalicjaobywatelska* – 1583 razy.

Rysunek 135. Liczba najczęściej występujących słów w tweetach – KO

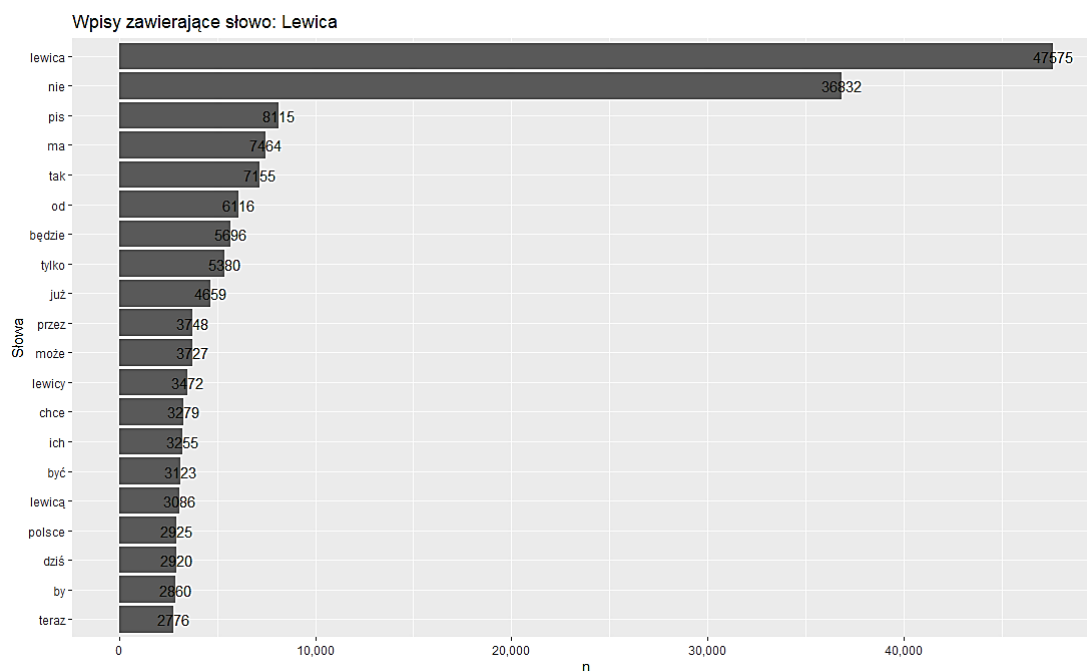


Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Lewica

Do najczęściej wypowiedzianych słów w *tweetach* zawierających słowo Lewica należą: *lewica* – 47575 razy, *nie* – 36832 razy oraz *pis* – 8115 razy.

Rysunek 136. Liczba najczęściej występujących słów w tweetach – Lewica

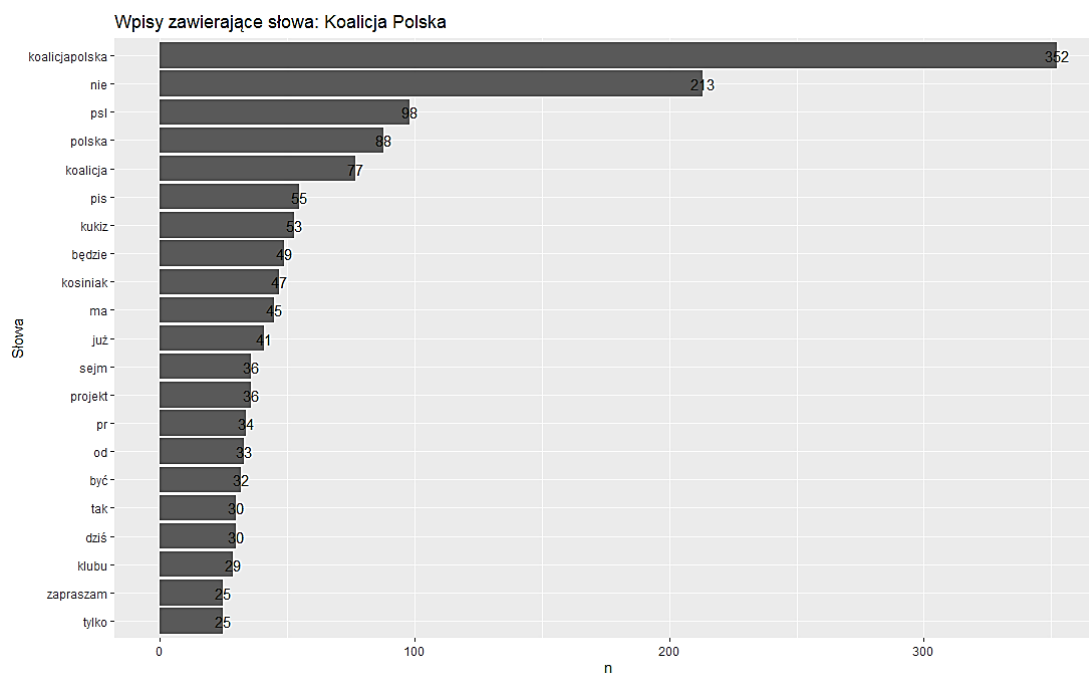


Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Koalicja Polska (KP)

Do najczęściej wypowiadanych słów w *tweetach* zawierających słowo Koalicja Polska należą: *koalicjapolska* – 352 razy, *nie* – 213 razy oraz *psl* – 98 razy.

Rysunek 137. Liczba najczęściej występujących słów w tweetach – KP

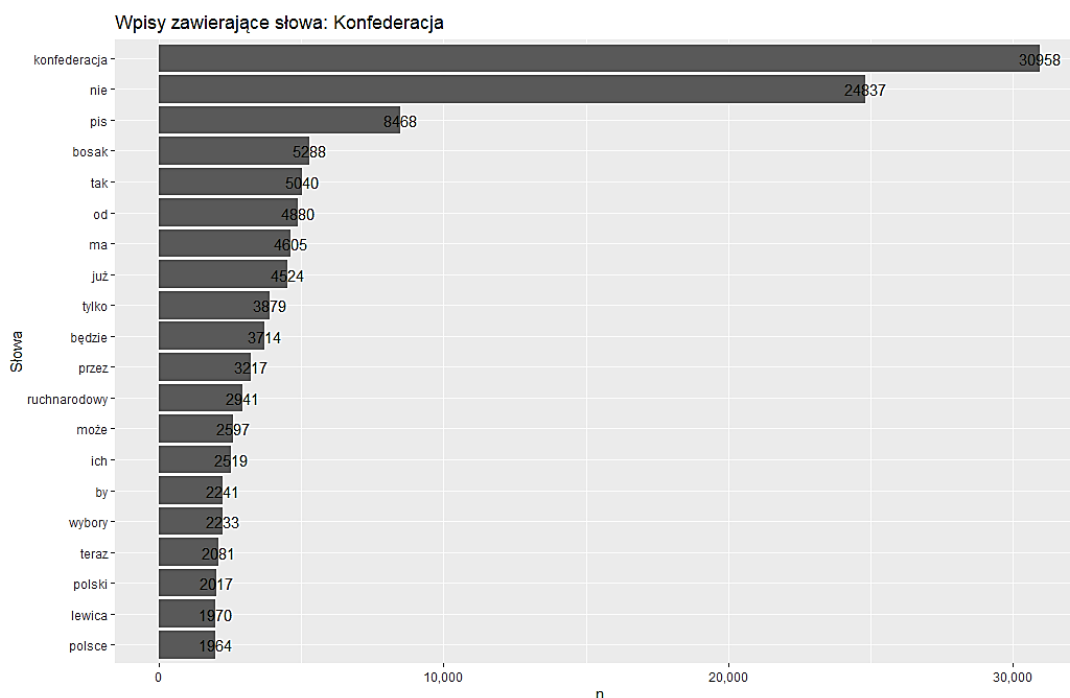


Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Konfederacja

Do najczęściej wypowiadanych słów w *tweetach* zawierających słowo Konfederacja należą: *konfederacja* – 30958 razy, *nie* – 24837 razy oraz *pis* – 8468 razy.

Rysunek 138. Liczba najczęściej występujących słów w tweetach – Konfederacja



Źródło: opracowanie własne z wykorzystaniem biblioteki ggplot2.

Podsumowanie

Podsumowując, widać wyraźnie, że w najczęściej używanych słowach są znaczące różnice jeśli chodzi o wypowiedzi posłów z mównicy sejmowej, a wypowiedzi internetowe. Jest co całkowicie naturalne i spodziewane, że wypowiedzi oficjalne, polityczne i nieoficjalne uczestników komunikowania politycznego będą miały inne formy i zawartość słów w swoich słownikach.

W słowach o najwyższej liczności na platformie Twitter wyraźnie widać tendencję do grupowania w jednej wypowiedzi nazw dwóch lub więcej partii, m.in. w grupie osób wspominających *koalicję obywatelską* również pojawia się słowo *pis*. W grupie osób wspominających *konfederację* również pojawia się słowo *pis* i *lewica*. Jest to wyraźna informacja, że jednoznaczna klasyfikacja na zwolenników czy przeciwników partii wymaga zbudowania osobnych modeli klasyfikacyjnych szacujących ich przynależność oraz stopień przypisania. Zbudowanie takich modeli predykcyjnych na podstawie samych wpisów internetowych powinno zostać ujęte w osobnym obszarze badawczym wraz z uwzględnieniem szeregu metod, jak analizy sentymentów każdego z wpisów oraz innych rozwiązań m. in. (I) jednym z przykładów mogłoby być zbudowanie modeli dla

partii politycznych na podstawie wpisów najaktywniejszych jej członków wypowiadających się w mediach ich komentarzy i reakcji internautów, (II) drugim podejściem mogłoby być wybór najpopularniejszych wpisów oraz ręczne przypisanie poparcia dla partii w przypadku zadeklarowanych politycznie profili i potraktowanie stworzonego zbioru jako zbiór testowy.

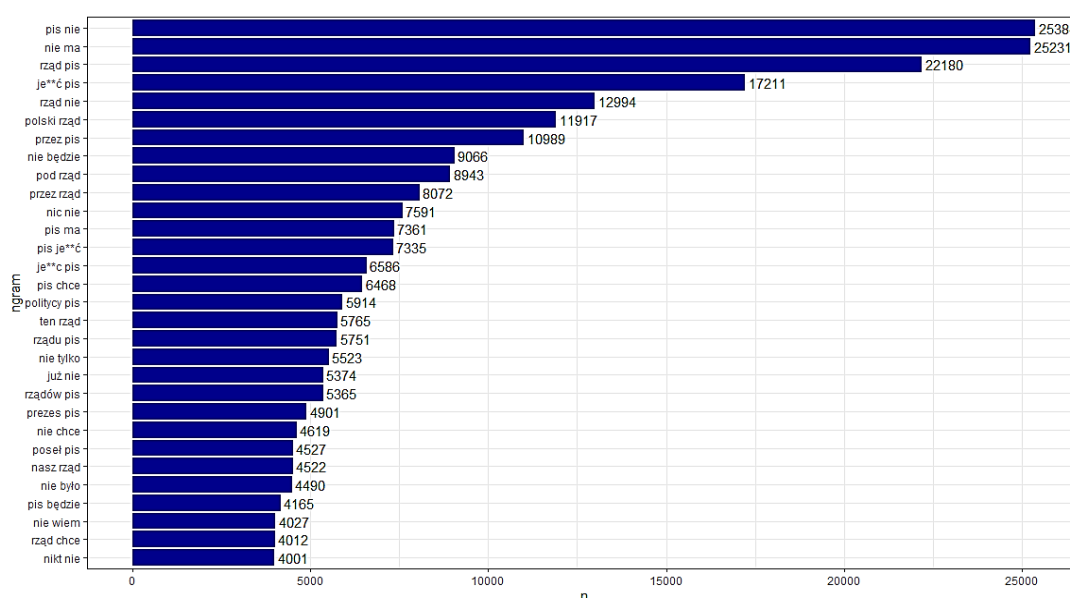
3.9.2. Ngramy

Prawo i Sprawiedliwość (PiS)

Do najczęściej występujących dwu-wyrazowych należą: *pis nie*, *nie ma*, *rząd pis*.

Do najczęściej występujących trzy-wyrazowych należą: *je**ć pis je**ć*, *pis je**ć pis*⁴⁷², *pis nie ma*.

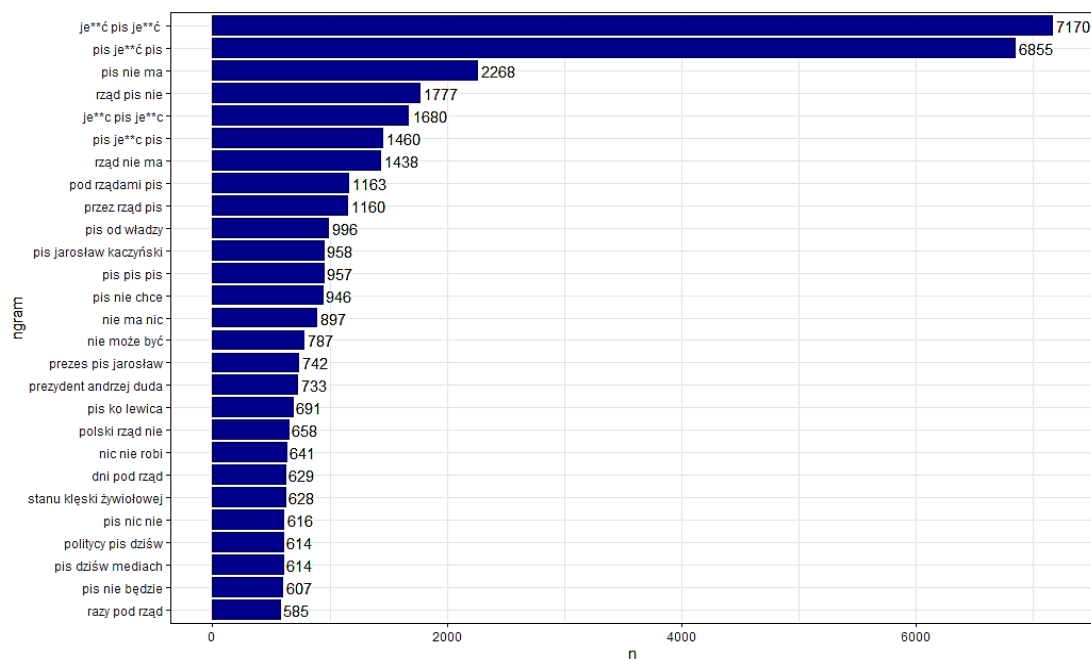
Rysunek 139. Najczęściej występujące wyrazy 2-częściowe. Partia: PiS



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

⁴⁷² Hasła głoszone w protestach przeciwko orzeczeniu Trybunału Konstytucyjnego z 22 października 2020, które delegalizuje w Polsce wykonywanie aborcji z powodu wad płodu, zastąpione symbolami ośmiu gwiazdek ***** ***.

Rysunek 140. Najczęściej występujące wyrazy 3-częściowe. Partia: PiS



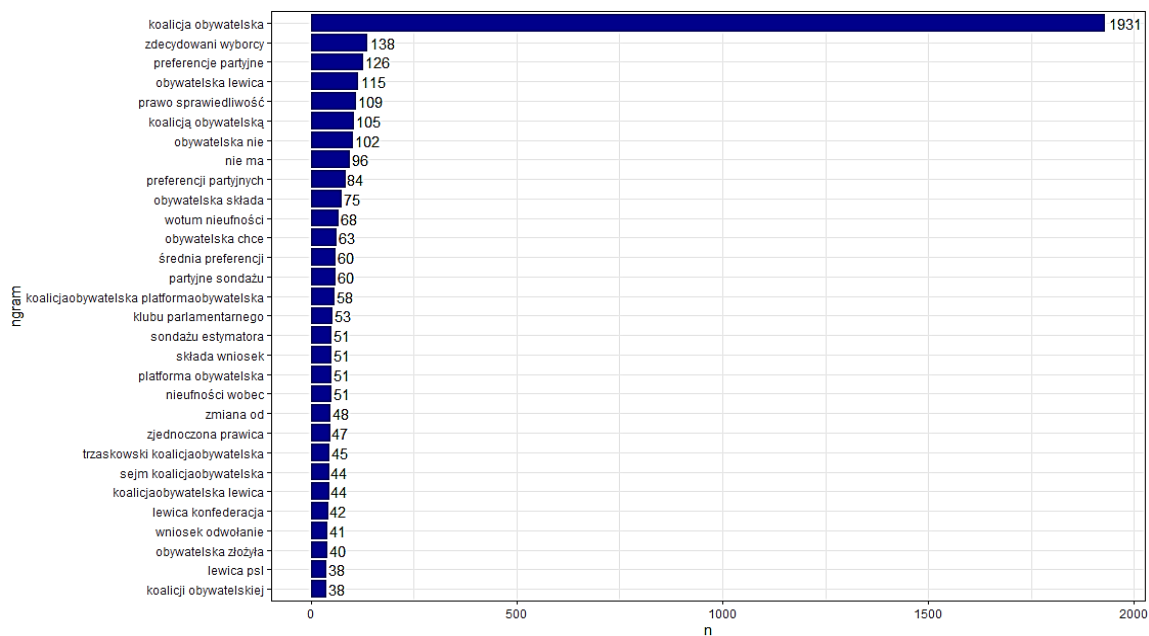
Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Koalicja Obywatelska (KO)

Do najczęściej występujących dwu-wyrazowych należą: *koalicja obywatelska*, *zdecydowani wyborcy*, *preferencje partyjne*.

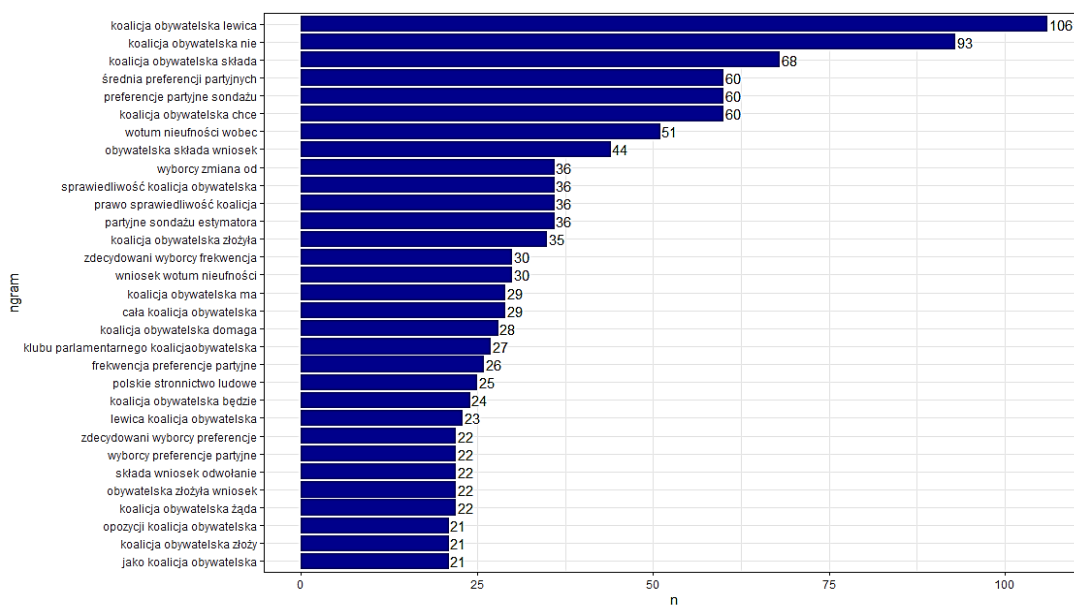
Do najczęściej występujących trzy-wyrazowych należą: *koalicja obywatelska lewica*, *koalicja obywatelska nie*, *koalicja obywatelska składa*.

Rysunek 141. Najczęściej występujące wyrazy 2-częściowe. Partia: KO



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 142. Najczęściej występujące wyrazy 3-częściowe. Partia: KO



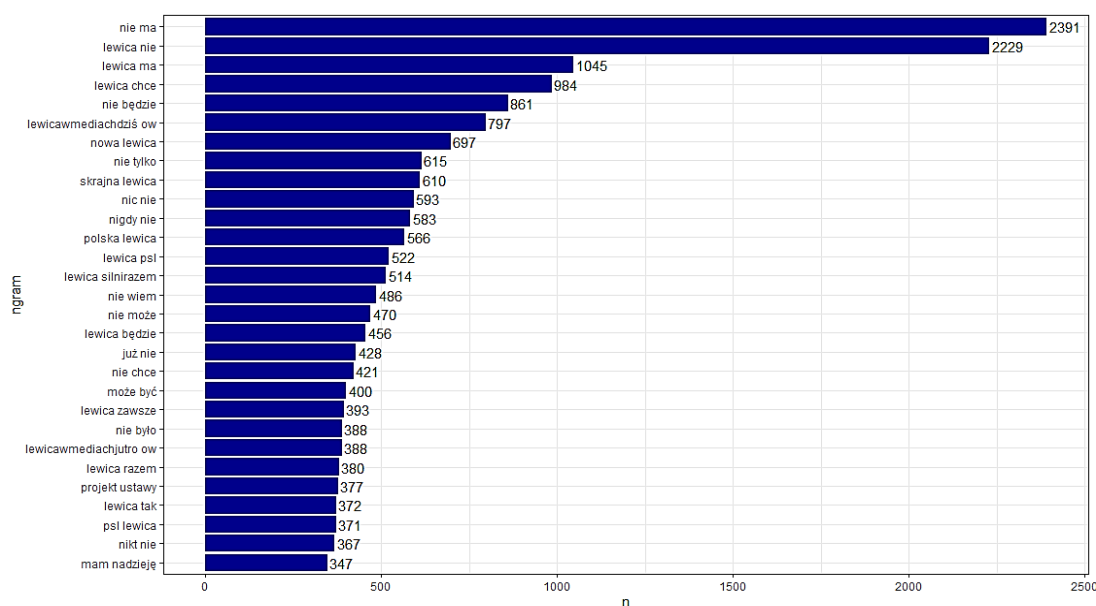
Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Lewica

Do najczęściej występujących dwu-wyrazowych należą: *nie ma*, *lewica nie*, *lewica ma*.

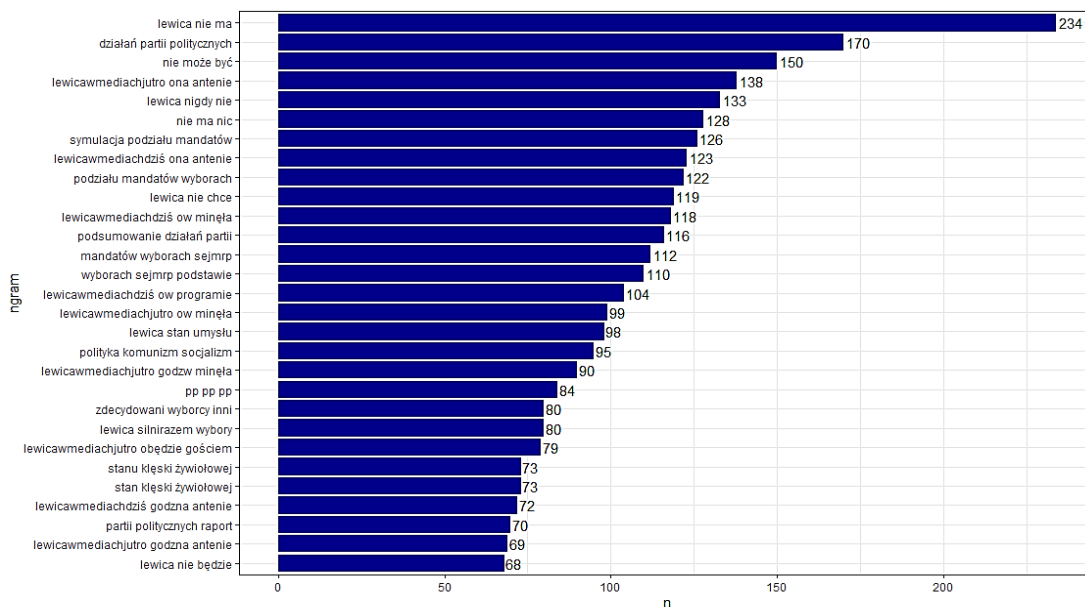
Do najczęściej występujących trzy-wyrazowych należą: *lewica nie ma*, *działań partii politycznych*, *nie może być*.

Rysunek 143. Najczęściej występujące wyrazy 2-częściowe. Partia: Lewica



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 144. Najczęściej występujące wyrazy 3-częściowe. Partia: Lewica



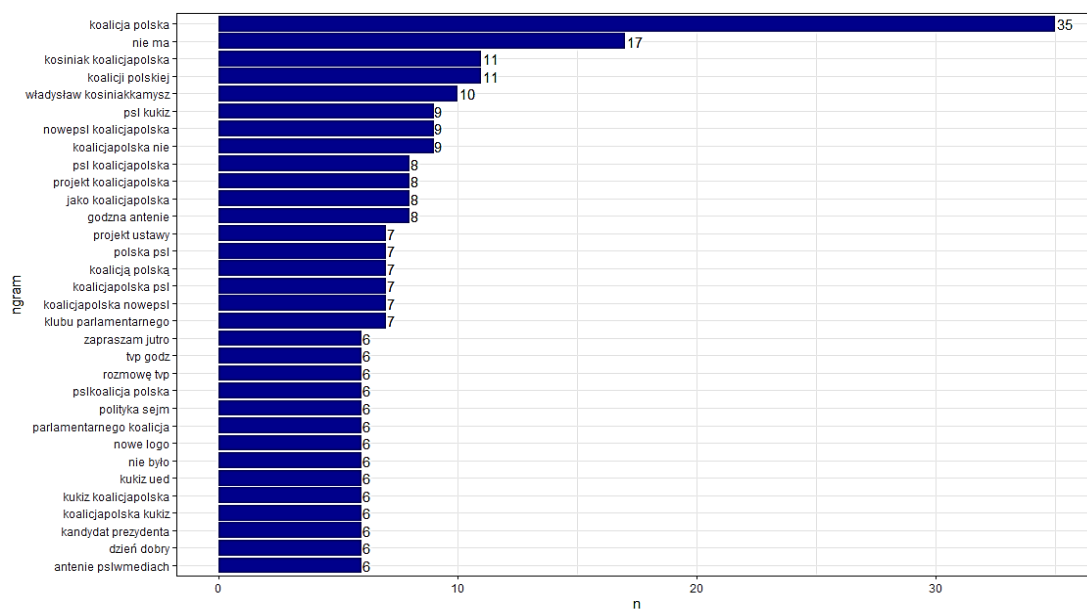
Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Koalicja Polska (KP)

Do najczęściej występujących dwu-wyrazowych należą: *koalicja polska*, *nie ma*, *ko-siniak koalicjapolska*.

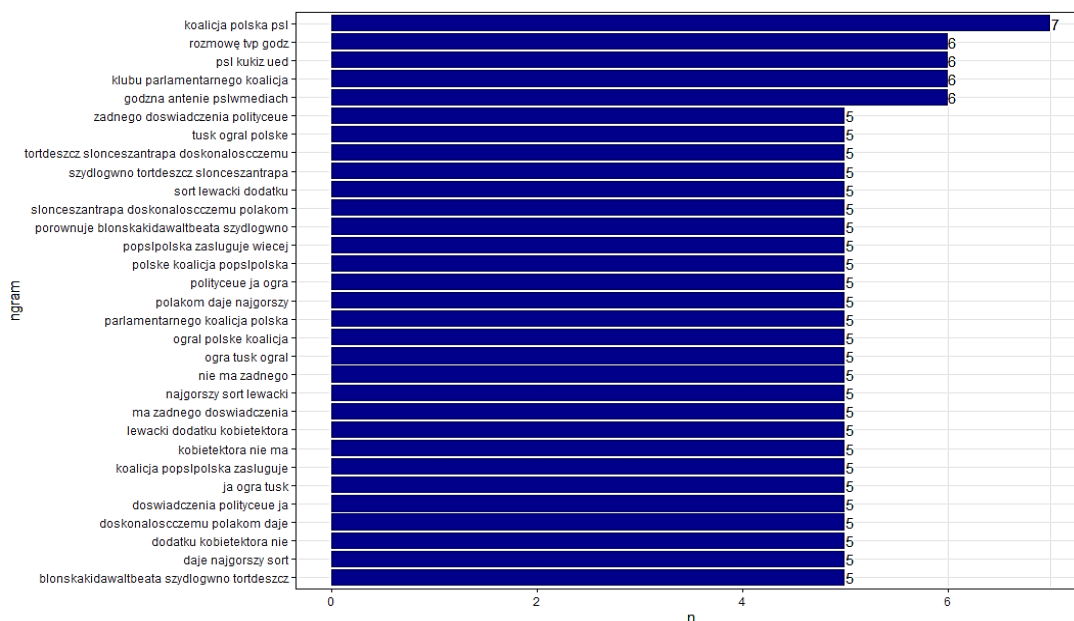
Do najczęściej występujących trzy-wyrazowych należą: *koalicja polska psl*, *rozmowa tvp godz*, *psl kukiz ued*.

Rysunek 145. Najczęściej występujące wyrazy 2-częściowe. Partia: KP



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 146. Najczęściej występujące wyrazy 3-częściowe. Partia: KP



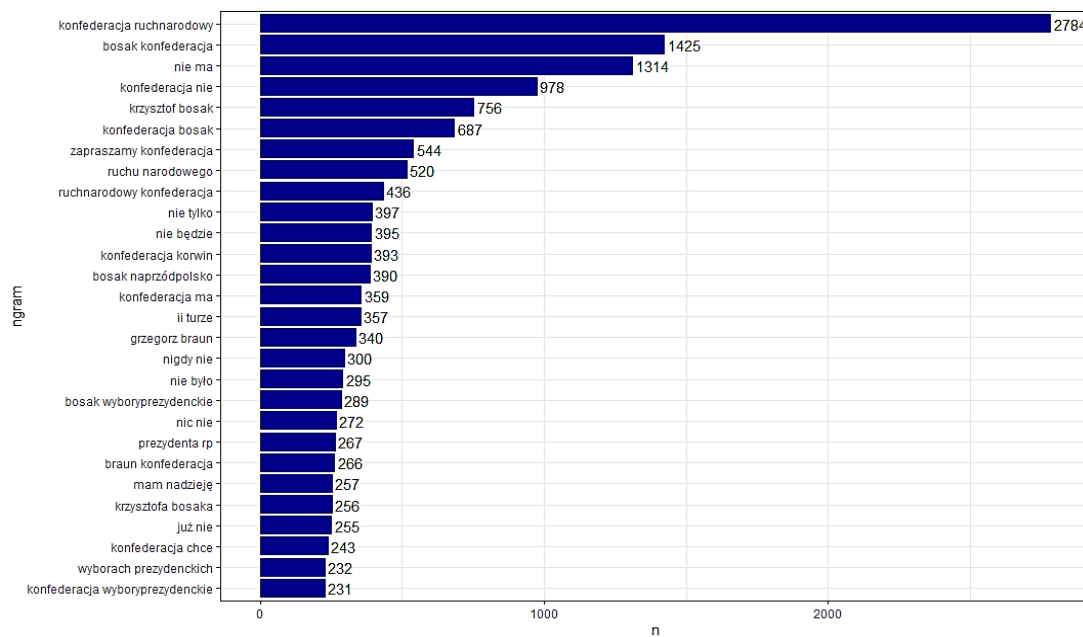
Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Konfederacja

Do najczęściej występujących dwu-wyrazowych należą: *konfederacja ruchnarodowy*, *bosak konfederacja*, *nie ma*.

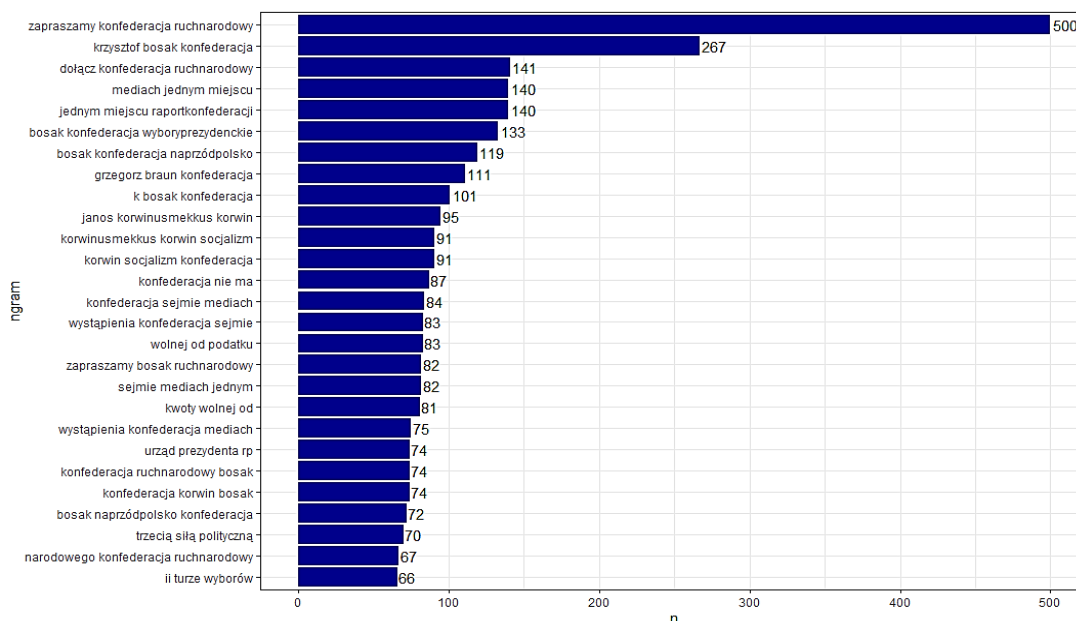
Do najczęściej występujących trzy-wyrazowych należą: *zapraszamy konfederacja ruchnarodowy*, *krzysztof bosak konfederacja*, *dołącz konfederacja ruchnarodowy*.

Rysunek 147. Najczęściej występujące wyrazy 2-częściowe. Partia: Konfederacja



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Rysunek 148. Najczęściej występujące wyrazy 3-częściowe. Partia: Konfederacja



Źródło: opracowanie własne z wykorzystaniem pakietu ggplot2.

Podsumowanie

Podsumowując widać wyraźnie, że w najbardziej używanych zwrotach są różnice jeśli chodzi o wypowiedzi posłów (z mównicy sejmowej), a wypowiedzi internetowe – podobnie jak w przypadku słów najczęściej używanych.

W przypadku osób wspominających partię rządzącą znaczącą częścią są wpisy zawierające niecenzuralne zwroty, szczególnie jeśli chodzi o 8 gwiazdek symbolizujących ruch krytyki rządu⁴⁷³, dodatkowo są wyraźne nawiązania do polityków PiS, prezesa PiS, prezydenta Andrzeja Dudy oraz zwrotu *pis od władzy*.

W przypadku KO, najczęstsze zwroty dotyczą nazwy partii oraz nazw innych partii jak PiS czy Lewica, również znaczącą liczbę prezentują tematy związane z wotum nieufności, składaniu wniosków oraz zwrotu: *preferencji partyjnych*.

Jeśli chodzi o wpis dotyczące klubu Lewicy, to jak w poprzednich przypadkach są wyraźne wzmianki o nazwie partii, jak i kombinacje słowa *nie*. W przypadku

⁴⁷³Co znaczy osiem gwiazdek ruch ośmiu gwiazd to inicjatywa anty-pis, <https://natemat.pl/314263,co-znaczy-osiem-gwiazdek-ruch-osmiu-gwiazd-to-inicjatywa-anty-pis-rozmowa>, 20.11.2022.

najczęstszych *bigramów* (obiektów dwuelementowych) na pierwszej pozycji plasuje się zwrot *nie ma*, na drugim *lewica nie*, a na piątym miejscu jeśli chodzi o licznosc kombinacja *nie będzie*. W przypadku *trigramów* (zwrotów trzelementowych) jest podobna sytuacja: na pierwszym miejscu zwrot *lewica nie ma*, na trzecim miejscu *nie może być*, na piątym *lewica nigdy nie*.

W przypadku Koalicji Polskiej zwroty *n*-elementowe stanowią nazwę partii oraz informację o uczestnictwie partii w mediach, imię i nazwisko przewodniczącego partii PSL oraz słowa *kukiz*.

Jeśli chodzi o Konfederację to znaczną część wpisów stanowią nazwa partii oraz posłów Krzysztofa Bosaka, Grzegorza Brauna oraz Janusza Korwin-Mikkego. Na wysokich pozycjach są również nawiązania do Ruchu Narodowego, informacji partii w mediach.

W zwrotach o najwyższej liczności na platformie *Twitter* wyraźnie widać, że wspominając daną partię również były nawiązania do innych partii. Ze względu na niedużą liczbę wspomnień partii *koalicji polskiej* istniejące wpisy zwykle sugerują wypowiedzi informacyjne. Każda z partii ma pewne charakterystyki i wzorce należące do grupy osób wspominających te partie.

3.9.3. Korelacje oraz podobieństwo słowników pomiędzy grupami

W kolejnym etapie badań obliczono siłę połączenia pomiędzy poszczególnymi słowami. Za siłę powiązań przyjęto współczynnik korelacji *phi-coefficient*⁴⁷⁴. Wyniki odfiltrowano – za graniczny wynik zaprezentowania na grafach przyjęto minimum 50 wystąpień jednej kombinacji słów⁴⁷⁵. Każda z grup z wspomnieniami partii w *tweetach* tworzy różne kategorie wiadomości. Do wizualizacji posłużono się biblioteką *ggraph*⁴⁷⁶.

⁴⁷⁴ C.P. Dancey, J. Reidy, *Statistics without Maths for Psychology*, Pearson Education 2017, s. 280-285.; Biblioteka widyr, https://www.rdocumentation.org/packages/widyr/versions/0.1.4/topics/pairwise_cor, 10.05.2021,

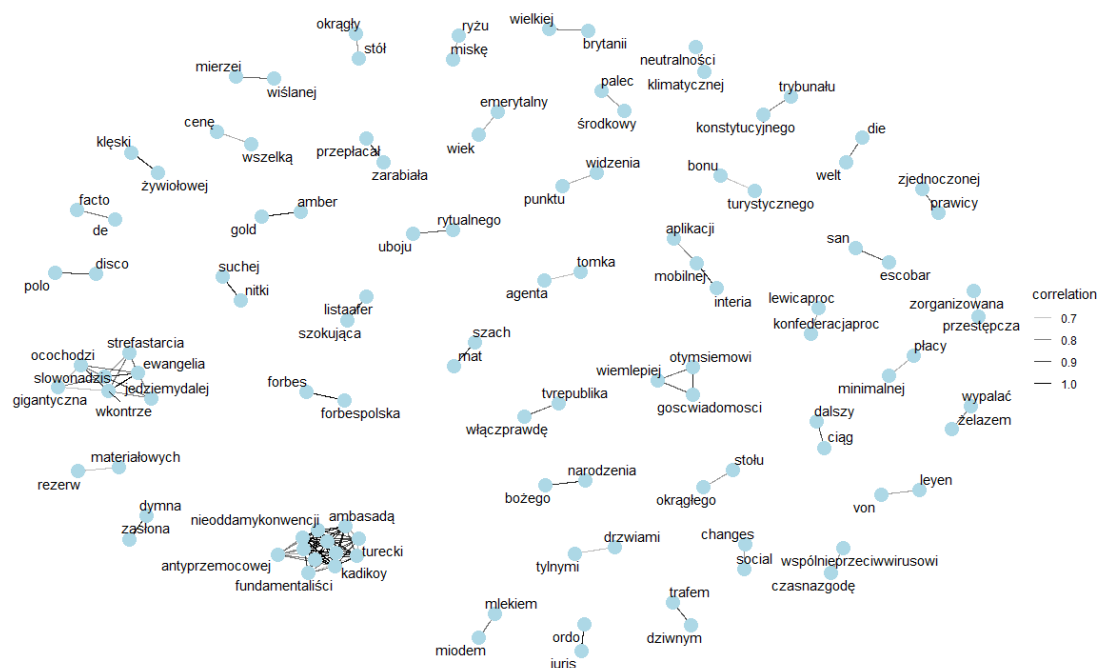
⁴⁷⁵ W przypadku *Koalicji Polskiej* zmieniono tę wartość na 10 ze względu na niską liczbę tweetów w stosunku do innych wspomnień partii.

⁴⁷⁶ Biblioteka *ggraph*, <https://www.rdocumentation.org/packages/ggraph/versions/2.1.0>, 11.05.2022.

Prawo i Sprawiedliwość (PiS)

Do najsilniejszych kombinacji należą między innymi informacje o *nieoddaniu konwencji antyprzemocowej* (lewy dolny róg grafu).

Rysunek 149. Najsilniejsze połączenia słów mierzone phi-coefficient. Partia: PiS.



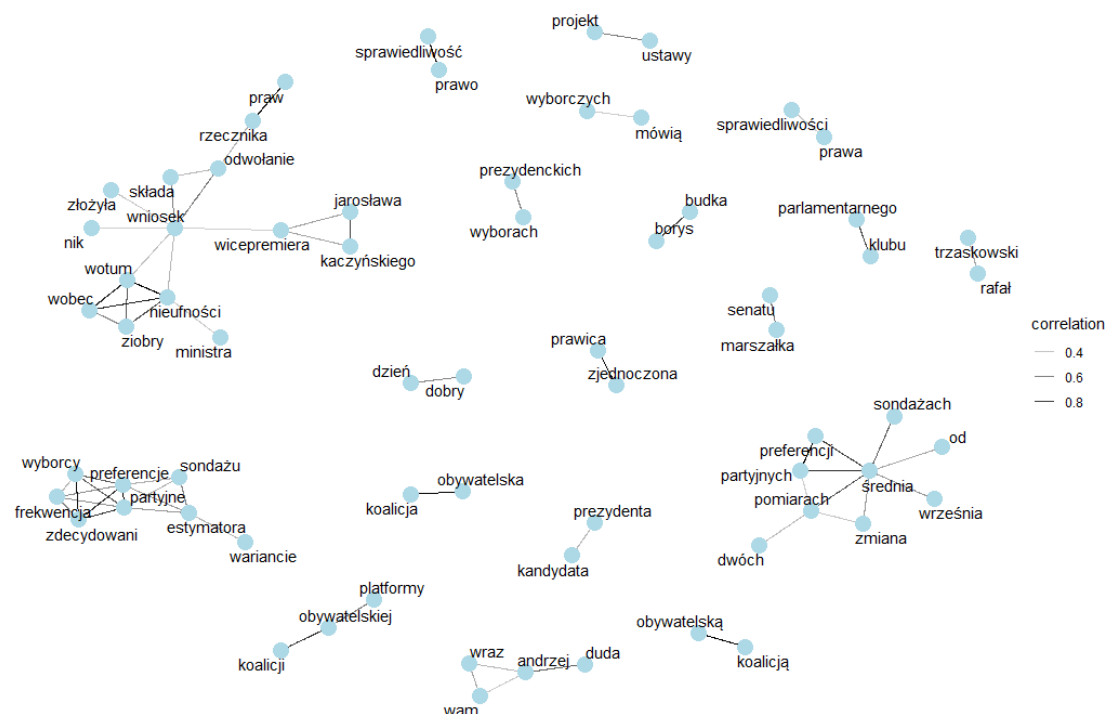
Źródło: opracowanie własne z wykorzystaniem pakietu ggraph.

Koalicja Obywatelska (KO)

Do najsilniejszych kombinacji należą między innymi informacje o *wotum nieufności wobec ministra ziobry* (lewa strona grafu).



Rysunek 150. Najsilniejsze połączenia słów mierzone phi-coefficient. Partia: KO

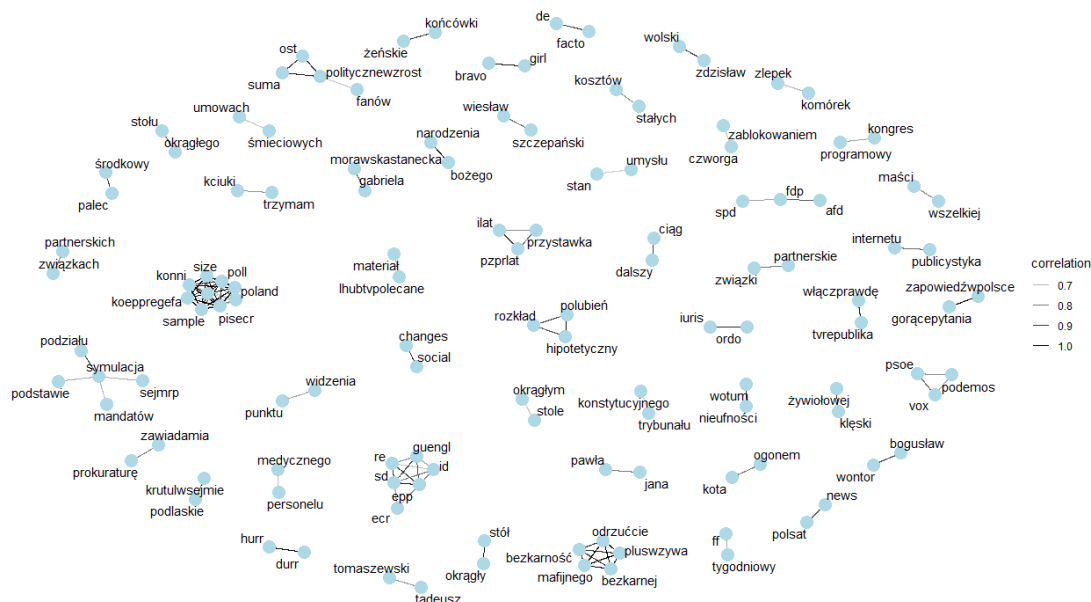


Źródło: opracowanie własne z wykorzystaniem pakietu ggraph.

Lewica

Do najsilniejszych kombinacji należą między innymi informacje o *odrzućcia bezkarności* (dolna część grafu).

Rysunek 151. Najsilniejsze połączenia słów mierzone phi-coefficient. Partia: Lewica.



Źródło: opracowanie własne z wykorzystaniem pakietu ggraph.

Koalicja Polska (KP)

Do najsilniejszych kombinacji należą między innymi informacje o zaproszeniu na program z udziałem psl (lewa strona grafu).

Rysunek 152. Najsilniejsze połączenia słów mierzone phi-coeficient. Partia: KP



Źródło: opracowanie własne z wykorzystaniem pakietu ggraph.

Konfederacja

Do najsilniejszych kombinacji należą między innymi informacje o *zapewnieniu rzetelnej opieki zdrowotnej* (dolna część grafu).

Tabela 32. Index Jaccarda pomiędzy partiami we wspomnieniach

	PiS	KO	Lewica	KP	Konfederacja
PiS		0,46	0,63	0,29	0,58
KO	0,46		0,47	0,28	0,46
Lewica	0,63	0,47		0,31	0,62
KP	0,29	0,29	0,31		0,31
Konfederacja	0,58	0,46	0,62	0,31	

Źródło: opracowanie własne

Zasób słów, tak jak w poprzednim przypadku (podrozdział 3.5.11.), gdzie zmierzone zostało podobieństwo słów polityków będzie na jakimś poziomie podobny dlatego, że słowa najczęściej używane czasami powtarzają się dla każdej z partii i wartości będą pokazywać jakąś część wspólną zbiorów. Z Tabeli 32. można odczytać, że najwyższą wartością podobieństwa cechuje się PiS oraz Lewica, gdzie parametr podobieństwa jest na poziomie 0,63. Najniższą wartość prezentuje kombinacja KP oraz KO (parametr *Jaccard index* równa się 0,29).

Podsumowanie

Każda z grup wspominających nazwy poszczególnych partii tworzy własne powiązania słów najczęściej skorelowanych ze sobą, na przykład internauci wspominający partię *pis* tworzą silne relacje, między innymi informację o *nieoddaniu konwencji antyprzemocowej*. Wspominający *koalicję obywatelską* mówią najczęściej o *wotum nieufności wobec ministra Ziobry*. Wspominający domagają się m.in. *odrzućenia bezkarności*. Wspominający *koalicję polską* zapraszają *na program z udziałem psl*, a wspominający *konfederację* mówią m.in. o *zapewnieniu rzetelnej opieki zdrowotnej*. Każda z grup wspominających dane partie tworzy swoje grupy wiadomości. Porównując indeks Jaccarda można zauważyć, że wspomnienia partii KP (których jest stosunkowo niedużo porównując ze wspomnieniami innych partii) odbiegają od reszty i nie są podobne do innych partii (*Jaccard index* waha się między 0,28 – 0,31).

3.10. Sentymenty w tweetach

Głównym celem tej części był pomiar emocji biorąc pod uwagę słowa używane przez internautów w serwisie społecznościowym Twitter. Aby dokonać tego pomiaru podobnie jak w przypadku części wypowiedzi z Sejmu RP przez posłów (Rozdział 3.8.) posłużono się biblioteką słów dostępną z serwisu *Słowsieć – Politechniki Wrocławskiej*

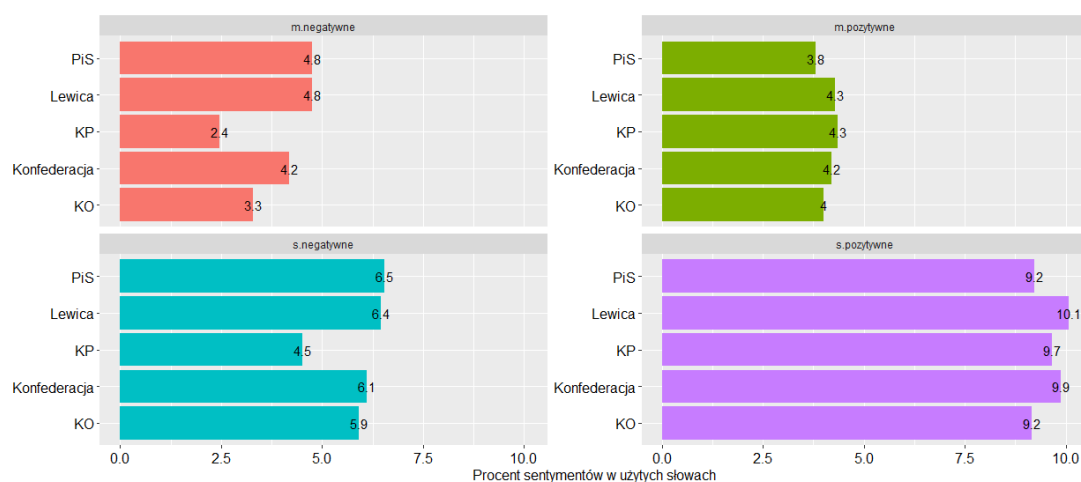
(plwordnet.pwr.wroc.pl)⁴⁷⁷. Dodatkowo, również aby wzmocnić pomiar i uzyskać nacechowanie podstawowych słów jak również ich odmian posłużono się biblioteką *PoliMorfologik*⁴⁷⁸. Złączono bazy z nacechowaniem (*Słowosieć*) oraz odmianami słów (*PoliMorfologik*). Trzeba mieć na uwadze, że wypowiedzi często mogły być pisane w sposób niechlujny, stąd błędy językowe w niektórych przypadkach nie umożliwiły pomiaru w sposób idealny. Nacechowanie słów pozytywne/negatywne definiuje słownik *Słowosieć*.

- Słów o nacechowaniu mocno negatywnym najwięcej notuje PiS (4,8%) oraz Lewica (4,8%), najmniej KP (2,4%);
- Słów o nacechowaniu słabo negatywnym najwięcej notuje PiS (6,5%), najmniej KP (4,5%);
- Słów o nacechowaniu mocno pozytywnym najwięcej notuje Lewica oraz KP (4,3%), najmniej PiS (3,8%);
- Słów o nacechowaniu słabo pozytywnym najwięcej notuje Lewica (10,1%), najmniej KO i PiS (9,2%).

⁴⁷⁷ A. Dziob, M. Piasecki, E. Rudnicka, *plWordNet 4.1—a Linguistically ...*, dz. cyt.

⁴⁷⁸ M. Miłkowski, Biblioteka odmian słów, <https://github.com/morfologik/polimorfologik/blob/master/src/LICENSE.txt>, 28.12.2021; Copyright (c) \$copyright_date, Marcin Miłkowski All rights reserved.

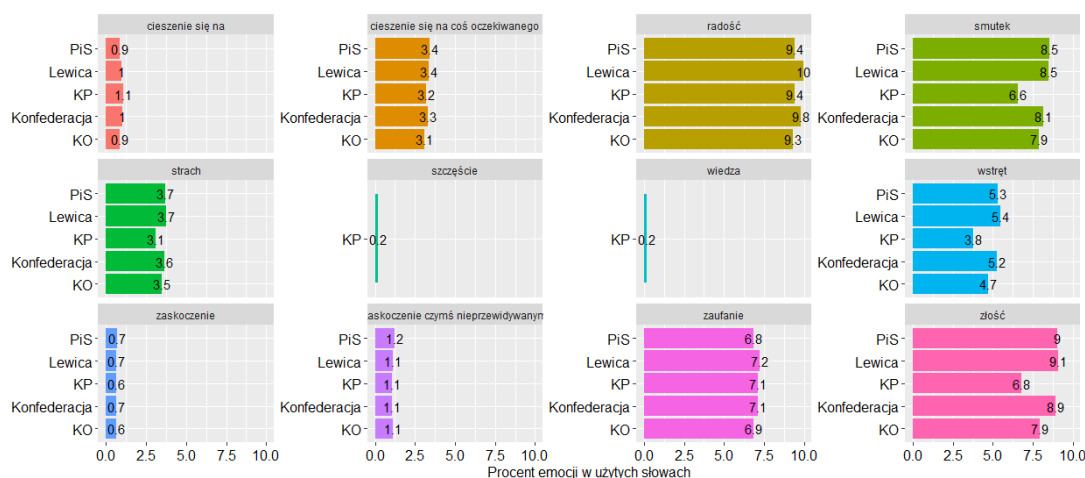
Rysunek 154. Procentowe rozłożenie sentymentów w użytych tweetach z podziałem na partie



Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Najczęściej spotykane emocje to radość, smutek, zaufanie oraz złość. Jeśli chodzi o radość najwyższe wartości notuje Lewica (10%). W przypadku emocji smutku najwyższe wartości procentowe notuje PiS i Lewica (po 8,5%). Jeśli chodzi o zaufanie najwyższą wartość uzyskuje Lewica (7,2%). Najwyższą wartość jeśli chodzi o złość notuje również Lewica (9,1%).

Rysunek 155. Procent emocji w użytych tweetach z podziałem na partie



Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Wyniki procentowe nacechowania w jednostce czasu

W przypadku słów o nacechowaniu mocno i słabo pozytywnym w ujęciu procentowym *tweety* wspominające partię KP przodują w badanym okresie. W przypadku słów mocno i słabo negatywnym różnicie się zacierają i najwyższe wartości w badanym okresie prezentują *tweety* wspominające partię: PiS i KP.

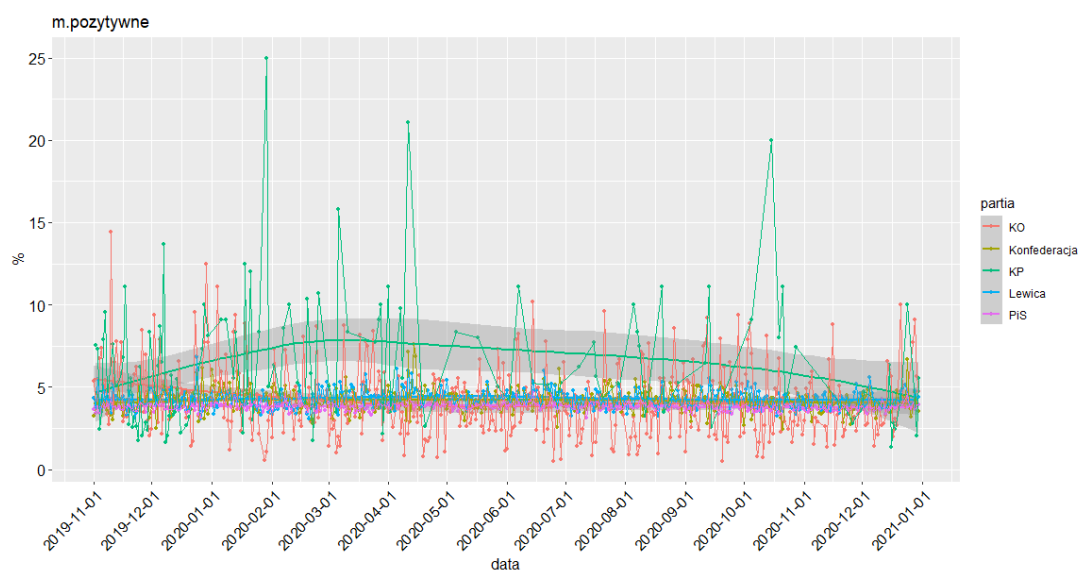
Z powyższej analizy można wyciągnąć następujące wnioski:

Słów o nacechowaniu mocno negatywnym najwięcej notuje wspomnienie partii PiS (4,8%) oraz Lewica (4,8%), najmniej KP (2,4%). Słów o nacechowaniu słabo negatywnym najwięcej notuje PiS (4,2%), najmniej KP (4,5%).

Słów o nacechowaniu mocno pozytywnym najwięcej notuje Lewica oraz KP (4,3%), najmniej PiS (3,8%). Słów o nacechowaniu słabo pozytywnym najwięcej notuje Lewica (10,1%), najmniej KO i PiS (9,2%).

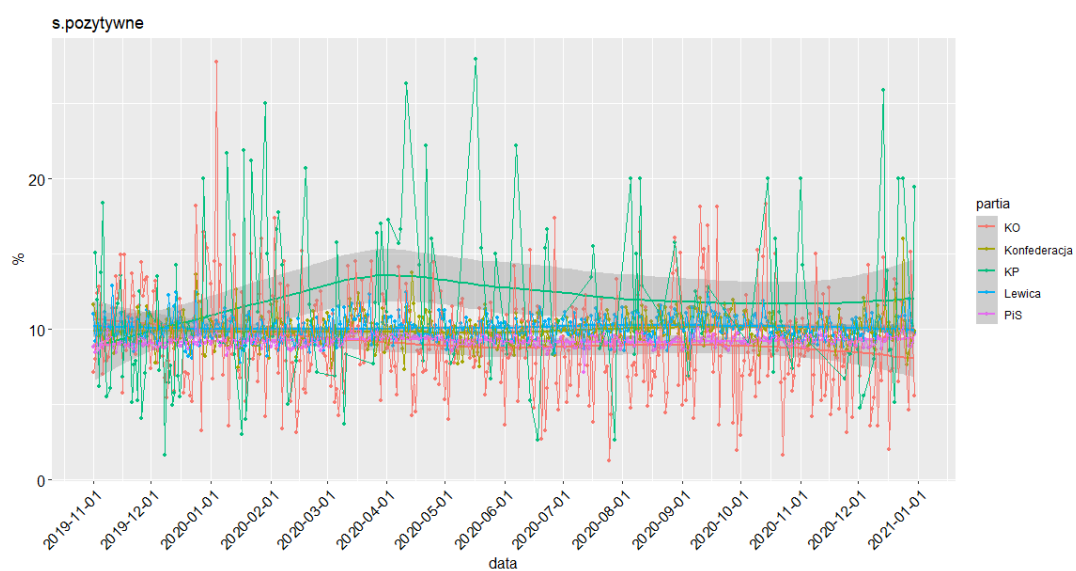
Biorąc pod uwagę emocję w użytych słowach, to jeśli chodzi o *radość*, *zauwanie*, *złość* najwyższe wartości notuje Lewica (kolejno 10%, 7,2% i 9,1%). W przypadku emocji *smutku* najwyższe wartości procentowe notuje PiS i Lewica (po 8,5%).

Rysunek 156. Procent emocji mocno pozytywnych w użytych tweetach z podziałem na wspomniane partie w czasie



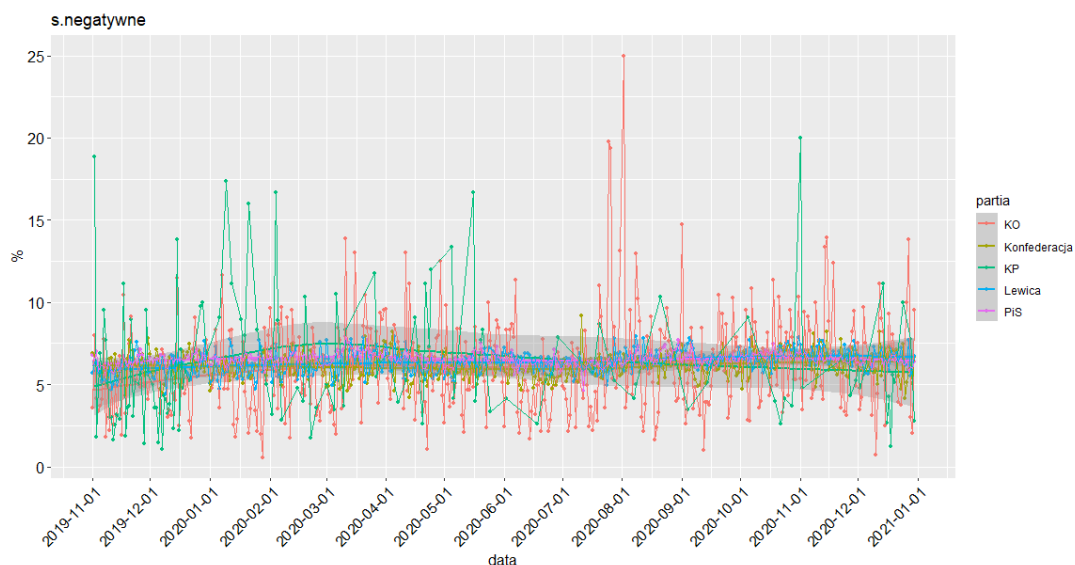
Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Rysunek 157. Procent emocji słabo pozytywnych w użytych tweetach z podziałem na wspomniane partie w czasie



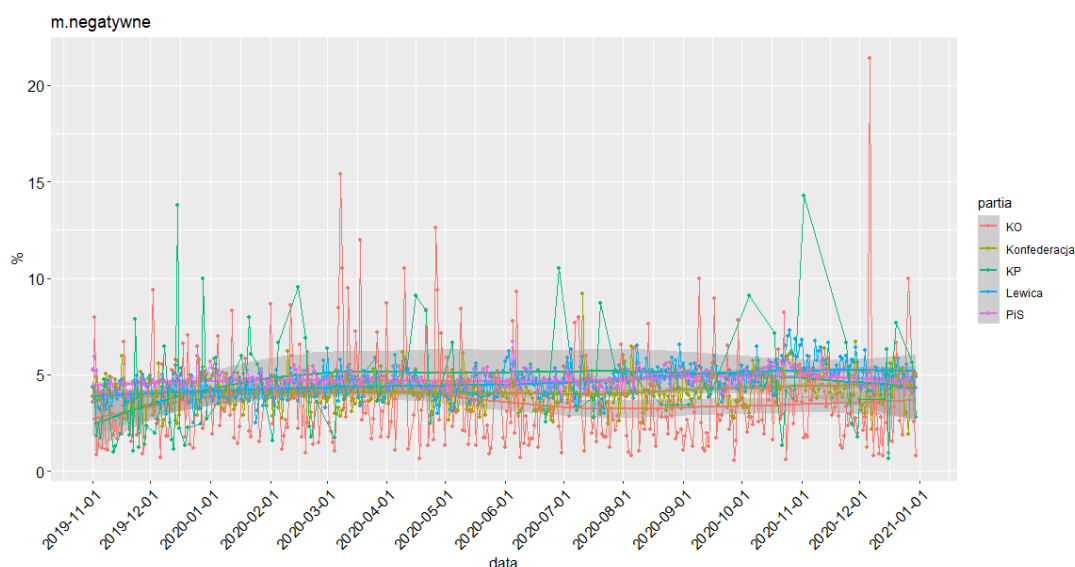
Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Rysunek 158. Procent emocji słabo negatywnych w użytych tweetach z podziałem na wspomniane partie w czasie



Źródło: opracowanie własne z użyciem biblioteki ggplot2.

Rysunek 159. Procent emocji mocno negatywnych w użytych tweetach z podziałem na wspomniane partie w czasie



Źródło: opracowanie własne z użyciem biblioteki ggplot2.

3.12. Podsumowanie rozdziału

W tym rozdziale opisano proces analizy danych z podziałem na wypowiedzi posłów (podrozdziały od 3.1 do 3.8) oraz wypowiedzi użytkowników mediów społecznościowych (podrozdziały 3.9 do 3.10) wyrażających poglądy polityczne. Dokonano szeregu

analiz statystycznych pokazujących częstości słów oraz zwrotów w postaci *bigramów* i *trigramów*. Wykonano klasteryzację słów oraz dokonano porównania podobieństwa w postaci dendrogramów. Dokonano analizy sentymentów w poszczególnych grupach oraz obliczono analizę koszykową słów współwystępujących razem dla polityków oraz analizę korelacji dla wypowiedzi użytkowników mediów społecznościowych.

Analizy pokazały znaczne różnice w wypowiedziach zarówno w grupach polityków, jak i użytkowników mediów społecznościowych.

W przypadku wypowiedzi polityków podobieństwa słów typu: (*pani, marszałek*), (*panie, marszałku*), (*wysoka, izba*), (*będą, będzie*) są widoczne dla każdej partii. *Jaccard index* pomiędzy 10 000 najczęściej używanymi słowami w każdej z partii to wskazuje najwyższe wartości pomiędzy KO oraz Lewica (0,62). Najniższą wartość prezentuje zestawienie PiS i Konfederacji (0,44). Słów o nacechowaniu mocno i słabo negatywnym najwięcej notuje Konfederacja (8%), najmniej PiS (6,2%). Słów o nacechowaniu mocno pozytywnym najwięcej notuje PiS (4,2%), najmniej KP (3,6%).

W przypadku wypowiedzi internautów porównując indeks Jaccarda można zauważyć, że wspomnienia partii KP (których jest stosunkowo niedużo porównując ze wspomnieniami innych partii) odbiegają od reszty (*Jaccard index* waha się między 0,28 – 0,31). Najwyższą wartością podobieństwa cechuje się PiS oraz Lewica, gdzie parametr podobieństwa jest na poziomie 0,63. Słów o nacechowaniu mocno negatywnym najwięcej notuje wspomnienie partii PiS (4,8%) oraz Lewica (4,8%), najmniej KP (2,4%). Słów o nacechowaniu mocno pozytywnym najwięcej notuje Lewica oraz KP (4,3%), najmniej PiS (3,8%).

Stanowi to sygnał że każda z grup ma swoje charakterystyki, że modele predykcyjne będą w stanie rozróżnić wypowiedzi poszczególnych grup politycznych na zadowalającym poziomie. Ze względu na to, że wypowiedzi internautów nie mają jasnego charakteru popierania konkretnej partii do dalszej części (budowy modeli predykcyjnych) posłużono się bazą wypowiedzi polityków.

Rozdział 4. Część Praktyczna – Modele predykcyjne

Po analizie danych i dokładnym sprawdzeniu poszczególnych grup politycznych przystąpiono do stworzenia struktur oraz ustawień do modeli predykcyjnych – jest to kolejny etap analizy części własnej badań.

W pierwszym kroku przeanalizowana została aktualna literatura związana z tym etapem badań. Jej wynik prezentowany jest na początku niniejszego rozdziału, aby wyraźnie wskazać wynikanie dobranych metod w badaniach własnych.

Wśród dostępnych opracowań warto wymienić pracę *Textual Analysis of Stock Market Prediction Using Financial News Articles*⁴⁷⁹, gdzie autorzy (R. Schumaker, H. Chen) przedstawiają rolę artykułów z wiadomościami finansowymi w trzech różnych reprezentacjach tekstowych. Używając algorytmów jak maszyny wektorów nośnych (SVM), pokazują, że model miał statystycznie istotny wpływ na przewidywanie przyszłych cen akcji w porównaniu z regresją liniową. Inne algorytmy używane w pracach to *Naive Bayesian* oraz algorytm genetyczny.

Inną pracę, którą warto przytoczyć to praca badaczy F. Zablith, I. H. Osman⁴⁸⁰ – z wykorzystaniem uczenia maszynowego stworzyli oni modele oceny produktów i usług. Zaproponowali platformę o nazwie *ReviewModus*, platformę eksploracji i przetwarzania tekstu, która, (1) opiera się na strukturze modelu i odpowiadających jej pytaniach oceniających, aby wytrenować algorytm uczenia maszynowego w celu przewidywania klasyfikacji recenzji; (2) przewiduje nastroje w ramach przeglądów na podstawie zestawów danych trenujących dotyczących przeglądów zewnętrznych; oraz (3) przekształca

⁴⁷⁹ R. Schumaker, H. Chen, *Textual Analysis...*, dz. cyt.

⁴⁸⁰ F. Zablith, I. H. Osman, *ReviewModus: Text ...*, dz. cyt., s. 569-583.

wyodrębnione miary z przeglądów do dalszej analizy. Badanie pokazuje obiecujące wyniki klasyfikacji wynikiem *F-score* oraz wysoką korelacją predykcji sentymentów.

Kolejnym tytułem, który warto przytoczyć jest ten należący do autorów H. Mueller, C. Rauh⁴⁸¹. Autorzy pokazują nową metodologię przewidywania konfliktów zbrojnych za pomocą tekstu w gazetach. Dzięki uczeniu maszynowemu ogromne ilości tekstu w gazetach zostają zredukowane do tematów możliwych do zinterpretowania. Tematy te następnie są wykorzystywane w *regresjach panelowych* w celu przewidzenia początku konfliktu. Autorzy proponują wykorzystanie tego rozwiązania szczególnie do przewidywania konfliktu w krajach, w których wcześniej jakiś konflikt wystąpił.

Pracą która jest tematycznie związana z zagadnieniami politycznymi jest praca J. Schwarz⁴⁸². W swojej pracy autorka omawia wyniki trzech różnych podejść do przewidywania ideologii politycznej w danych komentarzy na *YouTube*. Pierwszy klasyfikator został stworzony z 500 komentarzy platformy *YouTube*, drugi z postów na *Twitterze* kandydata na prezydenta, a ostatni z 4000 zdań z manifestów wyborczych partii amerykańskich. Wyniki wskazują, że we wszystkich rozwiązaniach ideologię polityczną można przewidzieć z ponad przypadkową dokładnością. Kolejną pracą związaną z przewidywaniami przynależności do grup politycznych jest praca badaczy M. Iyyer, P. Enns, J. Boyd-Graber, P. Resnik⁴⁸³. Autorzy zastosowali rekurencyjną sieć neuronową (*RNN*) do określenia stanowiska politycznego wyrażonego w zdaniach. Obydwie prace cechują klasyfikację binarną (poglądy demokratyczne/republikańskie) i zawierają wypowiedzi tylko w języku angielskim.

Niniejszy rozdział jest próbą stworzenia modelu predykcyjnego (1) na bazie języka polskiego, który nie ma w wielu przypadkach gotowych bibliotek do przetwarzania języka naturalnego oraz (2) stworzenia modeli posiadanych wiele klas w tym przypadku 5 – jako nazw partii (PiS, KO, Lewica, KP, Konfederacja). Model z określoną dokładnością powinien przewidywać przynależność partyjną na bazie wyuczonych wzorców z innych wypowiedzi polityków. Tworzenie modeli będzie poprzedzone oczyszczeniem

⁴⁸¹ H. Mueller, C. Rauh, *Reading Between the Lines: Prediction of Political Violence Using Newspaper Text*, Cambridge University Press, 2017.

⁴⁸² J. Schwarz, *Detecting political ...*, dz. cyt.

⁴⁸³ M. Iyyer, P. Enns, J. Boyd-Graber, P. Resnik, *Political Ideology...*, dz. cyt; *Proceedings of...* dz. cyt., s. 1113–1122.

danych opisanym w rozdziale (4.1), następnie podzielenie danych na trenujące (które posłużą do budowy modeli) oraz testowe (na których będziemy oceniać model). Podział na poszczególne grupy wymagał będzie wyrównania liczności klas – opisanego w rozdziale (4.1). Następnie każdy stworzony model będzie testowany na danych testowych parametrami: *Accuracy*, *Kappa*, *Mcnemar's Test P-Value* – również opisanych w rozdziale (4.1).

4.1. Przygotowanie danych

Dane do analizy zostały przygotowane w języku *R*, jak w przypadku rozdziału poprzedniego. Wykonane zostały przekształcenia występujące w analizie tekstu, jak: ujednolicenie do małych liter, usuwanie spójników, nawiasów i innych symboli⁴⁸⁴. Usunięto imię i nazwisko posła, tytuł wypowiedzi, wszelkie wypowiedzi marszałków oraz wypowiedzi innych posłów z sali.

Następnie usunięto z wypowiedzi spójniki oraz słowa mające marginalne znaczenie takie jak: *który, ze, że, także, również, więc, jeśli, jeżeli, może, którzy, nad, tam, ich, dla tego, aby, w, nie, czym, jak, tej, przez, po, bo, dla, za, po, są, ma, tego, które, od, te, żeby, tylko, ta, i, to, na, z, jest, że, się, o, do, a, ale, co, tym, czy, r, tych, ten, już, też, zł, która, oraz*.

Usunięto również słowa będące ogólnym przywitanie z marszałkiem, posłami bądź wysoką izbą. Należą do nich: *szanowni, dziękuję, pani, panie, marszałek, marszałku, ministrze, szanowna, szanowny, wysoka, izbo, panowie, posłowie, premierze*.

Stworzono dwa zbiory danych: trenujący zawierający 80% obserwacji, na którym budowano modele i 20%-owy zbiór testujący oceniający ich wynik.

Ze względu na różnicę w liczności klas wartości istniała potrzeba wyrównania do podobnych poziomów liczności danych trenujących, aby model miał szansę nauczyć się wzorców szczególnie, jeśli chodzi o klasy o mniejszej liczności. Jest kilka metod, które warto było rozważyć:

⁴⁸⁴ Symboli takich jak: przecinek, cudzysłów, znak procenta, cyfry

- *Upsampling* – ta metoda zwiększa rozmiar klasy mniejszościowej poprzez próbkowanie z zamianą, tak aby klasy miały ten sam rozmiar.
- Próbkowanie w dół (*ang. Downsampling*) – odwrotnie do metody *Upsampling*, zmniejsza wielkość klasy większości tak, aby była taka sama lub bliższa wielkości klasy mniejszości poprzez dobór losowej próby.
- Metody hybrydowe – dobrze znane metody hybrydowe to ROSE (*ang. Random oversampling examples*) i SMOTE (*ang. Synthetic minority oversampling technique*), które polegają na próbkowaniu w dół klasy większości i tworzeniu nowych sztucznych punktów w klasie mniejszości.⁴⁸⁵

Ze względu na stosunkowo niedużą liczbę przypadków zdecydowano się na *Upsampling* tylko dla zbioru trenującego.

Tabela 33. Zestawienie licznosci przypadków w klasach w zbiorze treningowym przed *Upsamplingiem*

KO	Konfederacja	KP	Lewica	PiS
3701	869	1009	2004	2254

Tabela 34. Zestawienie licznosci przypadków w klasach po *Upsamplingu*

KO	Konfederacja	KP	Lewica	PiS
3701	3701	3701	3701	3701

Do przygotowania danych do modelu użyto biblioteki *tm* w której przygotowano korpus słów. Obliczono występowanie 97% wszystkich słów biorąc pod uwagę ich częstotści⁴⁸⁶.

Modele oceniono na danych testowych (20%) wszystkich danych metodami z wykorzystaniem tablicy pomyłek (*ang. Confusion matrix*) oraz współczynnikami:

⁴⁸⁵ Methods for dealing with imbalanced data, www.r-bloggers.com/2019/04/methods-for-dealing-with-imbalanced-data, 04.01.2022.

⁴⁸⁶ W przypadku ostatniego modelu RNN użyto innego podejścia omówionego w rozdziale z wynikami dla tego algorytmu.

Rysunek 160. Tablica pomylek (ang. *Confusion matrix*)

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Źródło: *Understanding Confusion Matrix*, <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>, 10.01.2022

Accuracy ($TP+TN/Total$) – Dokładność to jedna z miar oceny modeli klasyfikacji. Dokładność to ułamek przewidywań, w których nasz model się nie mylił, gdzie TP = prawdziwie pozytywne przypadki, TN = prawdziwie negatywne przypadki, FP = fałszywie pozytywne przypadki, a FN = fałszywie negatywne przypadki⁴⁸⁷.

Sensitivity (Recall) ($TP/(TP+FN)$) – Odpowiada na pytanie: Jaka część rzeczywistych przypadków pozytywnych została prawidłowo zidentyfikowana?⁴⁸⁸

Specificity ($TN/(TN+FP)$) – Specyficzność wybieramy, jeśli chcemy objąć wszystkie prawdziwie negatywne przypadki (TN), co oznacza, że redukujemy fałszywe alarmy klasyfikowania (FP)⁴⁸⁹.

Cohen's kappa (Kappa) – statystyka ta jest miernikiem wydajności modeli, który bardzo dobrze radzi sobie zarówno z problemami wieloklasowymi (w przypadku pracy jest to model pięcioklasowy gdzie każda klasa jest nazwą partii: PiS, KO, KP, Lewica, Konfederacja), jak i nie zrównoważonymi klasami (gdzie liczność klas jest niezbalansowana:

⁴⁸⁷ Classification accuracy, <https://developers.google.com/machine-learning/crash-course/classification/accuracy>, 02.06.2022.

⁴⁸⁸ Precision and recall, <https://developers.google.com/machine-learning/crash-course/classification/precision-and-recall>, 02.06.2022.

⁴⁸⁹ Accuracy Recall Precision f-score, <https://towardsdatascience.com/accuracy-recall-precision-f-score-specificity-which-to-optimize-on-867d3f11124>, 20.05.2022.

w pracy autor wyrównał w danych trenujących licznosc przypadków poprzez technikę *upsamlingu* i *downsamlingu* opisaną wcześniej)⁴⁹⁰.

$$\kappa = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e},$$

gdzie p_o jest obserwowanym wynikiem, a p_e oczekiwanym wynikiem. Mówi, o ile lepszy jest klasyfikator w porównaniu z wydajnością klasyfikatora, który zwyczajnie zgaduje losowo zgodnie z licznoscia każdej klasy. *Kappa Cohena* jest zawsze mniejsza lub równa 1. Wartości 0 lub mniejsze wskazują, że klasyfikator jest bezużyteczny. Nie ma znormalizowanego sposobu interpretacji jego wartości. Według schematu Landisa oraz Kocha⁴⁹¹ wartość < 0 oznacza brak zgodności, 0–0,20 jako nieznaczną, 0,21–0,40 jako dostateczną, 0,41–0,60 jako umiarkowaną, 0,61–0,80 jako dużą, a 0,81–1 jako prawie doskonałą zgodność.⁴⁹²

McNemar's Test P-Value – jest nieparametrycznym testem hipotezy statystycznej.

$p > \alpha$. Brak odrzucenia H_0 (np. leczenie nie przyniosło efektu).

$p \leq \alpha$. Odrzucenie H_0 , istotna różnica (np. leczenie miało wpływ).⁴⁹³

Wykorzystano bibliotekę *caret*⁴⁹⁴ do oceny modeli.

4.2. Model z wykorzystaniem drzew decyzyjnych – R

Techniki drzew decyzyjnych są szeroko stosowane do budowania modeli klasyfikacyjnych. Modele te bardzo przypominają ludzkie rozumowanie i są łatwe do interpretacji. Stosowane w przypadku klasyfikacji, jak i regresji. Drzewa decyzyjne są modelami sekwencyjnymi, które logicznie łączą sekwencję prostych testów. Każdy test porównuje atrybut liczbowy z wartością progową lub atrybut nominalny zestaw możliwych wartości.

⁴⁹⁰ Cohens Kappa in spss statistics, <https://statistics.laerd.com/spss-tutorials/cohens-kappa-in-spss-statistics.php>, 02.04.2022.

⁴⁹¹ J. R. Landis, G.G. Koch, *The measurement of observer agreement for categorical data*, „Biometrics”, 1977, nr 33 (1), s 159–174.

⁴⁹² Performance measures Cohens Kappa statistic, <https://thedata scientist.com/performance-measures-cohens-kappa-statistic>, 10.01.2022.

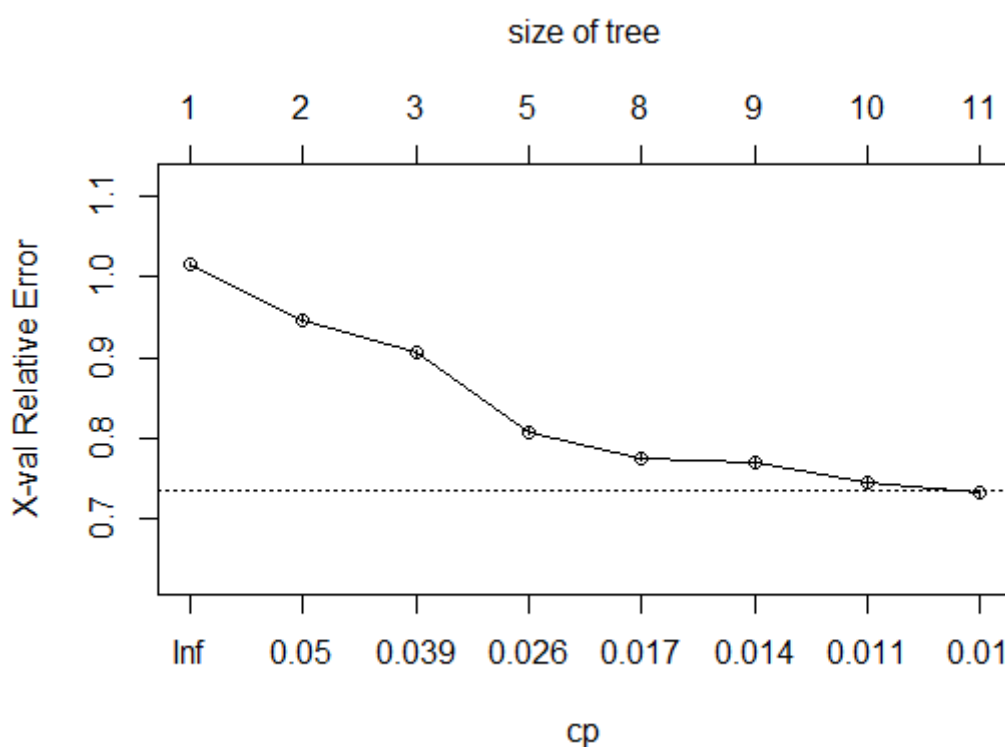
⁴⁹³ McNemars test for machine learning, <https://machinelearningmastery.com/mcnemars-test-for-machine-learning>, 10.01.2022.

⁴⁹⁴ Biblioteka caret, <https://www.rdocumentation.org/packages/caret/versions/6.0-93>, 02.04.2022.

Takie klasyfikatory mają przewagę nad modelami *czarnoskrzynkowymi*, takimi jak sieci neuronowe, pod względem zrozumiałości.⁴⁹⁵

Do tworzenia modelu wykorzystano bibliotekę *rpart* a do wizualizacji bibliotekę *rpart.plot*⁴⁹⁶. Stworzono drzewo, a następnie zoptymalizowano model o parametr *cp* (*ang. complexity parameter*) odpowiadający za przycięcie drzewa, a co za tym idzie zoptymalizowanie modelu.

Rysunek 161. Ustawienie *cp* w funkcji *rpart()* – przycięcie drzewa

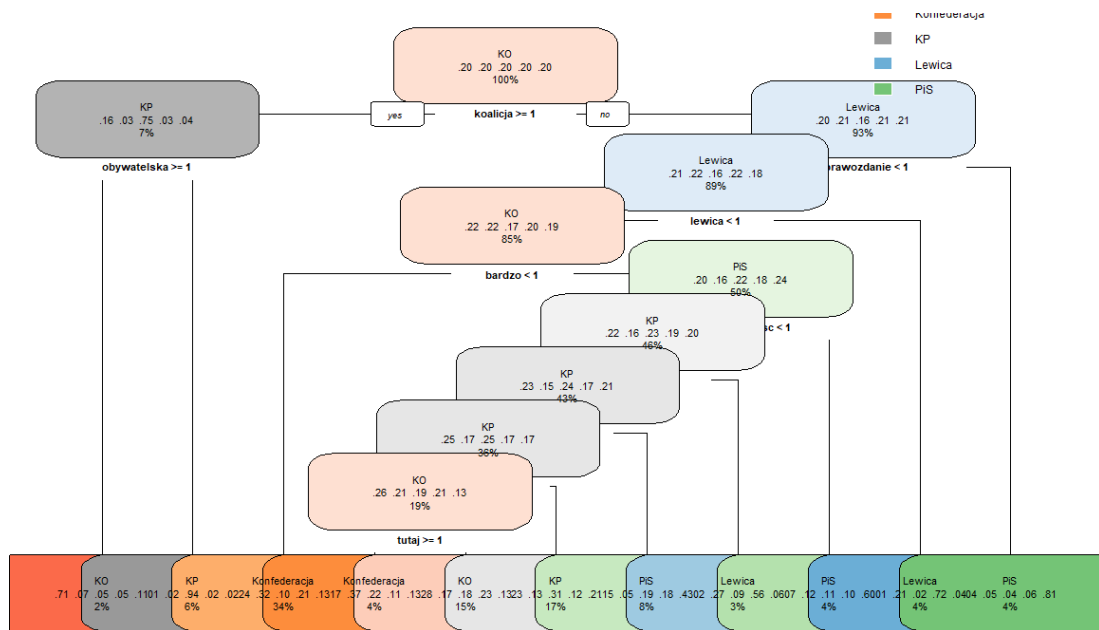


Źródło: opracowanie własne z wykorzystaniem biblioteki *rpart*.

⁴⁹⁵ S. B. Kotsiantis, *Decision trees: a recent overview*, „Artificial Intelligence Review”, 2013, nr 39, s. 261-283.

⁴⁹⁶ Biblioteka *rpart*, <https://www.rdocumentation.org/packages/rpart/versions/4.1-15/topics/rpart>, 12.01.2021; Biblioteka *rpart.plot*, <https://www.rdocumentation.org/packages/rpart.plot/versions/3.1.1/topics/rpart.plot> 12.01.2021.

Rysunek 162. Drzewo decyzyjne



Źródło: opracowanie własne z wykorzystaniem biblioteki rpart.plot.

Tabela predykcji pokazuje, że zoptymalizowany model miał duże problemy, jeśli chodzi o właściwą predykcję. Dobrze sklasyfikowanych przypadków jest 44% wypowiedzi. Model jest istotny statystycznie. Współczynnik Kappa wynosi 0,28, co wg schematu oceny Landisa oraz Kocho klasyfikuje model jako dostateczny.

Tabela 35. Tablica predykcji algorytmu drzew decyzyjnych oraz wyniki wydajności modelu

predictDTree					
	KO	Konfederacja	KP	Lewica	PiS
KO	381	140	102	193	109
Konfederacja	55	74	22	41	25
KP	60	20	109	38	25
Lewica	118	73	37	224	49
PiS	100	53	51	57	303

Overall Statistics

Accuracy : 0.4437
95% CI : (0.4239, 0.4636)
No Information Rate : 0.3762
P-Value [Acc > NIR] : 4.482e-12

Kappa : 0.2784

Mcnemar's Test P-Value : 5.295e-16


```
Statistics by Class:
Class: KO
Sensitivity: 0.4119
Specificity: 0.7829
Balanced Accuracy: 0.5974

Class: Konfederacja
Sensitivity: 0.34101
Specificity: 0.87244
Balanced Accuracy: 0.60672

Class: KP
Sensitivity: 0.43254
Specificity: 0.90394
Balanced Accuracy: 0.66824

Class: Lewica
Sensitivity: 0.44711
Specificity: 0.83197
Balanced Accuracy: 0.63954

Class: PiS
Sensitivity: 0.5372
Specificity: 0.8902
Balanced Accuracy: 0.7137
```

Źródło: opracowanie własne z wykorzystaniem biblioteki caret.

4.3. Model z wykorzystaniem metody najbliższego sąsiada kNN – R

Algorytm k-najbliższych sąsiadów (*kNN*) jest prostym, łatwym do wdrożenia algorytmem nadzorowanego uczenia maszynowego, którego można użyć do rozwiązywania problemów zarówno z regresją jak i klasyfikacją. Algorytm *kNN* zakłada, że istnieją podobne rzeczy blisko siebie mierzone różnymi sposobami miar odległości.⁴⁹⁷

Odległość między dwiema próbkami (q, p) można ją określić wykorzystując przykładowo miarę odległości euklidesowej.

$$d(p, q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}.$$

⁴⁹⁷ Machine learning basics with the k-nearest neighbors algorithm, <https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761> , 10.01.2022.

Ważną jest parametr k , który decyduje o tym, ilu sąsiadów zostanie wybranych do algorytmu kNN . Właściwy wybór k ma istotny wpływ na wydajność diagnostyczną algorytmu kNN . Duże k zmniejsza wpływ wariancji spowodowanej błędem losowym, ale wiąże się z ryzykiem zignorowania wzorca⁴⁹⁸. Rozwiązaniem do wybrania odpowiedniej wartości k jest zachowanie równowagi między zbytnim i ogólnym dopasowaniem⁴⁹⁹. Niektórzy autorzy sugerują, aby k było równe pierwiastkowi kwadratowemu liczby obserwacji w uczącym zbiorze danych⁵⁰⁰.

Do stworzenia modelu posłużono się biblioteką *class*. Stworzono model z parametrem k na poziomie równym 1 z tego względu, że każde zwiększenie tego parametru wiązało się ze słabszym wynikiem poprawnych klasyfikacji. Model osiągnął 36% poprawnych klasyfikacji. Tak jak w poprzednim algorytmie, model jest istotny statystycznie. Współczynnik Kappa wynosi 0,15 co wg schematu oceny Landisa oraz Kocha klasyfikuje model jako ten o nieznacznej zgodności.

Tabela 36. Tablica predykcji algorytmu kNN oraz wyniki wydajności modelu

kNNModel					
	KO	Konfederacja	KP	Lewica	PiS
KO	341	133	46	162	243
Konfederacja	60	57	15	31	54
KP	73	24	48	30	77
Lewica	139	74	19	103	166
PiS	96	47	38	54	329

Overall Statistics	
Accuracy :	0.3571
95% CI :	(0.3381, 0.3764)
No Information Rate :	0.3762
P-Value [Acc > NIR] :	0.9763
Kappa :	0.1541
Mcnemar's Test P-Value :	<2e-16

Statistics by Class:

⁴⁹⁸ Z. Zhang, *Introduction...*, dz., cyt.

⁴⁹⁹ Z. Zhang, *Too much covariates in a multivariable model may cause the problem of overfitting*, J Thorac Dis, 2014 Sep, nr 6(9), s. 196-197.

⁵⁰⁰ B. Lantz, *Machine learning with R. 2nd ed.* Birmingham, Packt Publishing, 2015.

```
Class: KO
Sensitivity: 0.3686
Specificity: 0.7601
Balanced Accuracy: 0.5644

Class: Konfederacja
Sensitivity: 0.26267
Specificity: 0.87600
Balanced Accuracy: 0.56934

Class: KP
Sensitivity: 0.19048
Specificity: 0.94653
Balanced Accuracy: 0.56850

Class: Lewica
Sensitivity: 0.20559
Specificity: 0.85853
Balanced Accuracy: 0.53206

Class: PiS
Sensitivity: 0.5833
Specificity: 0.7150
Balanced Accuracy: 0.6492
```

Źródło: opracowanie własne z wykorzystaniem biblioteki caret.

4.4. Model z wykorzystaniem lasów losowych – R

Lasy losowe (*ang. Random forest- RF*) to metoda klasyfikacji i regresji oparta na agregacji dużej liczby drzew decyzyjnych. Jest to zespół drzew skonstruowanych na podstawie zestawu danych uczących. Jest kilka wariantów RF, które charakteryzują: 1) *sposób zbudowania każdego drzewa z osobna*, 2) *procedura użyta do wygenerowania zmodyfikowanego zbioru danych, na podstawie których budowane jest każde pojedyncze drzewo*, 3) *sposób, w jaki prognozy każdego pojedynczego drzewa są agregowane w celu uzyskania predykcji*⁵⁰¹.

W oryginalnej metodzie RF zaproponowanej przez Leo Breimana⁵⁰², każde drzewo jest standardowym drzewem klasyfikacyjnym lub regresyjnym (CART), które wykorzystuje tzw. zmniejszenie entropii Gini (DGI) jako kryterium podziału i wybiera dzielenie

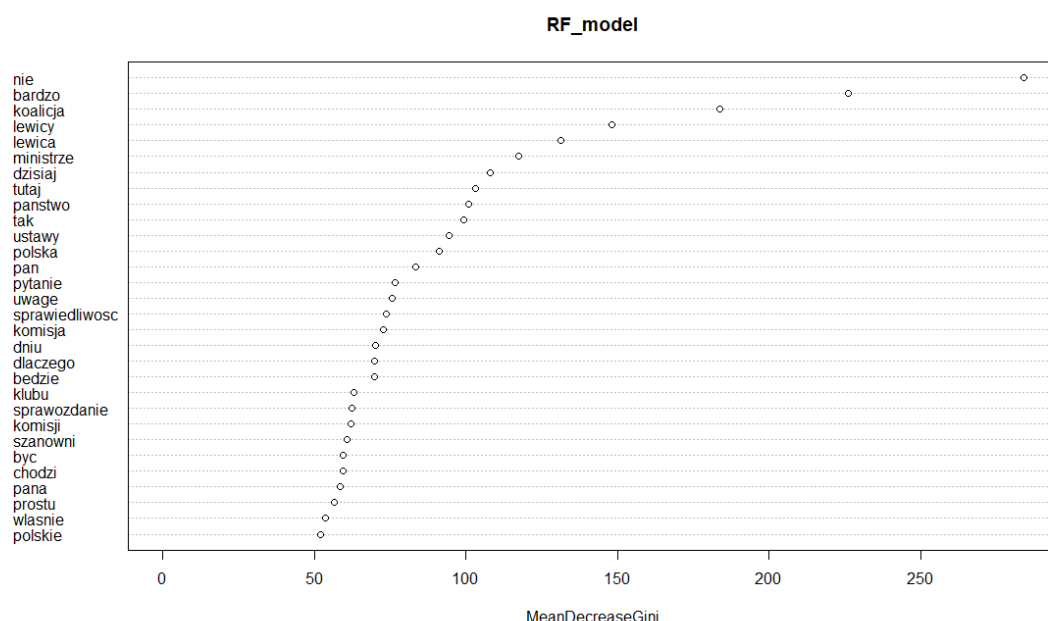
⁵⁰¹ A. L. Boulesteix, S. Janitza, J. Kruppa, I. R. König, *Overview of Random Forest Methodology and Practical Guidance with Emphasis on Computational Biology and Bioinformatics*, „Department of Statistics University of Munich Technical Report”, 2012, nr 129, s. 3.

⁵⁰² L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, *Classification and regression trees*, Chapman & Hall, 1984.

predyktora z losowo wybranego podzbioru predyktorów (podzbiór jest inny w każdym podziale). Każde drzewo jest zbudowane z próbki stworzonej z zamianą z oryginalnego zbioru danych, a przewidywania wszystkie drzewa są ostatecznie agregowane w drodze głosowania większościowego.⁵⁰³

Do stworzenia modelu posłużyła biblioteka *randomForest*⁵⁰⁴ z ustawieniem 500 pojedynczych drzew. Wykorzystano funkcję, która odpowiada za sprawdzenie ważności zmiennych wejściowych (w tym przypadku słów). Do najważniejszych słów, które miały wpływ na podział należą: *nie*, *bardzo*, *koalicja* oraz *lewica* i *lewicy*.

Rysunek 163. Ważność zmiennych po stworzeniu modelu z wykorzystaniem algorytmu *RandomForest*



Źródło: opracowanie własne z wykorzystaniem biblioteki *randomForest*.

Wyniki predykcji wskazują na obiecującą wydajność tego modelu. Model osiągnął 64% poprawnych klasyfikacji. Tak, jak w poprzednich algorytmach, model jest istotny

⁵⁰³ A. L. Boulesteix, S. Janitza, J. Kruppa, I. R. König, *Overview of ...*, dz. cyt., s. 3.

⁵⁰⁴ Biblioteka *randomforest*, <https://www.rdocumentation.org/packages/randomForest/versions/4.7-1.1/topics/randomForest>, 20.02.2021.

statystycznie. Współczynnik Kappa wynosi 0,49, co wg schematu oceny Landisa oraz Kocho klasyfikuje go jako model o umiarkowanej zgodności.

Tabela 37. Tablica predykcji algorytmu *RandomForest* oraz wyniki wydajności modelu.

predictRF					
	KO	Konfederacja	KP	Lewica	PiS
KO	810	15	5	38	57
Konfederacja	95	95	3	18	6
KP	127	3	85	15	22
Lewica	245	7	3	214	32
PiS	154	17	2	17	374

Overall Statistics

Accuracy : 0.6417
 95% CI : (0.6224, 0.6607)
 No Information Rate : 0.3762
 P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.4891

Mcnemar's Test P-Value : < 2.2e-16

Statistics by Class:

Class: KO
 Sensitivity: 0.8757
 Specificity: 0.5952
 Balanced Accuracy: 0.7354

Class: Konfederacja
 Sensitivity: 0.43779
 Specificity: 0.98127
 Balanced Accuracy: 0.70953

Class: KP
 Sensitivity: 0.33730
 Specificity: 0.99411
 Balanced Accuracy: 0.66571

Class: Lewica
 Sensitivity: 0.42715
 Specificity: 0.95506
 Balanced Accuracy: 0.69110

Class: PiS
 Sensitivity: 0.6631
 Specificity: 0.9383
 Balanced Accuracy: 0.8007

Źródło: opracowanie własne z wykorzystaniem biblioteki caret.

4.5. Model z wykorzystaniem algorytmu wzmacniającego GBM Boosting – R

Gradient Boosting Machine (GBM) to algorytm nadzorowanego uczenia się, który jest szeroko stosowany w wielu dziedzinach. *GBM* rutynowo występuje jako wiodący algorytm w konkursach uczenia maszynowego, takich jak Kaggle i KDDCup.⁵⁰⁵

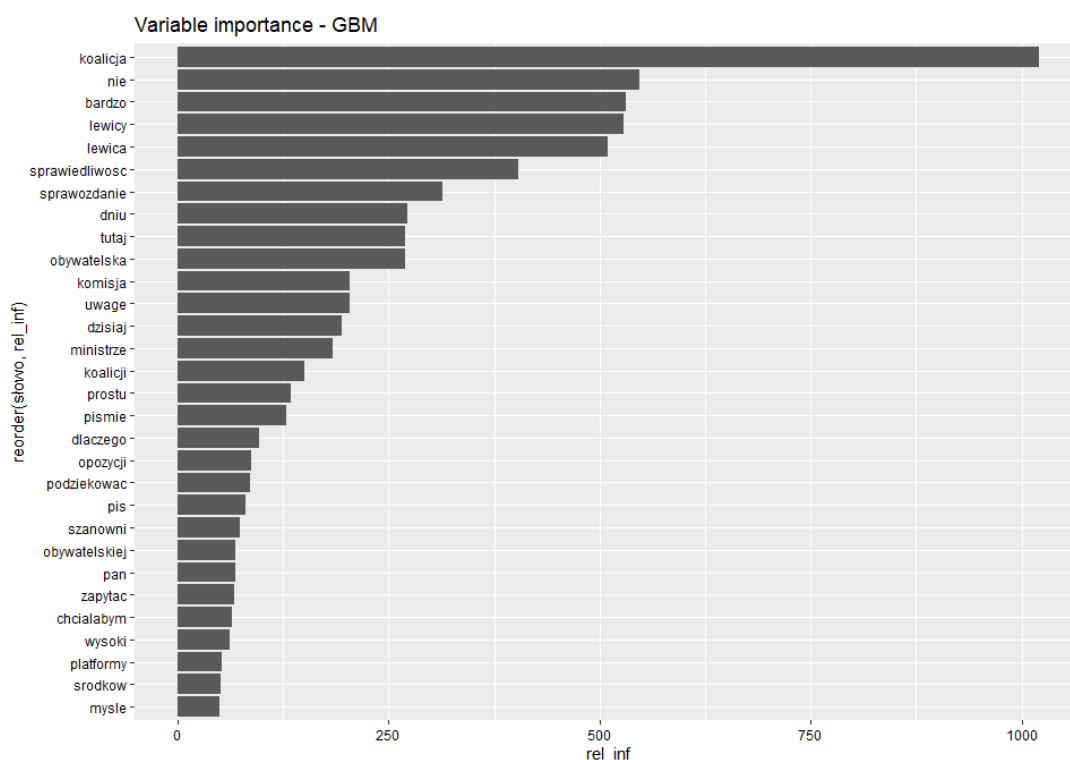
W *GBM*, procedura uczenia się dopasowuje kolejno nowe modele, aby zapewnić dokładniejsze oszacowanie zmiennej, którą przewidujemy. Główną ideą tego algorytmu jest skonstruowanie nowych elementów uczących się tak, aby były maksymalnie skorelowane z ujemnym gradientem funkcji straty. Zastosowane funkcje strat mogą być dowolne, ale dla lepszej intuicji, jeśli funkcja błędu jest klasyczną stratą błędu do kwadratu, procedura uczenia spowoduje dopasowanie błędów. Generalnie wybór funkcji straty należy do badacza, zarówno przy bogatym zróżnicowaniu funkcji straty, jak i możliwości realizacji własnej funkcji straty.⁵⁰⁶

Do budowy modelu wykorzystano bibliotekę *gbm*. Wykorzystano również podobnie, jak w przypadku algorytmu RF funkcję, która odpowiada za sprawdzenie ważności zmiennych wejściowych (słów). Do najważniejszych słów, które miały wpływ na podział należą: *koalicja*, *nie*, *bardzo* oraz *lewicy* i *lewica*. Ważność słów ma podobny zakres co w algorytmie *RandomForest* – inna jest tylko ich kolejność.

⁵⁰⁵ H. Lu, S. P. Karimireddy, N. Ponomareva, V. Mirrokni, *Accelerating Gradient ...*dz.cyt., s.1.

⁵⁰⁶ A. Natekin, A. Knoll, *Gradient boosting machines, a tutorial*, Front Neurorobot, 2013, s. 7-21.

Rysunek 164. Ważność zmiennych po stworzeniu modelu z wykorzystaniem algorytmu *gbm*.



Źródło: opracowanie własne z wykorzystaniem biblioteki *gbm*.

Wyniki predykcji wskazują na zadowalającą wydajność tego modelu. Model osiągnął 53% poprawnych klasyfikacji. Tak, jak w poprzednich algorytmach, model jest istotny statystycznie. Współczynnik Kappa wynosi 0,40, co wg schematu oceny Landisa oraz Kocho klasyfikuje go jako model o umiarkowanej zgodności.

Tabela 38. Tablica predykcji algorytmu *gbm* oraz wyniki wydajności modelu

predictGBMclass					
	KO	Konfederacja	KP	Lewica	PiS
KO	461	151	83	108	122
Konfederacja	27	141	8	27	14
KP	47	30	119	25	31
Lewica	99	79	21	236	66
PiS	60	57	53	39	355

Overall Statistics

Accuracy : 0.5336
 95% CI : (0.5136, 0.5534)
 No Information Rate : 0.3762
 P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.3979

McNemar's Test P-Value : < 2.2e-16

Statistics by Class:

Class: KO

Sensitivity: 0.4984

Specificity: 0.8481

Balanced Accuracy: 0.6732

Class: Konfederacja

Sensitivity: 0.64977

Specificity: 0.85861

Balanced Accuracy: 0.75419

Class: KP

Sensitivity: 0.47222

Specificity: 0.92524

Balanced Accuracy: 0.69873

Class: Lewica

Sensitivity: 0.47106

Specificity: 0.89837

Balanced Accuracy: 0.68471

Class: PiS

Sensitivity: 0.6294

Specificity: 0.8770

Balanced Accuracy: 0.7532

Źródło: opracowanie własne z wykorzystaniem biblioteki caret.

4.6. Model z wykorzystaniem maszyn wektorów nośnych (SVM) – R

Maszyny wektorów nośnych (SVM) to zestaw nadzorowanych metod uczenia używanych do klasyfikacji, regresji i wykrywania wartości odstających. Wdrożony został przez firmę *AT&T Bell Laboratories* przez Vladimira Vapnika⁵⁰⁷.

SVM są jedną z najsolidniejszych metod przewidywania, opartą na statystycznych ramach uczenia się. W porównaniu z innymi metodami uczenia maszynowego, SVM jest bardzo wydajny w rozpoznawaniu subtelnych wzorców w złożonych zestawach danych i może być używany do rozpoznawania pisma ręcznego, rozpoznawania fałszywych kart kredytowych, identyfikacji mówcy, a także wykrywania twarzy.⁵⁰⁸

⁵⁰⁷ C. Cortes, V. Vapnik, *Support-vector...*, dz. cyt.

⁵⁰⁸ S. Huang, N. Cai, P. Penzuti, S. Narandes, Y. Wang, X. Wayne, *Applications of Support Vector Machine (SVM) Learning in Cancer Genomics*, *Cancer Genomics Proteomics*, 2018, nr 15(1), s.1.

Do budowy modelu na bazie algorytmu SVM wykorzystano bibliotekę *e1071*⁵⁰⁹.

Tabela 39. Ustawienia modelu SMV

```
Call:
svm(formula = partia ~ ., data = trainSparse2, type = "C-classi-
fication", kernel = "radial", cost = 1)

Parameters:
  SVM-Type:  C-classification
  SVM-kernel: radial
    cost:    1

Number of Support Vectors: 14864

( 3488 2349 2853 3395 2779 )

Number of Classes: 5

Levels:
KO Konfederacja KP Lewica PiS
```

Źródło: opracowanie własne.

Wyniki predykcji wskazują na bardzo dobrą wydajność tego modelu. Model osiągnął 61% poprawnych klasyfikacji. Tak, jak w poprzednich algorytmach, model jest istotny statystycznie. Współczynnik Kappa wynosi 0,47, co wg schematu oceny Landisa oraz Kocho klasyfikuje go jako model o umiarkowanej zgodności.

Tabela 40. Tablica predykcji algorytmu SVM oraz wyniki wydajności modelu

predictSVM					
	KO Konfederacja	KP	Lewica	PiS	
KO	592	51	31	166	85
Konfederacja	62	116	4	20	15
KP	68	7	108	27	42
Lewica	156	27	20	256	42
PiS	65	23	19	39	418

Overall statistics

Accuracy : 0.6059
 95% CI : (0.5863, 0.6253)
 No Information Rate : 0.3762
 P-Value [Acc > NIR] : < 2.2e-16

⁵⁰⁹ Biblioteka e1071, <https://cran.r-project.org/web/packages/e1071/index.html>, 20.02.2021.

Kappa : 0.4688

McNemar's Test P-Value : 0.0005341

Statistics by Class:

Class: KO

Sensitivity: 0.6400

Specificity: 0.7712

Balanced Accuracy: 0.7056

Class: Konfederacja

Sensitivity: 0.53456

Specificity: 0.95183

Balanced Accuracy: 0.74320

Class: KP

Sensitivity: 0.42857

Specificity: 0.96647

Balanced Accuracy: 0.69752

Class: Lewica

Sensitivity: 0.5110

Specificity: 0.8713

Balanced Accuracy: 0.6911

Class: PiS

Sensitivity: 0.7411

Specificity: 0.9029

Balanced Accuracy: 0.8220

Źródło: opracowanie własne z wykorzystaniem biblioteki caret.

4.7. Model z wykorzystaniem algorytmu sieci neuronowych – Python

Do stworzenia modelu z wykorzystaniem rekurencyjnych sieci neuronowych użyto języka *Python* oraz biblioteki *Tensorflow*. *TensorFlow* jest kompleksową platformą na licencji *open source* do uczenia maszynowego. Posiada wszechstronny ekosystem bibliotek, narzędzi i zasobów, który pozwala naukowcom na wprowadzanie najnowocześniejszych rozwiązań w obszarze uczenia maszynowego (*ML*), a programistom tworzenie i wdrażanie aplikacji opartych na *ML*⁵¹⁰.

⁵¹⁰ Tensorflow, <https://www.tensorflow.org>, 09.01.2022.

W tym podejściu zastosowano rozwiązanie oparte o *Downsampling* czyli zredukowanie wszystkich klas w losowy sposób do klasy najmniejszej.

Wyniki przed *downsamplingiem*:

partia	
KO	4626
KP	1261
Konfederacja	1086
Lewica	2505
Pis	2818

Wyniki po *downsamplingu*:

partia	
KO	1086
KP	1086
Konfederacja	1086
Lewica	1086
Pis	1086

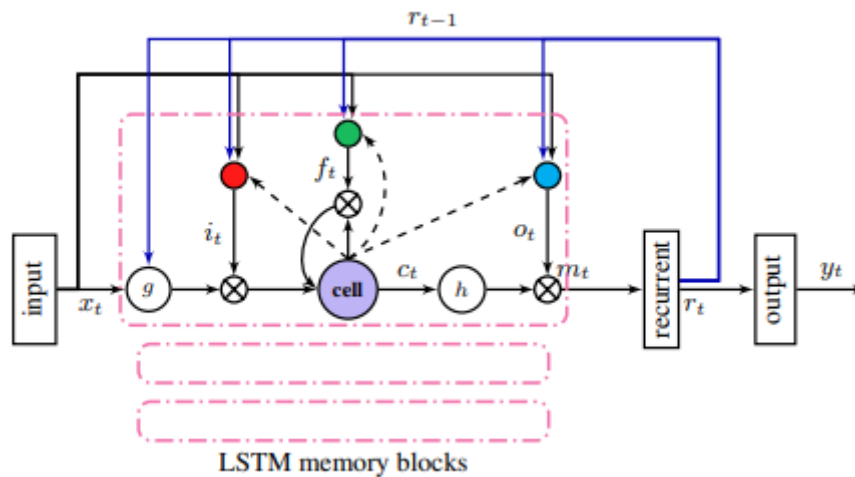
Rekurencyjne sieci neuronowe (*ang. Recurrent Neural Network – RNN*) są rodzajem sztucznych sieci neuronowych (*ang. Artificial Neural Networks*), które mają cykliczne połączenia, dzięki czemu są wydajne w modelowaniu sekwencji. Zostały wykorzystane do etykietowania sekwencji i zadań przewidywania sekwencji, takie jak rozpoznawanie pisma ręcznego, modelowanie języka oraz znakowania fonetycznego.⁵¹¹

LSTM (*ang. Long short-term memory*), które są elementem *RNN* zawierają specjalne jednostki zwane blokami pamięci w powtarzalnej warstwie ukrytej (*ang. Recurrent Hidden Layer*). Bloki pamięci zawierają pamięć komórki z własnymi połączeniami, przechowujące czasowy stan sieci, oprócz specjalnych jednostek multiplikatywnych zwanych bramkami do kontroli przepływu informacji. Każdy blok pamięci w oryginalnej architekturze zawiera bramkę wejściową i bramkę wyjściową.⁵¹²

⁵¹¹ H. Sak, A. Senior, F. Beaufays, *Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling*, Google, USA, 2014.

⁵¹² H. Sak, A. Senior, F. Beaufays, *Long Short-Term ...* dz. cyt., s. 2.

Rysunek 165. Architektura RNN LSTM.



Źródło: H. Sak, A. Senior, F. Beaufays, *Long Short-Term ...*, dz. cyt., s. 2.

Bramka wejściowa steruje przepływem aktywacji wejścia do komórki pamięci. Bramka wyjściowa steruje przepływem wyjściowym komórki aktywacji do reszty sieci. Później brama odpowiadająca za zapominanie (*ang. Forget Gate*) została dodana do bloku pamięci⁵¹³. To rozwiązało problem słabości modeli LSTM, uniemożliwiając im przetwarzanie ciągłych strumieni wejściowych, które nie są podzielone na podsekwencje. Bramka zapomnienia skaluje stan wewnętrzny komórki przed dodaniem jako wejście do komórki przez samopowtarzalne połączenie komórki, w ten sposób adaptacyjnie zapominając lub resetując komórkę pamięci⁵¹⁴.

Stworzony model na bazie RNN LSTM z wykorzystaniem języka *Python* i biblioteki *Tensorflow* przyjmuje kilka warstw m. in. warstwę LSTM.

Tabela 41. Opis konfiguracji modelu z wykorzystaniem Tensorflow

```
Shape of data tensor: (5430, 1500)
Shape of label tensor: (5430, 5)
```

```
Model: "sequential"
```

⁵¹³ F. A. Gers, J. Schmidhuber, F. Cummins, *Learning to forget: Continual prediction with LSTM*, *Neural Computation*, 2000, nr 12(10), s. 2451–2471.

⁵¹⁴ H. Sak, A. Senior, F. Beaufays, *Long Short-Term ...* dz. cyt., s.2.

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 1500, 100)	999900
spatial_dropout1d (SpatialDr	(None, 1500, 100)	0
lstm (LSTM)	(None, 90)	68760
dense (Dense)	(None, 5)	455
Total params: 1,069,115		
Trainable params: 1,069,115		
Non-trainable params: 0		

Źródło: opracowanie własne z wykorzystaniem biblioteki tensorflow⁵¹⁵.

Wyniki predykcji wskazują na bardzo dobrą/znakomitą wydajność tego modelu. Model osiągnął 67% poprawnych klasyfikacji. Tak jak w poprzednich algorytmach model jest istotny statystycznie. Współczynnik Kappa wynosi 0.58 – co jest najlepszym wynikiem spośród uzyskanych przez wszystkie modele – wg schematu oceny Landisa oraz Kocho klasyfikuje go jako model o umiarkowanej zgodności dochodząc do wartości bliskiej uzyskania zgodności dużej.

Tabela 42. Tablica predykcji algorytmu RNN oraz wyniki wydajności modelu

	Predicted				
	KO	KP	Konfederacja	Lewica	Pis
KO	119	21	14	60	16
KP	41	147	4	21	12
Konfederacja	20	5	165	12	4
Lewica	33	13	8	139	7
Pis	21	16	6	27	155
Accuracy: 0.668					
Kappa : 0.5847					
McNemar's Test P-Value : 0.01582					

Źródło: opracowanie własne z wykorzystaniem biblioteki keras/tensorflow.

⁵¹⁵ Biblioteka tensorflow, <https://pypi.org/project/tensorflow>, 01.05.2022.

4.8. Porównanie modeli

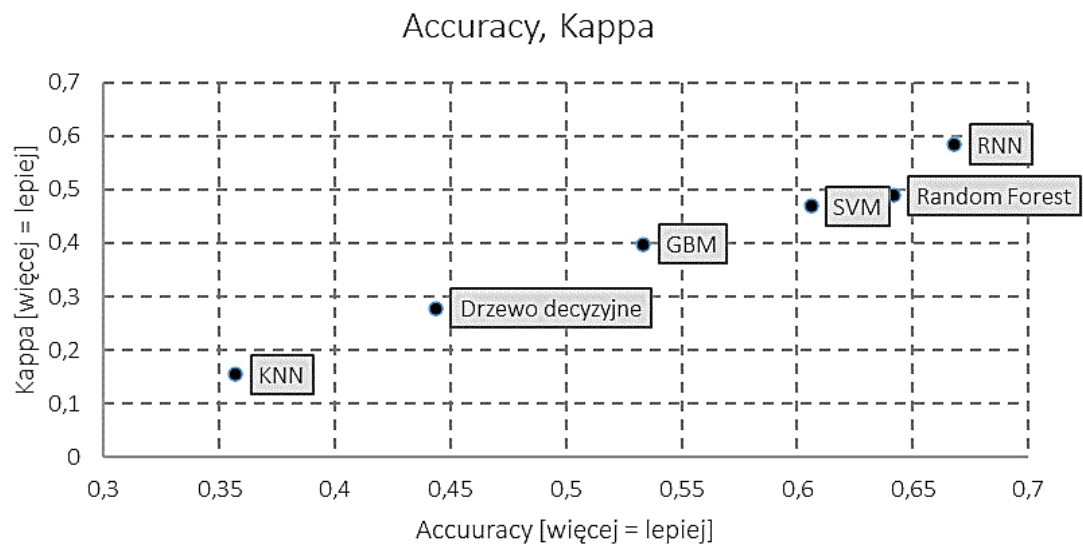
Wyniki w tabeli poniżej prezentują wydajność modeli na danych testowych omawianymi wcześniej parametrami *Accuracy* i *Kappa* wraz z poziomem istotności. Wyniki wskazują, że najlepszą wydajność prezentują modele: *RNN LSTM*, *Random Forest* i *SVM*. Pozostałe modele prezentują wydajność na wyższym poziomie niż losowe przypisywanie prognozy, ale znacznie odbiegają od wspomnianych trzech modeli.

Tabela 43. Porównanie wyników predykcji modeli

Algorytm	Drzewo decyzyjne	KNN	Random Forest	GBM	SVM	RNN
Accuracy	0.4437	0.3571	0.6417	0.5336	0.6059	0.668
Kappa	0.2784	0.1541	0.4891	0.3979	0.4688	0.5847
Mcnemar's Test P-Value	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05

Źródło: Opracowanie własne.

Rysunek 166. Porównanie wyników predykcji modeli



Źródło: Opracowanie własne

4.9. Testowanie na kilku przypadkach

Po sprawdzeniu klasyfikacji całego zbioru testowego w niniejszej części przedstawiono działanie narzędzia. W celu unaocznienia działania algorytmów przeprowadzono na niewielkiej próbie (wypowiedzi pięciu posłów) z różnych partii z różnymi długościami wypowiedzi. Sprawdzone predykcję do partii względem ich wypowiedzi wszystkimi wytrenowanymi modelami. Najkrótsza wypowiedź posła KO Borysa Budki ma po oczyszczeniu jedynie 100 wyrazów, najdłuższa wypowiedź posła Lewicy Włodzimierza Czarzastego ponad 2000 słów. Każda z wypowiedzi została sklasyfikowana wartościami w przedziale 0-1 do każdej z partii osobno. Jedyny wyjątek stanowi algorytm *kNN* który ocenia najbliższy przypadek w postaci przyporządkowania do partii w przestrzeni użytych słów i przewiduje najbliższą partię w tej przestrzeni.

Tadeusz Cymański: PiS (oryginalna wypowiedź: 1175 wyrazów – oczyszczona wypowiedź: 822 wyrazów)

Wypowiedź oryginalna:

Cieężko mi mówić, pani marszałek, drodzy państwo, koleżanki i koledzy, bo rzeczywiście mam perspektywę 22 lat, bo tu kiedyś przyszedłem. Niestety to nie jest czas wylotu polskiego parlamentu, przykro mi to powiedzieć. Nie jestem lepszy od was, to jasne, ale nie można w taki sposób... Wszystko jest nagrywane bardziej niż kiedyś. Teraz jest Internet, coraz trudniej kłamać. Tak nie można. Poseł RP musi rozróżniać złotówkę od tysiąca, od miliona i od miliarda. Ja to mówię, bo mi łatwiej, bo byłem w banku. Nie można takiej ściemy tutaj robić i nie rozróżniać pewnych rzeczy. Trzeba miarę zachowywać. Dziękuję tym kolegom, którzy wyczaili, mówiąc po studencku, i rzeczywiście dotknęli sprawy. Uśmiechają się teraz do mnie. Tak, jest problem wykorzystania funduszy, nie tylko tego funduszu, ale także Funduszu Pracy, gwarantowanych świadczeń. Chodzi o praktykę od wielu, wielu lat. Marzę sobie, żeby dać odpór jednemu z koleśków, kolegów, bo mówią do nas tutaj: Co wy wyprawiacie? A oni? Proszę bardzo: 2010 r. - 7,5 mld, 2011 r. - 4,2, 2012 r. - 7,5. Razem - 16,8, czyli tyle, ile było w 2012 r. Kto to robił?

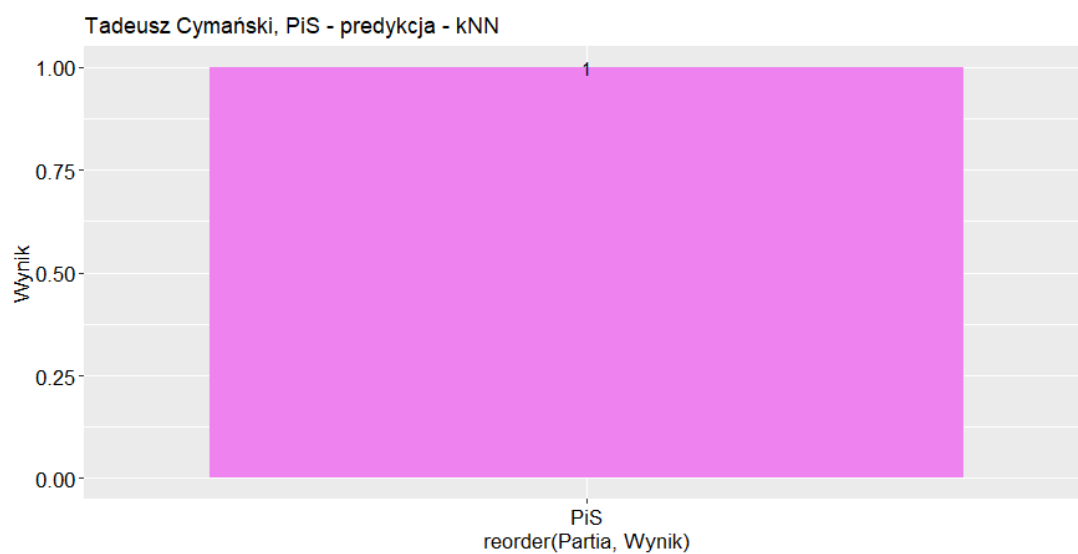
Krasnoludki robiły? A kto pytał? Przemysław Wipler. Odpowiedzi ... ciąg dalszy wypowiedzi

Wypowiedź oczyszczona:

ciężko mi mówić drodzy państwo koleżanki koledzy rzeczywiście mam perspektywę lat tu kiedyś przyszedłem niestety nie czas wlotu polskiego parlamentu przykro mi powiedzieć nie jestem lepszy od was jasne nie można taki sposób wszystko nagrywane bardziej niż kiedyś teraz internet coraz trudniej kłamać tak nie można poseł rp musi rozróżniać złotówkę od tysiąca od miliona od miliarda ja mówię mi łatwiej byłem banku nie można takiej ściemy tutaj robić nie rozróżniać pewnych rzeczy trzeba miarę zachowywać kolegom wyczaili mówiąc studencku rzeczywiście dotknęli sprawy uśmiechają teraz mnie tak problem wykorzystania funduszy nie tylko funduszu także funduszu pracy gwarantowanych świadczeń chodzi praktykę od wielu wielu lat marzę sobie żeby dać odpór jednemu koleśków kolegów mówią nas tutaj wy wyprawiacie oni proszę bardzo mld razem czyli tyle ile było kto robił krasnoludki robiły kto pytał przemysław wipler odpowiedzi nie uzyskał praktyki naszych poprzedników my nie jesteśmy święci naprawdę zanim kogoś uderzysz spójrz siebie nazywa moralne prawo krytyki my też mamy problemy też czasami idziemy skróty tu problemy wykorzystywanie niepełnosprawnych proszę państwa mam wewnętrzne przekonanie powiem teraz coś siłą owszem pieniądze budżecie różnych źródeł największe wiadomo vat akcyza ... ciąg dalszy wypowiedzi

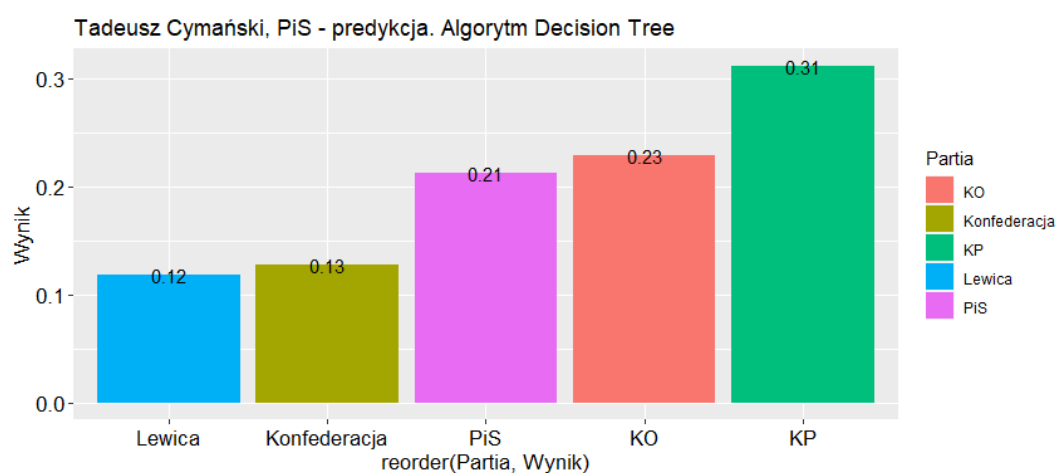
Cztery modele przewidują wyraźnie klasyfikację do partii PiS (*k*NN, SVM, Random Forest i RNN), model drzew klasyfikacyjnych przewiduje partię KP, lecz klasyfikacja nie jest tak jednoznaczna (0,31 dla KP, 0,23 KO, 0,21 PiS). Model na bazie GBM wskazuje wyraźnie klasyfikację na partię Konfederacja.

Rysunek 167. Predykcja modelem z użyciem algorytmu kNN – wypowiedź 1.



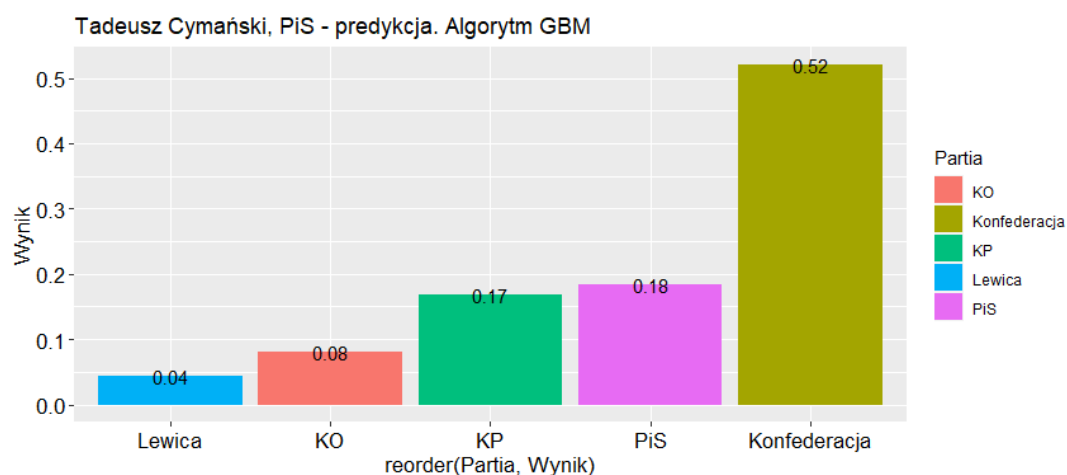
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 168. Predykcja modelem z użyciem algorytmu Drzewa decyzyjne – wypowiedź 1.



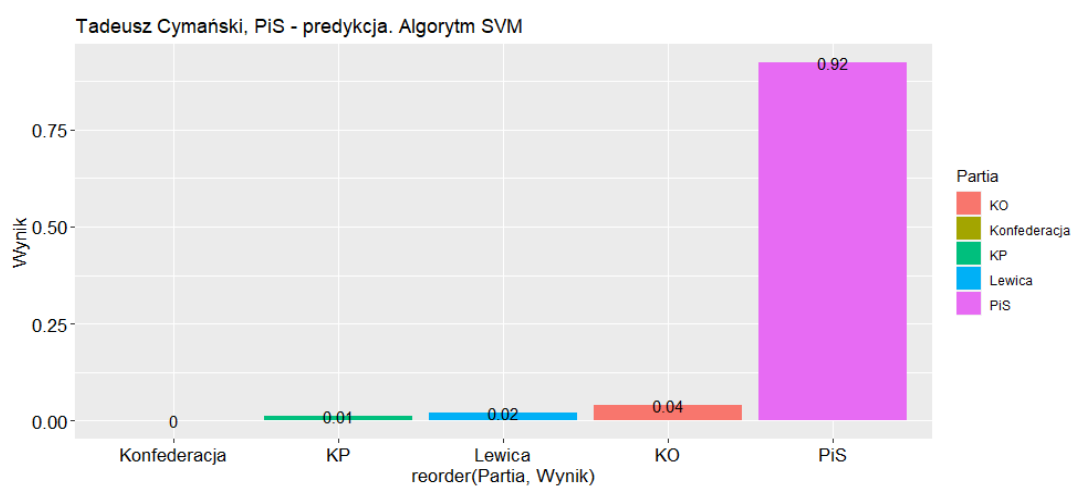
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 169. Predykcja modelem z użyciem algorytmu GBM – wypowiedź 1.



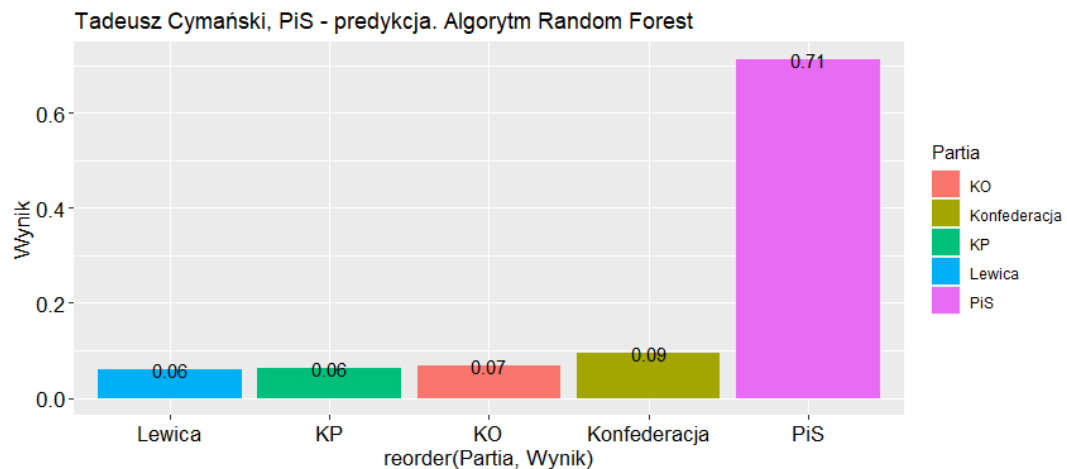
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 170. Predykcja modelem z użyciem algorytmu SVM – wypowiedź 1.



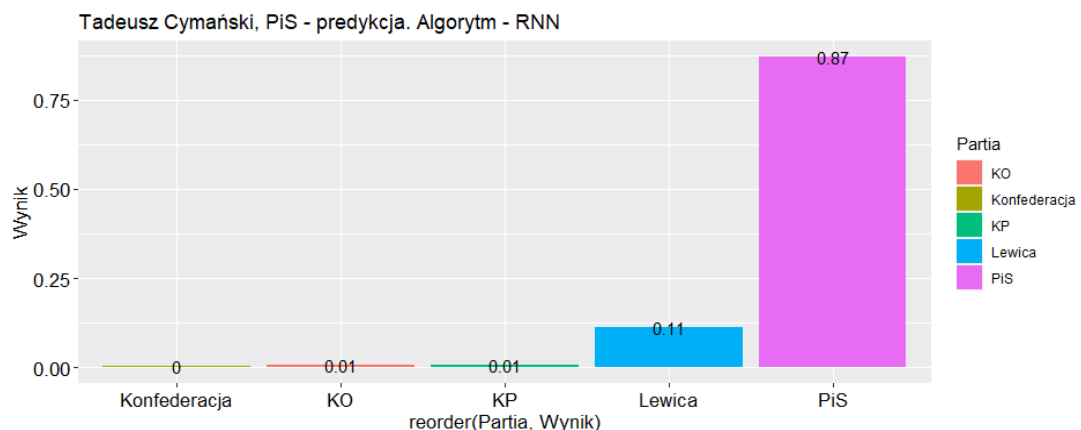
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 171. Predykcja modelem z użyciem algorytmu Random Forest – wypowiedź 1.



Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 172. Predykcja modelem z użyciem algorytmu RNN – wypowiedź 1.



Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Grzegorz Braun (Konfederacja) (oryginalna wypowiedź: 290 wyrazów – oczyszczona wypowiedź: 214 wyrazów)

Wypowiedź oryginalna:

Dziękuję, panie marszałku.

Szczęść Boże raz jeszcze.

Wykorzystam te krótkie minuty, żeby wezwać pana marszałka do tego, by zdyscyplinować służby sejmowe, które czynią przeszkody osobom zaproszonym na posiedzenie dwóch połączonych zespołów parlamentarnych. Posiedzenie to trwa już od

kilkudziesięciu minut, ale osoby z listy przedstawionej wcześniej - pomijając już problematyczność i arbitralność tych procedur - i zaakceptowanej przez Kancelarię Sejmu nie zostały dopuszczone. Chodzi o posiedzenie połączonych zespołów parlamentarnych: do spraw bezpieczeństwa szczepień dzieci i dorosłych i do spraw nadużyć i naruszeń prawa w związku z COVID-19. W tej chwili w trakcie tego posiedzenia trwa jeszcze wypowiedź pana doktora, ministra Zbigniewa Hałata, który jako szef sanepidu przy rządach w poprzednich kadencjach, jeszcze w latach 90., był w tej sprawie praktykiem i człowiekiem odpowiedzialnym... ciąg dalszy wypowiedzi

Wypowiedź oczyszczona:

szczęść boże raz jeszcze wykorzystam te krótkie minuty żeby wezwać pana marszałka by zdyscyplinować służby sejmowe czynią przeszkody osobom zaproszonym posiedzenie dwóch połączonych zespołów parlamentarnych posiedzenie trwa już od kilkudziesięciu minut osoby listy przedstawionej wcześniej pomijając już problematyczność arbitralność tych procedur zaakceptowanej przez kancelarię sejmu nie zostały dopuszczone chodzi posiedzenie połączonych zespołów parlamentarnych spraw bezpieczeństwa szczepień dzieci dorosłych spraw nadużyć naruszeń prawa związku covid chwili trakcie posiedzenia trwa jeszcze wypowiedź pana doktora ministra zbigniewa hałata który jako szef sanepidu przy rządach poprzednich kadencjach jeszcze latach był sprawie praktykiem człowiekiem odpowiedzialnym epidemiolog doświadczony również badaniach działaniach dziedzinie granicą mówi wszystko tu dokonuje pod pretekstem walki pandemią ni mniej ni więcej tylko eksperymentem medycznym... ciąg dalszy wypowiedzi

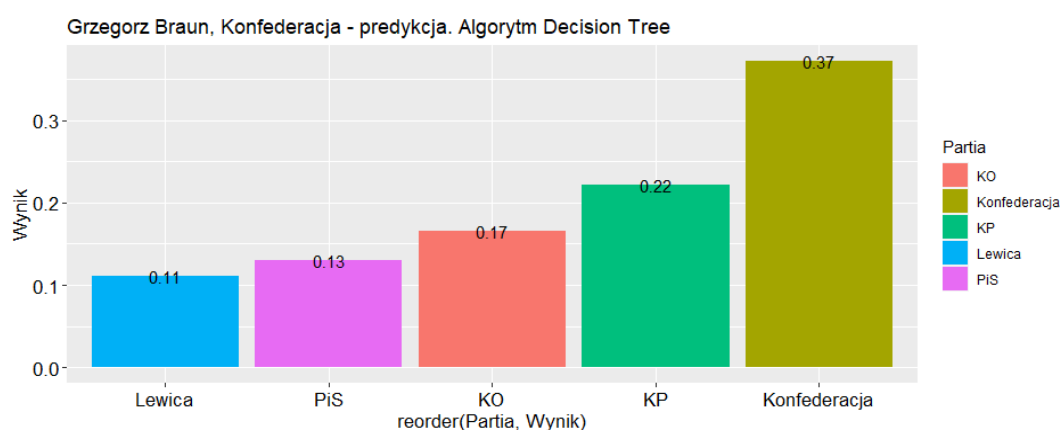
Wypowiedź posła Grzegorza Brauna z wykorzystaniem wszystkich modeli została sklasyfikowana poprawnie. Poseł stosuje charakterystyczne zwroty przywitania *szczęść boże* stąd prawdopodobnie w większości modeli stuprocentowa poprawność klasyfikacji.

Rysunek 173. Predykcja modelem z użyciem algorytmu kNN – wypowiedź 2.



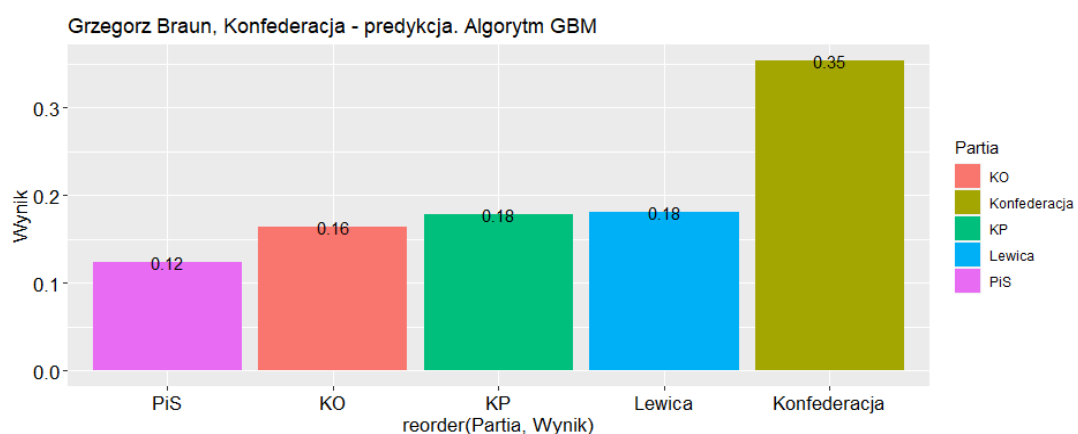
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 174. Predykcja modelem z użyciem algorytmu Drzewa decyzyjne – wypowiedź 2.



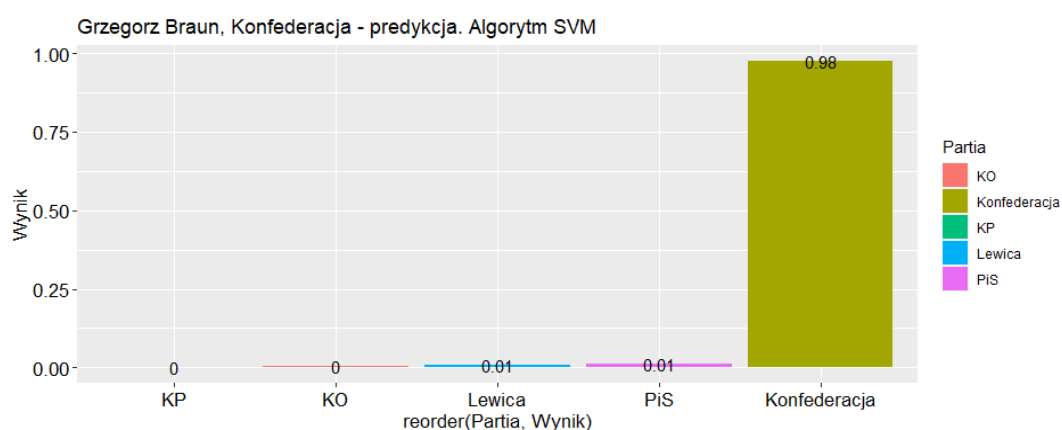
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 175. Predykcja modelem z użyciem algorytmu GBM – wypowiedź 2.



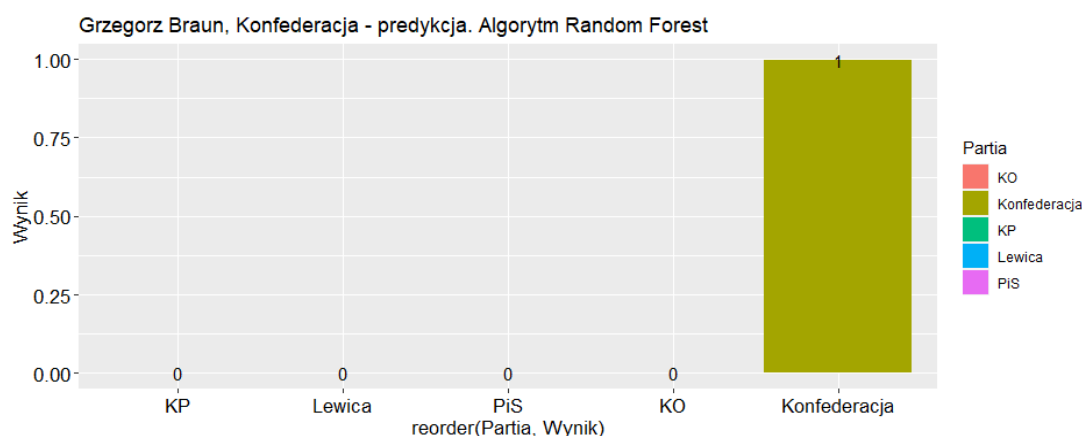
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 176. Predykcja modelem z użyciem algorytmu SVM – wypowiedź 2.



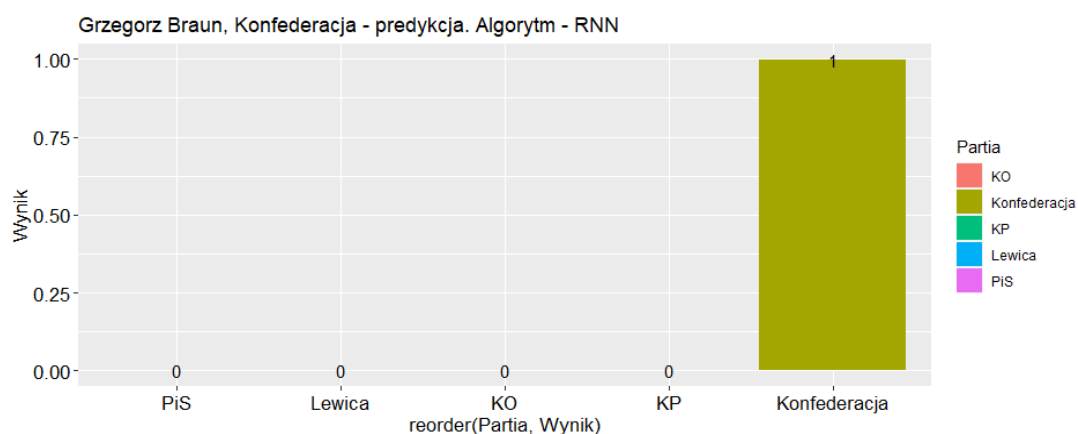
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 177. Predykcja modelem z użyciem algorytmu Random Forest – wypowiedź 2.



Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 178. Predykcja modelem z użyciem algorytmu RNN – wypowiedź 2.



Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Borys Budka (KO) (oryginalna wypowiedź: 134 wyrazy - oczyszczona wypowiedź: 100 wyrazów)

Najkrótsza wypowiedź w tej części pisma Borysa Budki została sklasyfikowana 3 razy (modele *kNN*, *SVM* i *Random Forest*) do partii KO i 3 razy jako posła KP (model Drzew decyzyjnych, *GBM* i *RNN*).

Wypowiedź oryginalna:

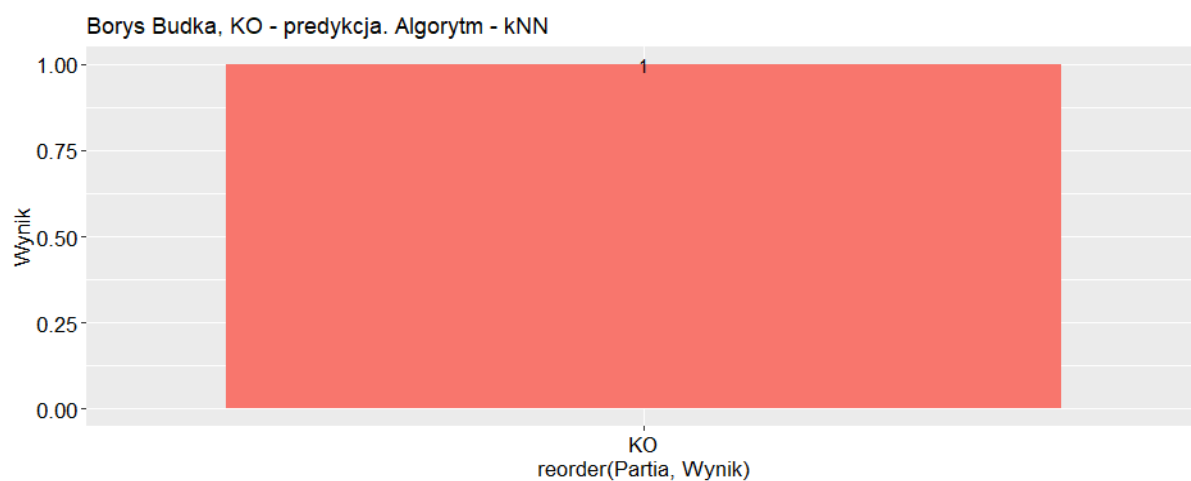
Dziękuję bardzo.

Szanowna Pani Marszałek! Wysoka Izbo! Mam wielką prośbę do pani marszałek, ponieważ klub Koalicji Obywatelskiej 3 grudnia zgłosił projekt uchwały w sprawie wyrażenia uznania dla laureatki Nagrody Nobla Olgi Tokarczuk. Rozumiem, pani marszałek, że na posiedzeniu Prezydium zapadły ustalenia, że ta uchwała będzie podjęta podczas tego posiedzenia, ale w przyszłym tygodniu. Uważamy, że cały świat już wyraził swoją radość, oddał honory naszej noblistce. Sejm Rzeczypospolitej będzie ostatnią instytucją, ale ważne, żeby to się stało. Dlatego bardzo proszę panią marszałek, żeby zobligować przewodniczącego komisji kultury do tego, by ta uchwała była przepracowana na tym posiedzeniu Sejmu... ciąg dalszy wypowiedzi

Wypowiedź oczyszczona:

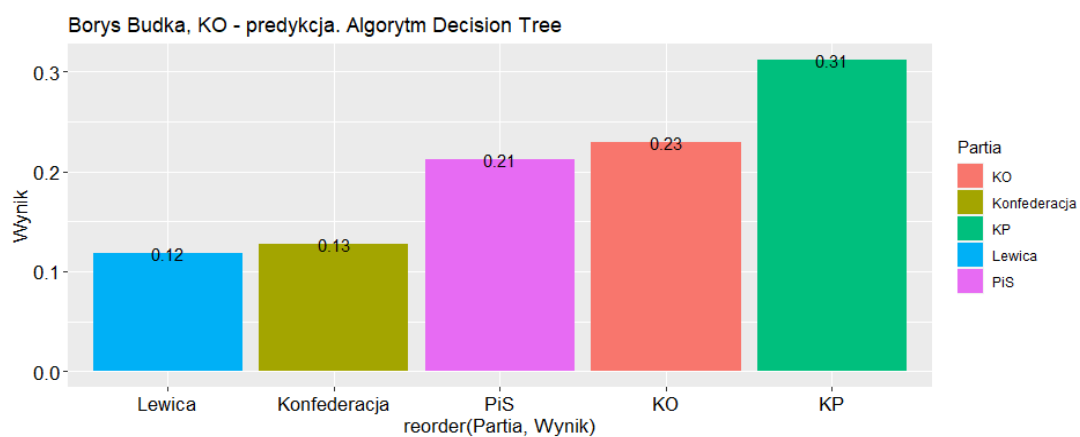
bardzo szanowna mam wielką prośbę ponieważ klub koalicji obywatelskiej grudnia zgłosił projekt uchwały sprawie wyrażenia uznania laureatki nagrody nobla olgi tokarczuk rozumiem posiedzeniu prezydium zapadły ustalenia ta uchwała będzie podjęta podczas posiedzenia przyszłym tygodniu uważamy cały świat już wyraził swoją radość oddał honory naszej noblistce sejm rzeczypospolitej będzie ostatnią instytucją ważne żeby stało dlatego bardzo proszę panią żeby zobligować przewodniczącego komisji kultury by ta uchwała była przepracowana posiedzeniu sejmu ażeby komisja kultury zebrała jeszcze dzisiaj sprawa zbyt ważna tu nie może być żadnej polityki literatura ta nagroda powinny nas wszystkich łączyć sejm rzeczypospolitej powinien podjąć tę uchwałę przez aklamację bardzo

Rysunek 179. Predykcja modelem z użyciem algorytmu kNN – wypowiedź 3.



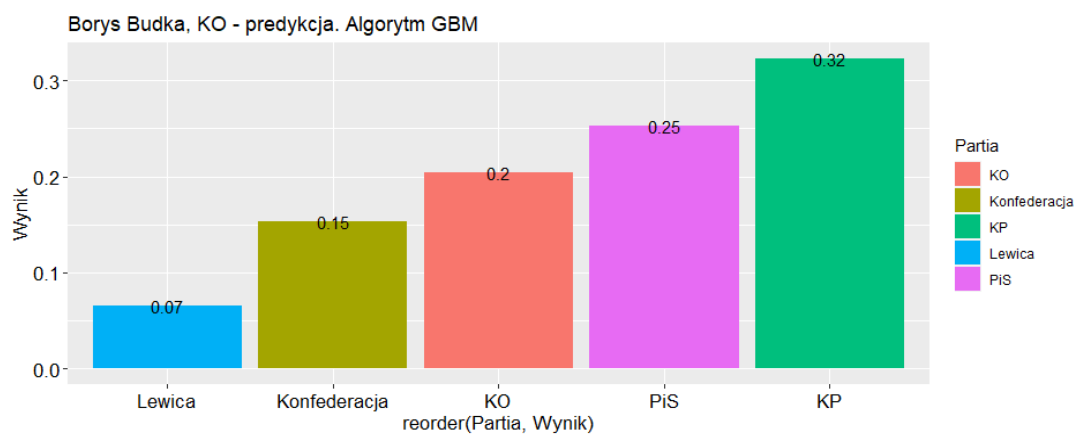
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 180. Predykcja modelem z użyciem algorytmu Drzewa decyzyjne – wypowiedź 3.



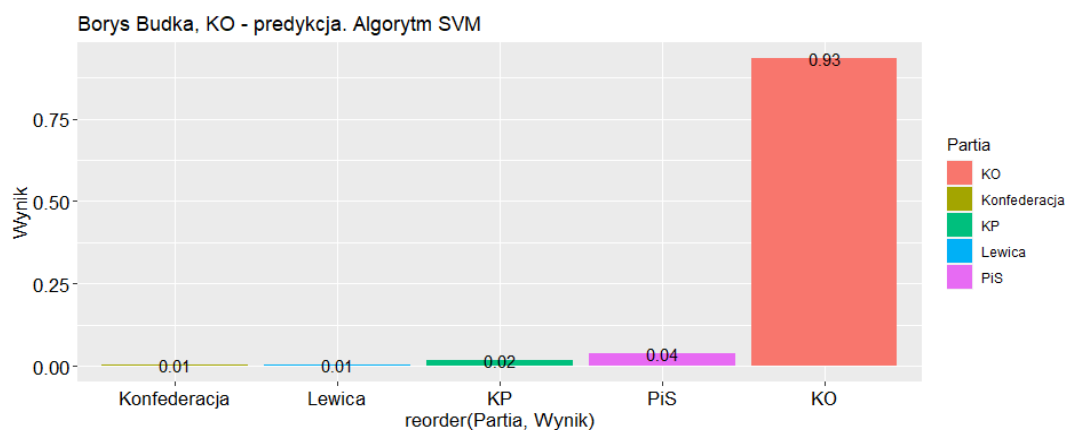
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 181. Predykcja modelem z użyciem algorytmu GBM – wypowiedź 3.



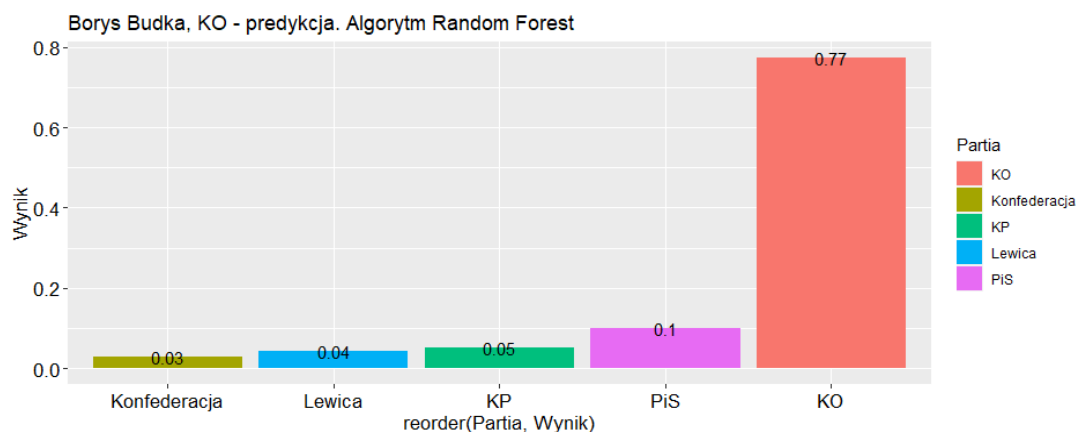
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 182. Predykcja modelem z użyciem algorytmu SVM – wypowiedź 3.



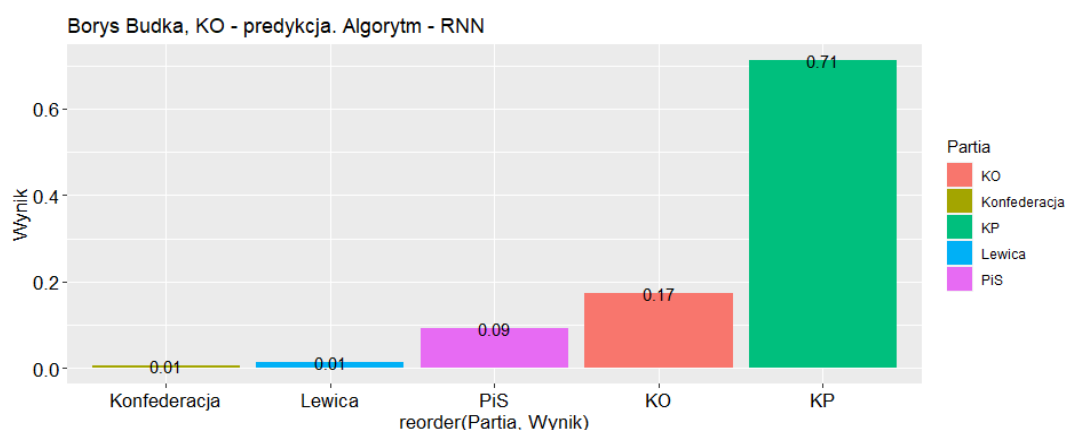
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 183. Predykcja modelem z użyciem algorytmu Random forest – wypowiedź 3.



Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 184. Predykcja modelem z użyciem algorytmu RNN – wypowiedź 3.



Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Włodzimierz Czarzasty Lewica, (oryginalna wypowiedź: 2823 wyrazów – oczyszczona wypowiedź: 2067 wyrazów)

Wypowiedź oryginalna:

Pani Marszałek! Wysoka Izbo! Panie Ministrze! Dziękuję panu za prezentację budżetu. Zresztą dokonał pan jej w sposób bardzo błyskotliwy i fantastyczny, ale pozwoli pan, że będę odnosił się do słów pana premiera, który ten budżet przedstawił nam rano, i kilkanaście razy zaczął w tej sprawie Lewicę, w związku z tym będę mówił do niego, mam nadzieję, że kiedyś tego wysłucha, bo to przecież ważna sprawa – budżet kraju.

Panie Premierze! Przedstawiając budżet, po raz kolejny mija się pan z prawdą. Przed 4 laty zasłynął pan swoją prezentacją w PowerPointcie, pana strategią na rzecz odpowiedzialnego rozwoju. Miały być stocznie, lukstorki, setki tysięcy mieszkań i samochodów elektrycznych. Mieliśmy zaraz, jeśli chodzi o dobrobyt, dogonić Niemcy. Dzisiaj wychodzi, jakie to tak naprawdę były bajdy. Bardzo niedawno, kilka tygodni temu, siedzieliśmy na tej Sali i znów słuchaliśmy festiwalu pięknych i wielkich obietnic w pana exposé, zresztą wiele było tych samych, co w pana pierwszym exposé 2 lata wcześniej. I co się z nimi stało, panie premierze? Gdzie się podziały? Bo gdy patrzę na ten budżet, widzę głównie braki i niedotrzymanie obietnic, braki środków na najbardziej palące potrzeby Polek i Polaków, a także braki odpowiedzi na największe problemy, przed którymi stoi cała polska, mimo że obiecaliście te problemy rozwiązać. Od miesięcy zarzucacie nas opowieściami o zrównoważonym budżecie, chwalicie się nim na każdym kroku. Świetnie. Zobaczmy zatem, jak wygląda wasze państwo zrównoważonego budżetu. Czy dzięki zrównoważonemu budżetowi będą rozwiązane podstawowe problemy Polaków? Co mają ludzie ze zrównoważonego budżetu? Czy jest w tym zrównoważonym budżecie ukryty deficyt? ... ciąg dalszy wypowiedzi

Wypowiedź oczyszczona:

ministrze panu prezentację budżetu zresztą dokonał pan jej sposób bardzo błyskotliwy fantastyczny pozwoli pan będę odnosił słów pana premiera który ten budżet przedstawił nam rano kilkanaście razy zaczął sprawę lewicę związku będę mówił niego mam nadzieję kiedyś wysłucha przecież ważna sprawa budżet kraju premierze przedstawiając budżet raz kolejny mija pan prawdą przed laty zasłynął pan swoją prezentacją powerpointcie pana strategią rzecz odpowiedzialnego rozwoju miały być stocznie lukstorki setki tysięcy mieszkań samochodów elektrycznych mieliśmy zaraz jeśli chodzi o dobrobyt dogonić Niemcy dzisiaj wychodzi jakie tak naprawdę były bajdy bardzo niedawno kilka tygodni temu siedzieliśmy sali znów słuchaliśmy festiwalu pięknych wielkich obietnic pana exposé zresztą wiele było tych samych pana pierwszym exposé lata wcześniej nimi stało premierze gdzie podziały gdy patrzę ten budżet widzę głównie braki niedotrzymanie obietnic braki środków najbardziej palące potrzeby Polek Polaków także braki odpowiedzi na największe problemy przed którymi stoi cała polska mimo obiecaliście te problemy rozwiązać od miesięcy zarzucacie nas opowieściami zrównoważonym budżecie chwalicie nim każdym kroku świetnie zobaczmy zatem wygląda wasze państwo zrównoważonego

budżetu dzięki zrównoważonemu budżetowi będą rozwiązane podstawowe problemy polaków mają ludzie ze zrównoważonego budżetu zrównoważonym budżecie ukryty deficyt porozmawiajmy merytorycznie konkretach fundamentalnych pól polaków ... ciąg dalszy wypowiedzi

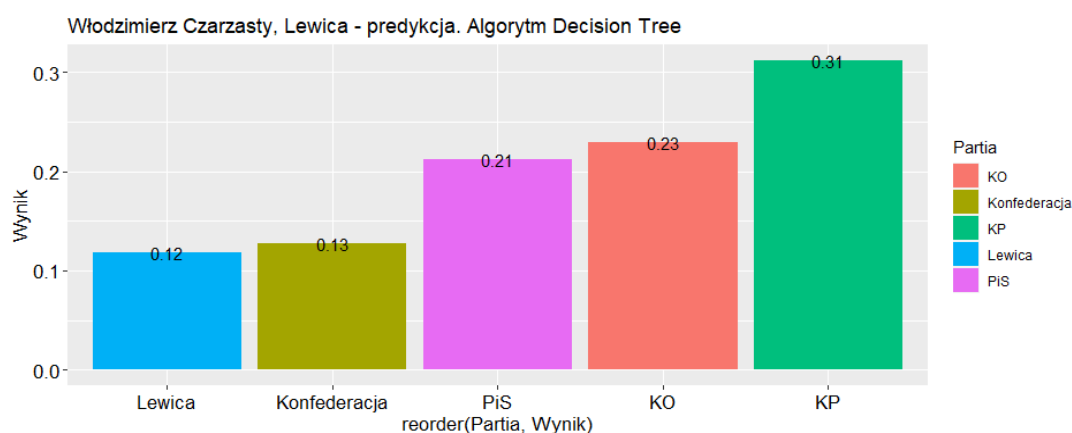
Najdłuższa wypowiedź w tej części należąca do Włodzimierza Czarzastego cztery razy została dobrze sklasyfikowana modelami (*kNN*, *GBM*, *Random forest*, *RNN*). Model na bazie algorytmu drzew decyzyjnych sklasyfikował ją jako wypowiedź KP – co ciekawe algorytm przypisał najmniejszą predykcję partii Lewica. Model na bazie SVM wskazał przynależność wypowiedzi do partii PiS.

Rysunek 185. Predykcja modelem z użyciem algorytmu kNN – wypowiedź 4.



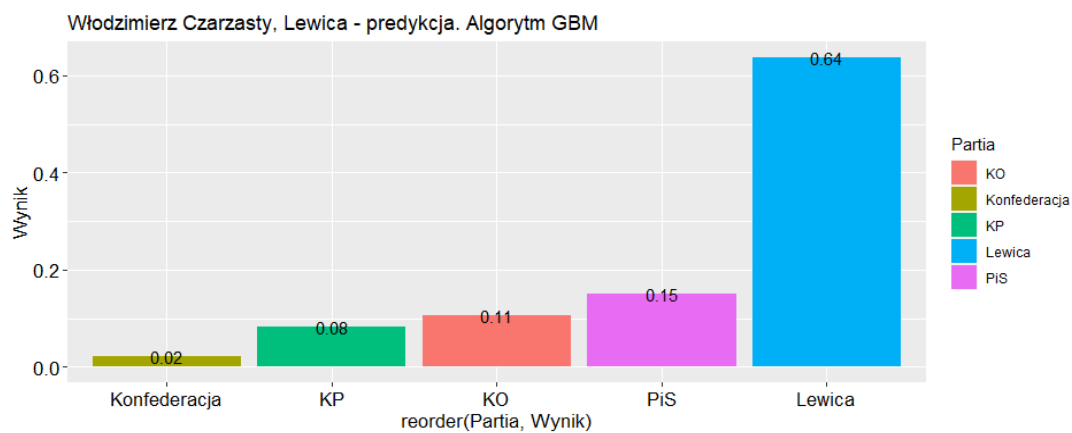
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 186. Predykcja modelem z użyciem algorytmu Drzewa decyzyjne – wypowiedź 4.



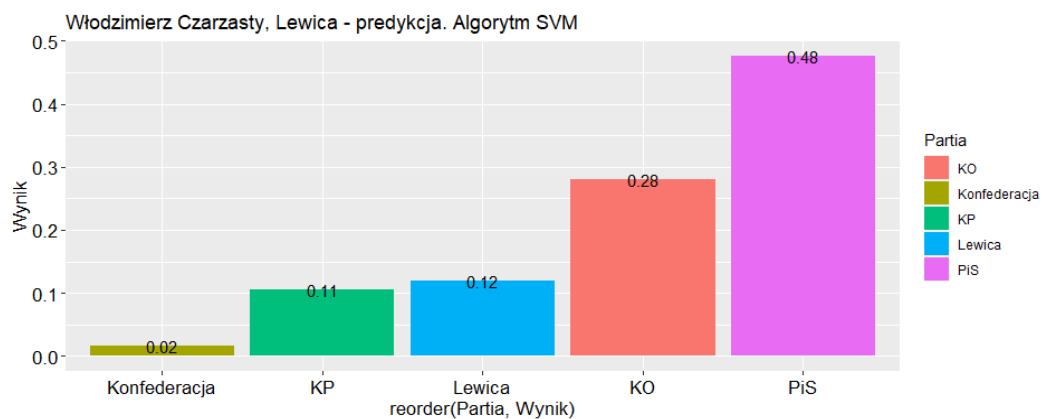
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 187. Predykcja modelem z użyciem algorytmu GBM – wypowiedź 4.



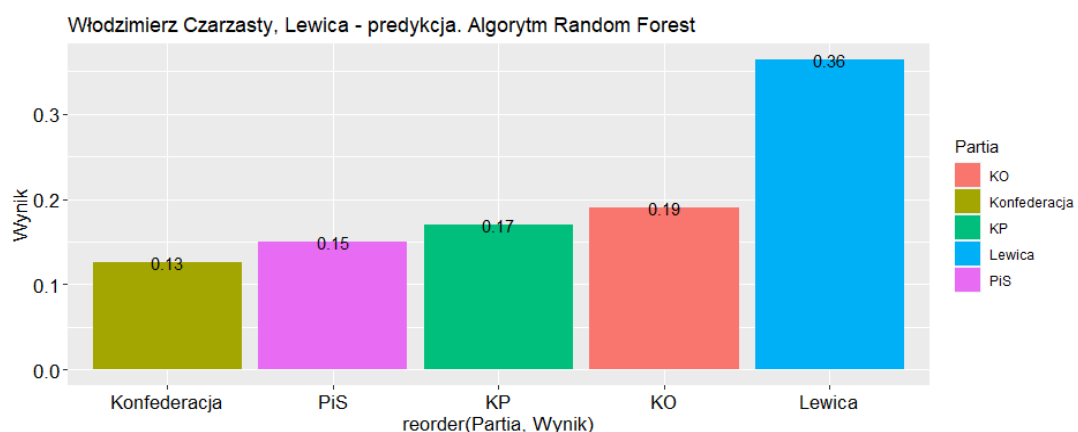
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 188. Predykcja modelem z użyciem algorytmu SVM – wypowiedź 4.



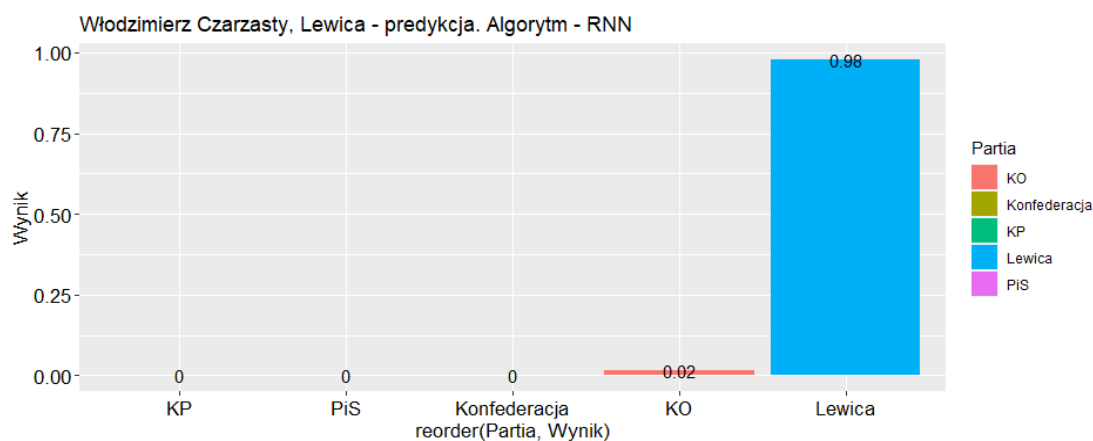
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 189. Predykcja modelem z użyciem algorytmu Random forest – wypowiedź 4.



Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 190. Predykcja modelem z użyciem algorytmu RNN – wypowiedź 4.



Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Wypowiedź posła KP Władysława Kosiniaka-Kamysza składająca się 470 słów została sklasyfikowana 5 razy poprawnie, co ciekawe najlepszy algorytm według oceny na całym zbiorze testowym (RNN) okazał się błędnie sklasyfikować wypowiedź jako posła Konfederacji.

Władysław Kosiniak-Kamysz, KP (oryginalna wypowiedź: 645 wyrazów – oczyszczona wypowiedź: 470 wyrazów)

Wypowiedź oryginalna:

Panie Marszałku! Wysoka Izbo! Panie Premierze! Słuchać Polaków, zmieniać Polskę - to było kiedyś wasze hasło. Jak słuchacie Polaków, to wiemy - po liczbie protestów i grup społecznych, które w tych protestach uczestniczą, od rolników po przedsiębiorców, od kobiet po osoby niepełnosprawne.

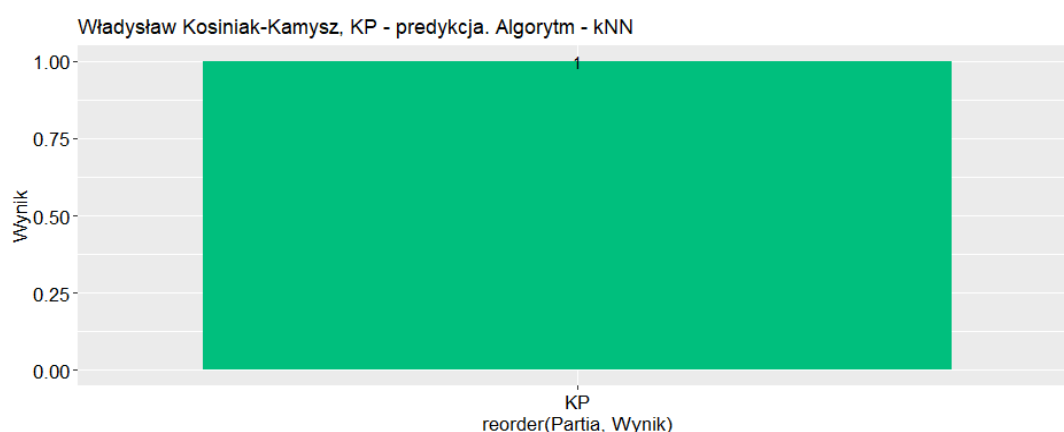
O tym, jak zmieniacie Polskę, było dzisiaj już bardzo dużo. Domagaliście się, żeby dzisiaj w tej debacie było więcej o Polsce, to usłyszycie teraz prawdę o tym, jak tę Polskę w ostatnich latach zmieniliście - według zasady: dzielić i zawłaszczać. Podział i zawłaszczenie. Podział na lepszy i gorszy sort. Podział medyków, prawdziwych bohaterów współczesnej Polski, na tych, którym się należą dodatki, i tych, którym pieniądze tydzień temu odebraliście. Pracodawców i przedsiębiorców podzieliliście według PKD, a nie według wielkości strat, które ponoszą przez COVID, i nie daliście im wsparcia. Samorządowców wczoraj podzieliliście na tych, którzy mają PiS-owską legitymację i dostali pieniądze z funduszu inwestycji lokalnych, i tych, którzy tej legitymacji nie posiadają - i mieszkańcom tej wspólnoty obywatelskiej się nic nie należy... ciąg dalszy wypowiedzi

Wypowiedź oczyszczona:

premierze słuchać polaków zmieniać polską było kiedyś wasze hasło słuchacie polaków wiemy liczbie protestów grup społecznych tych protestach uczestniczą od rolników przedsiębiorców od kobiet osoby niepełnosprawne zmieniacie polską było dzisiaj już bardzo dużo domagaliście żeby dzisiaj debacie było więcej polsce usłyszycie teraz prawdę tę polską ostatnich latach zmieniliście według zasady dzielić zawłaszczać podział zawłaszczenie podział lepszy gorszy sort podział medyków prawdziwych bohaterów współczesnej polski tych którym należą dodatki tych którym pieniądze tydzień temu odebraliście pracodawców przedsiębiorców podzieliliście według pkd nie według wielkości strat ponoszą przez covid nie daliście im wsparcia samorządowców wczoraj podzieliliście tych mają pisowską legitymację dostali pieniądze funduszu inwestycji lokalnych tych legitymacji nie

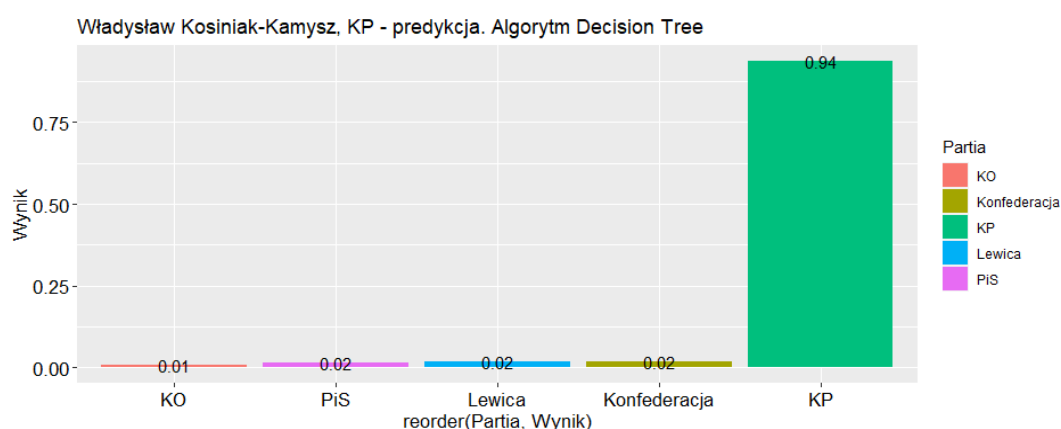
posiadają mieszkańcom wspólnoty obywatelskiej nic nie należy podzielić dziennikarzy media podzielić dziennikarzy tych którym odpowiada pytania tych których opinie przy konferencjach prasowych podzielić media te patriotyczne te przecież wiadomo ukrytej opcji podzielić nawet koła gospodyń wiejskich te którym należą dotacje te arimr nie zarejestrowały którym dotacji nie dacie najgorsze podzielić polską podzielić nasz ukochany kraj podzielić ten sposób nie ma gwarancji bezpieczeństwa nie liczba f gwarantuje bezpieczeństwo wspólnota polaków wspólnota narodowa świadomy swoich praw obowiązków naród wy ten naród tę wspólnotę rozbiliście zawłaszczyliście instytucje państwowe których uczyniliście pisowski folwark sądy których... ciąg dalszy wypowiedzi

Rysunek 191. Predykcja modelem z użyciem algorytmu kNN – wypowiedź 5.



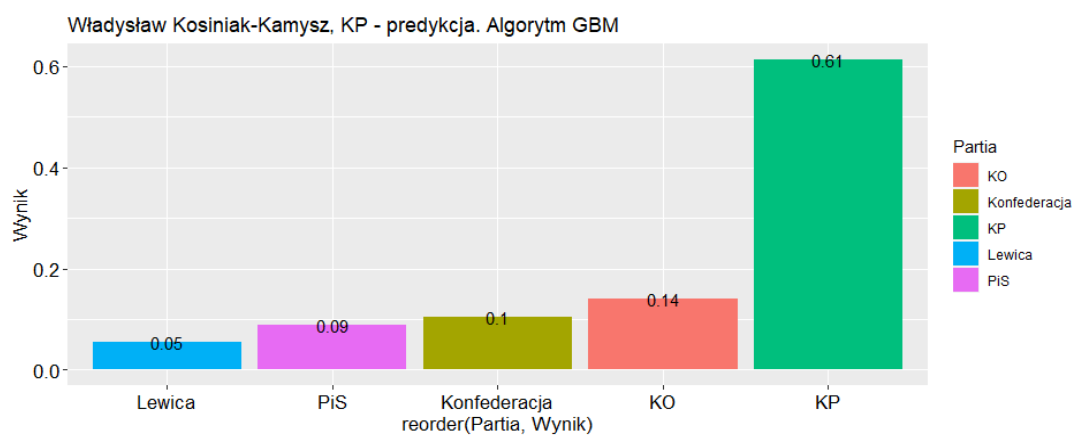
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 192. Predykcja modelem z użyciem algorytmu Drzewa decyzyjne – wypowiedź 5.



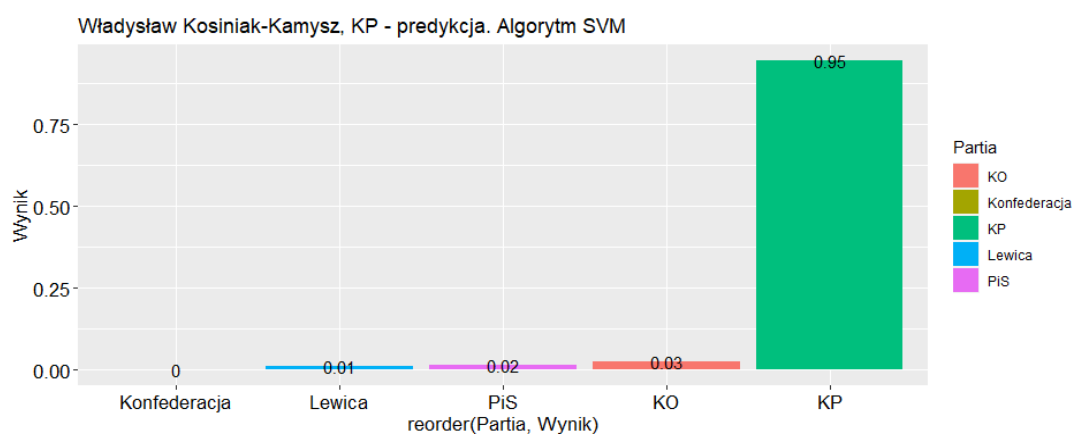
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 193. Predykcja modelem z użyciem algorytmu GBM – wypowiedź 5.



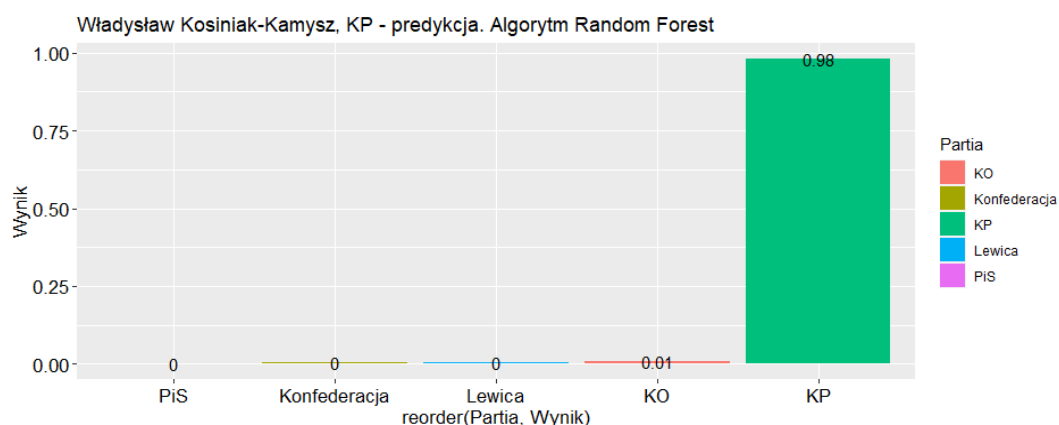
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 194. Predykcja modelem z użyciem algorytmu SVM – wypowiedź 5.



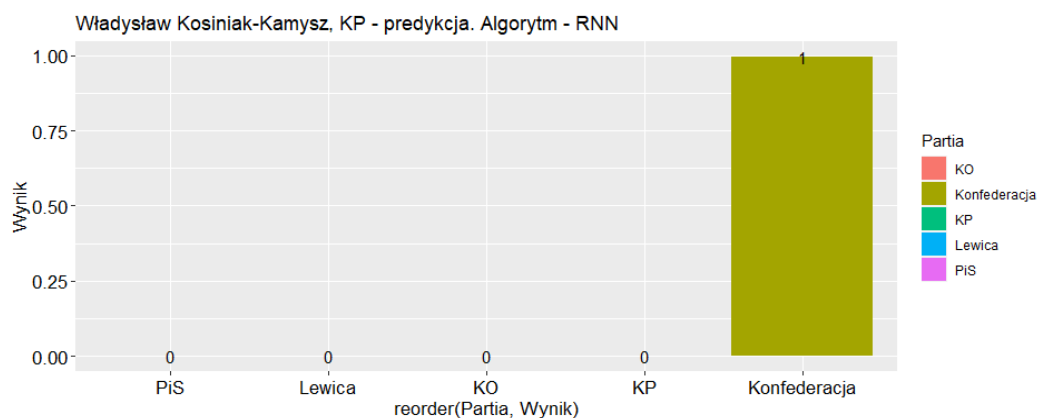
Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 195. Predykcja modelem z użyciem algorytmu Random forest – wypowiedź 5.



Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

Rysunek 196. Predykcja modelem z użyciem algorytmu RNN – wypowiedź 5.



Źródło: Opracowanie własne z wykorzystaniem biblioteki ggplot2

4.10. Podsumowanie rozdziału

W tej części skupiono się na zbadaniu możliwości, stworzeniu i testowaniu kilku modeli zbudowanych na bazie wcześniej oczyszczonego i odpowiednio przygotowanego zbioru danych. Stworzono sześć modeli predykcyjnych na bazie algorytmów sztucznej inteligencji w językach programowania *R* oraz *Python* z czego 3 osiągnęły bardzo dobry wynik 60% poprawnych klasyfikacji przy modelu 5 klasowym zbudowanym na danych w języku polskim. Reszta modeli uzyskała również zadowalający poziom. Bazując na pojedynczych wypowiedziach warto zastanowić się nad rozwiązaniem hybrydowym

wykorzystującym kilka algorytmów i na zasadzie głosowania większościowego dokonać predykcji przypadku.

5. Zakończenie

5.1. Weryfikacja hipotez

Ze względu na obszerny charakter poszczególnych rozdziałów, aby zachować przejrzystość odniesiono się w kolejności odpowiadającej hipotezom pomocniczym:

1/ Wraz z postępującą mediatyzacją polityki zwiększa się wykorzystanie mediów cyfrowych w komunikowaniu politycznym oraz zwiększa się możliwość analizy treści komunikowania politycznego.

Hipotezę można potwierdzić. Wraz z postępem mediatyzacji liczba użytkowników mediów społecznościowych rośnie. Polska jest na pierwszym miejscu, jeśli chodzi o liczbę tweetów politycznych w udziale wszystkich tweetów⁵¹⁶. Zarządzanie wizerunkiem polityka z wykorzystaniem Sieci staje się coraz powszechniejsze; Twitter zyskuje na popularności – w lipcu 2019 roku z Twittera w Polsce korzystało 5,1 miliona użytkowników, a w lipcu 2020 roku było ich 5,8 miliona (+14% użytkowników)⁵¹⁷. Autor zbudował model predykcyjny do 2024 roku z wykorzystaniem algorytmów autoregresyjnych ARIMA przewidujący liczbę kont w serwisie Twitter dla danych⁵¹⁸ dotyczących kandydatów do Parlamentu Europejskiego⁵¹⁹, z którego wynika, że liczba kont w 2024 będzie wahać się w przedziale: 614-688 kont (wzrost o około 40% do danych w 2019 r.). Stanowi to informację, że osoby zajmujące się polityką mają coraz większe pole do budowania zaufania i elektoratu politycznego w Sieci. Analizy pokazują, że Twitter czy Instagram, który ma tendencję wzrostową pod względem liczby użytkowników pozyskujących informacje

⁵¹⁶ C. Ducombe, *The Politics ...*, dz. cyt.

⁵¹⁷ W. Maguś, *Twitter jako ...*, dz. cyt., s. 507.

⁵¹⁸ W. Maguś, *Twitter jako ...*, dz. cyt., s. 519.

⁵¹⁹

właśnie z tego kanału codziennych informacji stanowią bazę do działań polityków w social media.

2/ Poszczególne grupy polityczne mają indywidualne słowniki słów oraz zwrotów wykorzystywane w komunikacji.

Hipotezę można potwierdzić. Podrozdział 3.4. pokazuje licznosc słów najczęściej używanych przez polityków poszczególnych partii politycznych. Podrozdział 3.9.1 jest odpowiednikiem takiego zestawienia dla danych z serwisu Twitter. Zestawy w każdej z grup pokazują, że mimo nieznacznych podobieństw słowniki różnią się. Z Tabeli 14. można odczytać, że najwyższą wartością podobieństwa cechuje się zawartość słownika partii KO oraz Lewica, gdzie parametr podobieństwa (*Jaccard index*) jest na poziomie 0,62. Najniższą wartość prezentuje kombinacja PiS oraz Konfederacja (parametr ten równa się 0,44) co pokazuje, że słowniki różnią się najbardziej. Podobna sytuacja dotyczy ngramów czyli zwrotów słów współwystępujących razem (rozdziały 3.7 oraz 3.9.2.) – zwroty, kolejności oraz zawartość różnią się między grupami.

3/ Słowa, zwroty oraz ich nacechowanie i emocje uczestników komunikowania politycznego (polityków i użytkowników mediów społecznościowych) zmieniają się w interwałach czasowych.

Hipotezę można potwierdzić. Wyraźnie trzeba zaznaczyć, że słowa zarówno od strony polityków, jak i internautów komentujących wydarzenia polityczne niosą duży ładunek emocjonalny. Osoby należące do partii rządzącej zwykle przez większość badanego okresu czasowego mają większy udział słów traktowanych jako mocno pozytywne i słabo pozytywne oraz mniejszy odsetek słów nacechowanych słabo i mocno negatywnie. Jeśli chodzi emocje, sytuacja wygląda podobnie – w partii rządzącej występuje większy udział słów o charakterze pozytywnych emocji (radość) niż negatywnych (smutek, złość). Na odwrót sytuacja wygląda jeśli chodzi o opozycję. W przypadku partii wspomnianych w *tweetach* nie widać takiej zależności.

4/ Istnieją podobieństwa między grupami polityków oraz użytkowników mediów społecznościowych wypowiadających się na tematy polityczne.

Hipotezę można potwierdzić. Rozdział 3.4. pokazujący licznosc słów oraz podział na klastry oraz rozdział 3.5, w którym porównywane są dendrogramy pokazują, że istnieją pewnego rodzaju powiązania, jeśli chodzi o konkretne grupy polityczne, jak i grupy słów

w poszczególnych grupach. Rozdział 3.9.1 oraz 3.9.3 pokazują, że każda z grup osób komentujących wydarzenia polityczne ma cechy wspólne i tematy się przenikają. Osoby wspominające o jednej partii, wspominają o innej i na odwrót. Ze względu na to, że każda z grup ma indywidualny słownik, nie jest możliwe obliczenie współczynnika splątania w klastrach, służącemu identyfikacji stopnia spokrewnienia tworzonych struktur, natomiast najwyższą wartością podobieństwa *Jaccard index* cechuje się zawartość słownika wspomnień partii KO oraz Lewica, gdzie parametr podobieństwa jest na poziomie 0,62. W przypadku danych z serwisu Twitter najwyższą wartością podobieństwa cechują się wpisy wspominające partię PiS oraz Lewicy, gdzie parametr podobieństwa jest na poziomie 0,63.

5/ Słowa wypowiediane przez uczestników w wypowiedziach komunikowania politycznego będą wskazywały do jakich partii pasują najbardziej.

Hipotezę można potwierdzić. Widać wyraźnie, że każda z partii prezentuje swój odrębny słownik słów i zwrotów, co jest czymś naturalnym. Grupy polityczne mają też swoje podobieństwa, na przykład powszechne zwroty. Niektóre z algorytmów predykcyjnych pozwoliły wyodrębnić słowa, które w sposób zdecydowany pozwalają na rozróżnienie partii: algorytm lasów losowych wskazuje na słowa jak: *nie, bardzo, koalicja* oraz *lewica* i *lewicy*; algorytm *GBM* wskazuje na podobny zbiór słów tylko w innej kolejności: *koalicja, nie, bardzo* oraz *lewicy* i *lewica*. Stanowi to sygnał, że algorytmy wychwytyują wzorce charakteryzujące daną grupę polityczną. Oznacza to, że Hipotezę 5 można uznać za potwierdzoną.

6/ Charakterystyka algorytmu, w tym jego wydajność, będzie miała wpływ na stopień dopasowania użytkownika do partii politycznej na podstawie jego wypowiedzi.

Hipotezę można potwierdzić. Każdy z wybranych algorytmów ma inne charakterystyki i prezentuje inną wydajność. Algorytmy o prostszych mechanizmach, jak *kNN* mają gorszą wydajność na danych testowych niż bardziej złożone, wymagające dłuższego czasu uczenia, jak maszyny wektorów nośnych czy sztuczne sieci neuronowe. Wzorce, które pasują do konkretnych partii wydają się być bardziej zauważalne dla algorytmów złożonych.

Po szeregu badań dla hipotezy głównej, która brzmi:

Wypowiadane słowa polityków oraz użytkowników mediów społecznościowych komentujących wydarzenia polityczne niosą ze sobą wartość informacyjną i po procesie kwantyfikacji oraz rafinacji danych mogą posłużyć do analizy treści i budowy modeli predykcyjnych dopasowania do grupy politycznej.

można uznać jednoznacznie, że słowa wypowiedane przez polityków oraz użytkowników mediów społecznościowych mają dużą wartość informacyjną i po odpowiednim ich przygotowaniu mogą posłużyć do budowy modeli predykcyjnych używanych następnie do przewidywania poglądów politycznych. Potwierdzeniem tego są wyniki wszystkich stworzonych modeli, które okazały się mieć wyższe wartości predykcji niż losowanie (przy modelu pięcioklasowym prawdopodobieństwo około 20% szans). Najlepsze z nich, wydajnościowo przekraczały 60%-ową poprawność klasyfikacji.

Osoby wspominające partie polityczne w serwisach społecznościowych nie oznaczają przynależności czy poczucia sympatii do partii. Często wyrażają niepochlebne opinie poprzez stosowanie słów niecenzuralnych, w istotnej części krytykę jednej lub kilku partii w swoich wypowiedziach. Zdecydowano się zbudować modele na danych osób przynależnych do partii politycznych czyli polityków.

Mając na uwadze, że wszystkie hipotezy szczegółowe zostały potwierdzone, można jednoznacznie uznać, że hipotezę główną można również uznać za potwierdzoną. Problemy badawcze, które zostały ujęte w niniejszej pracy potwierdzają, że zbadanie aktualnych kanałów, w których istnieje komunikowanie polityczne oraz ich zalet, zbudowanie i skonfigurowanie narzędzi do pobrania potrzebnych danych, sprawdzenie, określenie, pomiar językowych cech grup politycznych oraz utworzenie i ocena modeli predykcyjnych wykorzystujących kwestie językowe poszczególnych do klasyfikacji były obszarami, które należało zbadać w sposób kaskadowy wykorzystując otrzymane wnioski z poprzedniego etapu prac.

Głównym celem pracy było odpowiedzenie na pytanie: *czy, w jaki sposób i w jakim stopniu analiza treści w komunikowaniu politycznym i narzędzia sztucznej inteligencji mogą służyć do budowy modeli predykcyjnych dopasowania do grupy politycznej*⁵²⁰.

⁵²⁰ Grupy polityczne rozumiane w charakterze partii oraz koalicjantów (Prawo i Sprawiedliwość (PiS), Koalicja Obywatelska (KO), Koalicja Polska (KP), Konfederacja, Lewica).

Dowodem na osiągnięcie celu są przede wszystkim zbudowane istotne statystycznie modele spełniające te wymagania w bardzo zadowalającym poziomie. Każdy z rozdziałów pokazuje, że aspekt analizy treści komunikowania politycznego niesie ze sobą niezwykle korzyści w postaci zrozumienia kwestii w tych poruszanych treściach ich ilości, powiązań, nacechowania, emocji i znalezieniu wzorców, które są charakterystyczne dla każdej z grup.

5.2. Rekomendacje

Niniejsza praca jest dowodem, że wykorzystanie kilku dyscyplin naukowych może pozwolić na przełożenie analiz jakościowych na ilościowe. Analiza treści w połączeniu z metodami statystycznymi, algorytmami AI może dawać możliwości budowania narzędzi (w tym narzędzi komercyjnych) klasyfikujących poglądy polityczne, gospodarcze, czy społeczne. Warto zwrócić uwagę na kilka aspektów, które mogłyby zostać rozszerzone:

1. W obszarze mediów warto zauważyć, każda z grup politycznych ma swoje wzorce w zakresie komunikowania politycznego, zarówno jeśli chodzi o wypowiediane treści przez polityków, jak i internautów. Warto analizować je w różnych aspektach, aby poprawić poprawną komunikację pomiędzy wyborcami a politykami, aby odpowiadać na zapotrzebowanie i oczekiwania wyborców.
2. Internetowe wpisy, komentarze są idealnym zapleczem nie tylko do sprawdzania nastrojów, oczekiwań, zainteresowań grup ludzi, ale także są miejscem budowania poparcia politycznego w dzisiejszych czasach i świadomy polityk powinien mieć narzędzia do sprawdzania tych treści na bieżąco.
3. Ocena przynależności poglądów politycznych i zbudowanie modeli predykcyjnych na podstawie wpisów internetowych podobnych do tych użytych w pracy ze względu na swoją wielowektorowość powinna zostać ujęta w osobnym obszarze badawczym z uwzględnieniem szeregu metod, jak analizy sentymentów oraz innych rozwiązań (na przykład zbudowanie modeli partii na podstawie wpisów wszystkich jej członków wypowiadających się w mediach ich komentarzy i reakcji internautów) i stanowi ciekawy aspekt do dalszej rozbudowy.
4. Gromadzenie danych politycznych o publicznych dostępach do użycia. Wykonanie analiz wymagało napisania oprogramowania, które mogłoby gromadzić i układać dane o dostępie ogólnym w sposób wykonywalny do analiz. Opis procesu budowy

narzędzia z wykorzystaniem języka *Python* oraz biblioteki *Selenium* ze względu na obszerny oraz niepowiązany z dyscypliną nauki o mediach charakter został celowo pominięty. Należy pamiętać jednak, że zbudowanie i skonfigurowanie takiego narzędzia jest procesem pracochłonnym i wymagającym zasobu czasu. Dane z kilku źródeł mogłyby znacznie polepszyć siłę predykcji i wyniki poszczególnych wyników. Algoritmy mogłyby mieć lepszą możliwość nauczania się nowych, bardziej złożonych wzorców, a co za tym idzie posiadałby by schematy dzięki którym mogłyby klasyfikować dokładniej nowe przypadki. Stąd warto zaznaczyć, że oficjalne wypowiedzi polityków mogłyby być umieszczone w systemie ewidencji przeznaczone do szybkiej analizy.

5. Proces obróbki i budowania modeli tekstowych wymaga znacznych zasobów mocy obliczeniowej. Rozszerzenie baz oraz elementów potrzebnych do nauki nowych wzorców wymaga bardzo dużych nakładów sprzętowych. Dobrym rozwiązaniem byłoby sprawdzenie i rozwój możliwości rozwiązań chmurowych analizy tekstu (komercyjnych i niekomercyjnych) w tym zakresie.
6. Rozwijanie słowników nacechowania w języku polskim. Aby przystąpić do analiz statystycznych związanych z badaniem nacechowania poszczególnych wypowiedzi należałoby ciągle rozwijać słowniki typu: *SłowoSieć*. Wiadomo, że język ewaluuje, powstają nowe słowa, niektóre z nich nabierają innych znaczeń i warto aktualizować bazę słów pod tym względem.
7. Praca zawiera stworzone gotowe moduły w językach programowania *R* i *Python* do użycia w innych projektach komercyjnych lub niekomercyjnych. Umożliwia to ich wdrożenie i implementacje jako rozwiązania desktopowe, bądź internetowe jak serwis *REST* pozwalające na elastyczność w wykorzystaniu interfejsów *REST API* do łączenia komponentów i aplikacji w architekturach mikrousług⁵²¹. Zastosowanie tych modułów przyspieszy tworzenie aplikacji weryfikujących poglądy polityczne na podstawie wypowiadanych słów.

⁵²¹ Rest API, <https://www.ibm.com/pl-pl/cloud/learn/rest-api?msckid=c1354bfdaa9511ec809f18e7162e1eec>, 23.03.2022.

8. Pracę warto kontynuować rozbudowując (I) bazę do nauki modeli (źródeł danych) (II) poszukiwania innych algorytmów rozwiązujących problem klasyfikacji wypowiedzi politycznych oraz (III) stworzeniem rozwiązań predykcji na zasadzie głosowania większościowego puli modeli o najlepszych osiągnięciach z uwzględnieniem średniej ważonej tych modeli – co może jeszcze poprawić wydajność przewidywań.

Kończąc rozważania w niniejszej pracy trzeba zaznaczyć, że podjęty w niej problem analizy treści i budowania modeli predykcyjnych klasyfikowania poglądów politycznych z pewnością nie został wyczerpany. Warto zastanowić się nad innymi metodami analizy treści komunikowania politycznego i możliwością rozszerzania treści politycznych o nowe źródła. Obszar rozwoju kooperacji pomiędzy polityką, nauką o mediach, a analizą statystyczną i branżą IT należy do przyszłości budowania nowych strategii budowania poparcia politycznego.

6. Bibliografia

6.1. Źródła

Raporty

1. Grayling's report on how politicians across Europe are using social media, 2021, https://bpcc.org.pl/uploads/ckeditor/attachments/19242/GPol_-_European_survey_of_SM_use_amongst_MPs__Poland_highlighted_.pdf, 20.04.2022
2. GUS, Raport *Spółeczeństwo informacyjne w Polsce w 2021*, 06.06.2022
3. Raport *ePolacy 2019*, <https://www.gov.pl/documents>, 02.05.2022

Bazy danych oraz narzędzia ich pozyskania

Dane wypowiedzi polityków zostały zebrane z wykorzystaniem narzędzia *Selenium* skonfigurowanego przez autora pobierającego publiczne ogólnie dostępne wypowiedzi polityków z Sejmu ze strony rządowej (sejm.gov.pl). W przypadków danych wypowiedzi z Twittera wykorzystano skonfigurowaną przez autora bibliotekę *Bsi-sentiment* (GNU General Public License).

1. Biblioteka Selenium, <https://pypi.org/project/selenium/> 02.02.2021.
2. Bsi-sentiment, <https://pypi.org/project/bsi-sentiment/>, GNU General Public License v3 (GPLv3), 05.05.2021
3. Wypowiedzi sejmowe, <https://sejm.gov.pl>, 20.04.2021

Akty prawne

1. Dane dot. wyników Sejmu 2019, <https://sejmsenat2019.pkw.gov.pl/sejmse-nat2019/pl/wyniki/sejm/pl>, 01.09.2021
2. Dziennik Ustaw Rzeczypospolitej Polskiej, <http://isap.sejm.gov.pl/isap.nsf/download.xsp/WDU20190001955/O/D20191955.pdf>, 01.09.2021
3. Kodeks wyborczy (Dz.U. z 2011 r. nr 21, poz. 113 z późn. zm.), <https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20110210112>, 17.11.2022
4. Obwieszczenia PKW – Wybory Prezydenta Rzeczypospolitej Polskiej w 2020 r., <https://prezydent20200628.pkw.gov.pl/prezydent20200628/pl/obwieszczenia>, 01.09.2021
5. Regulamin Sejmu – tekst jednolity, <https://www.sejm.gov.pl/prawo/regulamin/kon7.htm>, 19.11.2022
6. Rozdział 4 ustawy z dnia 27 czerwca 1997 r. o partiach politycznych (Dz.U. Nr 98, poz. 604), https://pkw.gov.pl/uploaded_files/1498132393_ustawa_o_partiach_politycznych_-_wersja_aktualna.pdf, 08.09.2021

Strony internetowe

1. A short guide to the Polish political scene, <https://notesfrompoland.com/2019/11/13/a-short-guide-to-the-polish-political-scene>, 08.09.2021
2. Accuracy Recall Precision f-score, <https://towardsdatascience.com/accuracy-recall-precision-f-score-specificity-which-to-optimize-on-867d3f11124>, 20.05.2022
3. Analiza koszykowa, <https://businessintelligence.pl/analiza-koszykowa>, 01.02.2022
4. Association Mining with R, <http://r-statistics.co/Association-Mining-With-R.html>, 28.12.2021
5. Association rules, <https://towardsdatascience.com/association-rules-2-aa9a77241654>, 28.12.2021
6. Auto ARIMA, <https://www.rdocumentation.org/packages/forecast/versions/8.16/topics/auto.arima>, 20.07.2022
7. Biblioteka arules, <https://www.rdocumentation.org/packages/arules/versions/1.7-2>, 28.12.2021
8. Biblioteka arulesViz, <http://mhahsler.github.io/arulesViz>, 28.12.2021
9. Biblioteka caret, <https://www.rdocumentation.org/packages/caret/versions/6.0-93>, 02.04.2022
10. Biblioteka cluster, <https://www.rdocumentation.org/packages/cluster/versions/2.1.2/topics/agnes>, 28.12.2021
11. Biblioteka cluster, <https://www.rdocumentation.org/packages/cluster/versions/2.1.3>, 28.12.2021
12. Biblioteka dendextend, <https://www.rdocumentation.org/packages/dendextend/versions/1.16.0>, 04.05.2021
13. Biblioteka dendextend, <https://www.rdocumentation.org/packages/dendextend/versions/1.16.0>, 28.12.2021
14. Biblioteka dplyr, <https://www.rdocumentation.org/packages/dplyr/versions/0.7.8>, 02.05.2021
15. Biblioteka e1071, <https://cran.r-project.org/web/packages/e1071/index.html>, 20.02.2021
16. Biblioteka factoextra, <https://www.rdocumentation.org/packages/factoextra/versions/1.0.7>, 28.12.2021
17. Biblioteka forecast, <https://www.rdocumentation.org/packages/forecast/versions/8.16>, 20.02.2022
18. Biblioteka ggplot2, <https://www.rdocumentation.org/packages/ggplot2/versions/3.3.6>, 02.05.2021
19. Biblioteka ggraph, <https://www.rdocumentation.org/packages/ggraph/versions/2.1.0>, 11.05.2022
20. Biblioteka hclust, <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/hclust>, 28.12.2021
21. Biblioteka randomforest, <https://www.rdocumentation.org/packages/randomForest/versions/4.7-1.1/topics/randomForest>, 20.02.2021
22. Biblioteka rpart, <https://www.rdocumentation.org/packages/rpart/versions/4.1-15/topics/rpart>, 12.01.2021
23. Biblioteka rpart.plot, <https://www.rdocumentation.org/packages/rpart.plot/versions/3.1.1/topics/rpart.plot>, 12.01.2021
24. Biblioteka tensorflow, <https://pypi.org/project/tensorflow>, 01.05.2022
25. Biblioteka tidytext, <https://www.rdocumentation.org/packages/tidytext/versions/0.3.3>, 02.05.2021
26. Biblioteka tm, <https://www.rdocumentation.org/packages/tm/versions/0.7-8>, 02.05.2021
27. Biblioteka widyr, https://www.rdocumentation.org/packages/widyr/versions/0.1.4/topics/pairwise_cor, 10.05.2021

28. Big data by IBM, <https://www.redbooks.ibm.com/redpapers/pdfs/redp5070.pdf>, 20.08.2021
29. Classification accuracy, <https://developers.google.com/machine-learning/crash-course/classification/accuracy>, 02.06.2022
30. Co znaczy osiem gwiazdek ruch ośmiu gwiazd to inicjatywa anty-pis, <https://natemat.pl/314263,co-znaczy-osiem-gwiazdek-ruch-osmiu-gwiazd-to-inicjatywa-anty-pis-rozmowa>, 20.11.2022
31. Cohens Kappa in spss statistics, <https://statistics.laerd.com/spss-tutorials/cohens-kappa-in-spss-statistics.php>, 02.04.2022
32. COVID-19 w Polsce, https://pl.wikipedia.org/wiki/Pandemia_COVID-19_w_Polsce, 28.12.2021
33. De Fleur Model of Communication, <https://www.communicationtheory.org/de-fleur-model-of-communication>, 22.05.2022
34. Deep learning, <https://www.ibm.com/cloud/learn/deep-learning>, 12.05.2021
35. Definicja Sieci PWN, <https://encyklopedia.pwn.pl>, 22.03.2022
36. Dlaczego warto wprowadzić JOW?, <http://www.kukiz15.org/program/krajowy/jow>, 18.11.2022
37. Elon Musk sfinalizował zakup Twittera, <https://dorzeczy.pl/ekonomia/362503/elon-musk-sfinalizowal-zakup-twittera-agrawal-zwolniony.html>, 02.02.2023
38. Ile mandatów jest do zdobycia w okręgach wyborczych?, <https://www.rp.pl/polityka/art1083171-ile-mandatow-jest-do-zdobycia-w-okregach-wyborczych>, 19.11.2022
39. Introduction to hierarchical clustering, <https://towardsdatascience.com/introduction-hierarchical-clustering-d3066c6b560e>, 28.12.2021
40. Językowy obraz świata, <https://encyklopedia.pwn.pl/haslo/jezykowy-obraz-swiata;3917950.html>, 11.11.22
41. Lasswell's Model, <https://www.communicationtheory.org/lasswells-model>, 18.05.2022
42. M. Miłkowski, Biblioteka odmian słów, <https://github.com/morfologik/polimorfologik/blob/master/src/LICENSE.txt>, 28.12.2021, Copyright (c) \$copyright_date, Marcin Miłkowski All rights reserved
43. M. Szutiak, Prawie 90% gospodarstw domowych w Polsce ma dostęp do szerokopasmowego Internetu - podaje GUS, <https://www.telepolis.pl/wiadomosci/prawo-finanse-statystyki/gus-xszerokopasmowy-internet-90-procent-gospodarstw-d-mowych-w-polsce>, 01.09.2021
44. Machine learning basics with the k-nearest neighbors algorithm, <https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761>, 10.01.2022
45. McNemars test for machine learning, <https://machinelearningmastery.com/mcnemars-test-for-machine-learning>, 10.01.2022
46. Na czym polega metoda d'Hondta?, rp.pl, 20.11.2022
47. Nasza klasa znika z sieci, <https://businessinsider.com.pl/technologie/nasza-klasa-znika-z-sieci-w-szczycie-miala-14-mln-uzytkownikow/hswvc80>, 12.08.2021
48. Newcomb's model, <https://www.communicationtheory.org/the-newcomb%E2%80%99s-model/>, 20.05.2022
49. Papież Benedykt XVII, *Przemówienie z okazji zgromadzenia plenarnego Papieskiej Rady ds. Środków Społecznego Przekazu*, 28.02.2011, https://opoka.org.pl/biblioteka/W/WP/benedykt_xvi/przemowienia/massmedia_28022011.html, 01.11.2022
50. Parties and Elections, <http://www.parties-and-elections.eu/poland.html>, 08.09.2021
51. Performance measures Cohens Kappa statistic, <https://thedata scientist.com/performance-measures-cohens-kappa-statistic>, 10.01.2022
52. Pierwsze polskie komunikatory- przeżyjmy to jeszcze raz, <https://antyweb.pl/pierwsze-polskie-komunikatory-przezyjmy-to-jeszcze-raz>, 12.08.2021
53. Polska Rzeczpospolita Ludowa, <https://polskiemiesiace.ipn.gov.pl/mie/wszystkie-wydarzenia/czerwiec-1989/polska/historia/117922,PRL-Polska-Rzeczpospolita->

- Ludowa.html, 20.06.2022
54. Postulaty Kukiz '15, <https://pl.wikipedia.org/wiki/Kukiz%E2%80%9915>, 08.09.2021
 55. Precision and recall, <https://developers.google.com/machine-learning/crash-course/classification/precision-and-recall>, 02.06.2022
 56. PRL czerwiec 1989, <https://polskiemiesiace.ipn.gov.pl/mie/wszystkie-wydarzenia/czerwiec-1989/polska/historia/117922,PRL-Polska-Rzeczpospolita-Ludowa.html>, 20.06.2022
 57. Rest API, <https://www.ibm.com/pl-pl/cloud/learn/rest-api?msckid=c1354bfdaa9511ec809f18e7162e1eec>, 23.03.2022
 58. Riley & Riley Model of Communication, <https://www.communicationtheory.org/riley-riley-model-of-communication>, 20.05.2022
 59. Schramm Model, <https://www.managementstudyguide.com/schramm-model-of-communication.htm>, 21.05.2022
 60. Silhouette Method , <https://towardsdatascience.com/silhouette-method-better-than-elbow-method-to-find-optimal-clusters-378d62ff6891>, 28.12.2021
 61. Tensorflow, <https://www.tensorflow.org> , 09.01.2022
 62. Text Clustering with R, <https://medium.com/@SAPCAI/text-clustering-with-r-an-introduction-for-data-scientists-c406e7454e76>, 28.12.2021
 63. Timeline of social media, https://en.wikipedia.org/wiki/Timeline_of_social_media, 12.08.2021
 64. Twitter Statistics, <https://blog.hootsuite.com/twitter-statistics/> , 12.08.2021
 65. Użytkownicy Social Media w Polsce i na Świecie, <https://www.whysosocial.pl/uzytkownicy-social-media-w-polsce-i-na-swiecie> , 12.08.2021
 66. Warszawa będzie mieć więcej posłów? PKW wzywa Sejm do korekt, <https://warszawa.wyborcza.pl/warszawa/7,54420,29066108,warszawa-bedzie-miec-wiecej-poslow-panstwowa-komisja-wyborcza.html?disableRedirects=true>, 18.11.2022
 67. What is an ARIMA model, <https://towardsdatascience.com/what-is-an-arima-model-9e200f06f9eb>, 17.06.22
 68. What is Big Data, https://www.sas.com/pl_pl/insights/big-data/what-is-big-data.html, 26.08.2021
 69. Wybory do Senatu 2019 podział mandatów, <https://forsal.pl/artykuly/1420851,wybory-do-senatu-2019-podzial-mandatow-jak-liczone-sa-glosy.html>, 20.11.2022

6.2. Opracowania

1. Adamik-Szysia M., *Social media as a tool of political permanent campaign on the example of the activity of polish politicians*, „SWS Journal of Social Sciences and Art”, 2019, nr 2
2. Adamski A., *Media as the Intersphere of Human Life: Another View on the Mediatization of Communication Theory*, (red.) D. Petranová, S. Magá, „MEGATRENDS AND MEDIA Media Farm – Totems and Taboo”, 2015
3. Adamski A., *Media w analogowym i cyfrowym świecie. Wpływ cyfrowej rewolucji na rekonfigurację komunikacji społecznej*, Dom Wydawniczy ELIPSA, Warszawa 2012
4. Adamski A., Przywara B., Przybyło S., *The Church 'S Attitude To The Means And Philosophy Of Communication In The 9th And 16th Centuries: Courage And Trust Versus Escape And Censorship*, „Konštantínove listy/Constantine s Letters”, nr 13(2), 2020
5. Ampuja M., Koivisto J., Välvirronen E., *Strong and weak forms of mediatization theory: A critical review*, 2014
6. Anderberg M. R., *Cluster Analysis for Applications*, Academic Press, New York 1973
7. Annusewicz O., *Funkcje komunikowania politycznego w Internecie*, „Studia politologiczne”, 2020, nr. 14
8. Annusewicz O., *Marketing polityczny – pytania i kontrowersje* [w:] Pietrzyk-Zieniewicz E. (red.), *Scena polityczna i media, Miraże sukcesu – ryzyko autoprezentacji*, Warszawa 2004

9. Anusiewicz J., Dąbrowska A., Fleischer M., *Językowy obraz świata i kultura. Projekt koncepcji badawczej*, Język a kultura (tom: 13), Acta Universitatis Wratislaviensis (tom: 2218), Warszawa 2000
10. Anusiewicz J., *Problematyka językowego obrazu świata w poglądach niektórych językoznawców i filozofów niemieckich XX wieku*, [w:] *Językowy obraz świata*, (red.) Bartmiński J., Lublin 1990
11. Asp K., *Måktiga massmedier: Studier i politisk opinionsbildning [Powerful mass media: studies in political opinion-formation]*, Akademilitteratur, Stockholm 1986
12. Awdiejew A., *Wartościowanie wymuszone, a szacunek dla odbiorcy w dyskursie politycznym*, 2018
13. Balczyńska-Kosman A., *Język dyskursu publicznego w polskim systemie politycznym*, „Środkowoeuropejskie Studia Polityczne”, 2013, nr 2
14. Bartmiński J., *Językowe podstawy obrazu świata*, Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, 2006
15. Bartmiński J., Tokarski R., *Językowy obraz świata a spójność tekstu*, [w:] *Teoria tekstu. Zbiór studiów*, (red.) Dobrzyńska T., Wrocław, 1986
16. Bartnik Cz. S., *Personalizm*, Oficyna wydawnicza „Czas”, Lublin 1995
17. Bauer Z., *Dziennikarstwo wobec nowych mediów*, Universitas, 2009
18. Baylon Ch., Mignot X., *Komunikacja*, Kraków 2008
19. Bellio E., Crestini G., Arkoudas A., *Twitter and politics: Evidence from the US presidential elections 2016*, „Journal of Marketing Communications”, 2018
20. Bogołębska B., *Retoryka w odmianach dziennikarstwa*, „FOLIA LITTERARIA POLONICA”, 2016, nr 1(31)
21. Bordewijk J. L., van Kaam B., *Towards a new classification of teleinformation services*, [w:] D. McQuail (red.), *McQuail's Reader in Mass Communication Theory*, London, 1986
22. Boulesteix A. L., Janitza S., Kruppa J., König I. R., *Overview of Random Forest Methodology and Practical Guidance with Emphasis on Computational Biology and Bioinformatics*, „Department of Statistics University of Munich Technical Report”, 2012, nr 129
23. Bralczyk J., *O języku propagandy i polityki*, Wydawnictwo Trio, Warszawa 2007
24. Breiman L., Friedman J. H., Olshen R. A., Stone C. J., *Classification and regression trees*, Chapman & Hall, 1984
25. Bucur M., Ban A., *Study on the Westley Maclean model and the co-orientation model and their exemplification in an industrial enterprise*, „Acta Marisiensis Seria Technologica”, 2019, nr 16(2)
26. Bugajski M., *Język w komunikowaniu*, Warszawa 2006
27. Chawla N. V., Bowye K.W., Hall L. O., Kegelmeyer W. P., *SMOTE: Synthetic Minority Over-sampling Technique*, „Journal of Artificial Intelligence Research”, 2002, nr 16
28. Colleoni E., Rozza A., Arvidsson A., *Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data*, „Journal of Communication”, 2014, nr 64(2)
29. Cortes C., Vapnik V., *Support-vector networks*, „Machine Learning”, 1995, nr 20 (3)
30. Courth F., *Bóg trójjedyniej miłości*, tł. Kowalczyk M., adaptacja Bartel L., Mickiewicz F., (oryg. Der Gott der dreifaltigen Liebe), Pallotinum, Poznań 1997
31. Cox M., Ellsworth D., *Managing Big Data for Scientific Visualization*, ACM SIGGRAPH '97 Course #4, *Exploring Gigabyte Datasets in Real-Time: Algorithms, Data Management, and Time-Critical Design*, Los Angeles 1997
32. Czechowska-Derkacz B., Chomik D., Wojsław J. (red.), *Potencjał informacyjny rafinacji Big Data. Biznes i Media*, [w:] *Media Biznes Kultura. Media w procesie zmian – etyka i komunikacja*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk 2018
33. Czechowska-Derkacz B., *Inforozrywka i tabloidyżacja w polskich mediach po 1989 roku w kontekście komunikacji politycznej*, [w:] J. Kreft, R. Stopikowski (red.), *Media, biznes, kultura*, Marpress, 2011, nr 2

34. Czelakowska A., *Językowe środki wartościowania negatywnego w wypowiedziach polityków*, Annales Academiae Paedagogicae Cracoviensis, „Studia Linguistica IV”, 2008
35. Czyżewski M., Kowalski S., Piotrowski A., *Rytualny chaos. Studium dyskursu publicznego*, Areus, Kraków 1997
36. Czyżewski M., *Rytualny chaos. Studium dyskursu publicznego*, (red.) Czyżewski M., Kowalski S., Piotrowski A., Wrocław 2004
37. Czyżowski D., Porębski L., *Media społecznościowe w kampanii wyborczej. Wykorzystanie Twittera w polskich wyborach parlamentarnych w roku 2015*, „STUDIA I ANALIZY” Studia politologiczne, 2017, nr 45
38. Dahl M., *Wpływ transformacji ekonomicznej na sferę społeczną w Polsce po 1989 roku*, Myśl Ekonomiczna i Polityczna, 2010
39. Dancey C. P., Reidy J., *Statistics without Maths for Psychology*, Pearson Education 2017
40. de Kerckhove D., *Inteligencja otwarta. Narodziny społeczeństwa sieciowego*, Mikom, Warszawa 2001
41. DeFleur M., *Theories of Mass Communication*, David McKay, New York 1970
42. Denton R., Woodward G., *Political Communication in America*, Longman, New York 1998
43. Devika M. D., Sunitha C., Ganesh A., *Sentiment Analysis: A Comparative Study On Different Approaches*, „Procedia Computer Science”, 2016, nr 87
44. Dobek-Ostrowska B., *Komunikowanie polityczne i publiczne*, PWN, Warszawa 2006
45. Dobek-Ostrowska B., *Media masowe i aktorzy polityczni w świetle studiów nad komunikowaniem politycznym*, Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław 2004
46. Dobrzyńska T., *Tekst. Próba syntezy* [w:] *Współczesny język polski*, (red.) Bartmański J., Lublin 2001
47. Doug L., *Data Management: Controlling Data Volume, Velocity, and Variety*, „Application Delivery Strategies”, META Group (currently with Gartner), 2001
48. Downes E. J., McMillan S. J., *Defining Interactivity: A Qualitative Identification of Key Dimensions*, *New Media and Society*, 2000
49. Driessens O., *The celebrityization of society and culture: Understanding the structural dynamics of celebrity culture*, „International Journal of Cultural Studies”, 2013, nr 6 (16)
50. Ducombe C., *The Politics of Twitter: Emotions and the Power of Social Media*, Oxford Academic, 2019
51. Dudek A., *Historia polityczna Polski 1989–2012*, Znak, Kraków 2013
52. Dumitrescu D., Boiangiu C., *A Study of Image Upsampling and Downsampling Filters*, „MDPI”, 2019, nr 8(2)
53. Dziob A., Piasecki M., Rudnicka E., *plWordNet 4.1—a Linguistically Motivated, Corpus-based Bilingual Resource*. In: Fellbaum C., Vossen P., Rudnicka E., Maziarz M., & Piasecki M. (Eds.), *Proceedings of the 10th Global WordNet Conference: July 23-27, 2019, Wrocław (Poland)*. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, pp. 353–362. (pdf,Bibtex); Piasecki M., Szpakowicz S., Broda B., *A Wordnet from the Ground Up*. Wrocław, Oficyna Wydawnicza Politechniki Wrocławskiej, 2009. (pdf,Bibtex)
54. Eco U., *Wymyślanie wrogów*, Rebis, 2011
55. Eulau H., Eldersveld S. J., Janowitz M., *Political Behavior: A Reader in Theory and Research*, „American Journal of Sociology”, nr 63(1), 1956
56. Evans A., Twomey J., Talan S., *Twitter as a Public Relations Tool*, „Public Relations Journal”, nr 5(1)
57. Everitt B., *Cluster Analysis*, Heinemann Educ Books, London 1975
58. Falkowski A., Cwalina W., *Marketing polityczny. Perspektywa psychologiczna*, Gdańsk 2005
59. Filas R., *Dwadzieścia lat przemian polskich mediów (1989–2009) w ujęciu periodycznym*, Zeszyty prasoznawcze, Kraków 2010
60. Fletcher S., Islam M. Z., *Comparing sets of patterns with the Jaccard index*, „Australasian Journal of Information Systems”, 2018, nr 22

61. Frame A., Brachotte G., *Le tweet stratégique: Use of Twitter as a PR Tool by French Politicians*, „Public Relations Review”, nr 41(2)
62. Fras J., *Język propagandy politycznej*, [w:] B. Dobek-Ostrowska, J. Fras, B. Ociepa, *Teoria i praktyka propagandy*, Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław 1997
63. Fras J., *Komunikacja polityczna. Wybrane zagadnienia gatunków i języka wypowiedzi*, Wrocław 2005
64. Gajowniczek T., *Internet w komunikowaniu politycznym*, Wydział Nauk Społecznych Uniwersytet Warmińsko-Mazurski w Olsztynie, 2019
65. Garlicki L., *Polskie prawo konstytucyjne. Zarys wykładu*, Wolters Kluwer, 2020, s. 37–39
66. Gazdar G., *Pragmatics, Implicature, Presupposition and Logical Form*, New York 1979
67. Gers F. A., Schmidhuber J., Cummins F., *Learning to forget: Continual prediction with LSTM*, Neural Computation, 2000, nr 12(10)
68. Goban-Klas T., *Zarys historii i rozwoju mediów. Od malowideł naskalnych do multimedialnych*, Wydawnictwo Naukowe Akademii Pedagogicznej w Krakowie, Kraków 2001
69. Goban-Klas T., *Nauki o mediach i komunikacji społecznej jako nowa dyscyplina nauk humanistycznych*, „Studia Medioznawcze”, 2008, 2(33)
70. Goban-Klas T., *Cywilizacja medialna. Geneza, ewolucja, eksplozja*, Wydawnictwa Szkolne i Pedagogiczne, Warszawa 2005
71. Goban-Klas T., *Media i komunikowanie masowe. Teorie i analizy prasy, radia, telewizji i Internetu*, Wydawnictwo Naukowe PWN, Warszawa 2004
72. Goban-Klas T., *Nauki o mediach – baza czy nadbudowa nowych teorii nauk społecznych?*, Uniwersytet Jagielloński, Kraków 2012
73. Goban-Klas T., *Pochwała medioznawstwa czyli renesans McLuhana*, Studia Medioznawcze, 2000
74. Goban-Klas T., *Rwący nurt mediów mediocen – nowa fala mediatyzacji życia społecznego pisma z lat 2012–2020*, Wyższa Szkoła Informatyki i Zarządzania w Rzeszowie, Kraków–Rzeszów 2020
75. Goban-Klas T., *System informacji w państwie, a system komunikacji społecznej*, Przekazy i Opinie, 1989, nr 2
76. Goban-Klas T., *Termin „media” – krytyka, obrona i objaśnienie*, Między Klio a Themis Księga dedykowana Profesorowi Jackowi Sobczakowi, Adamowski J. W. (red.) Wallas T. (red.), Kakarek K. (red.), Wydawnictwo Naukowe SILVA RERUM, 2016
77. Goban-Klas T., *The Media in the “Liquid” Coronavirus Pandemic of 2020. Communication During Lockdown: Separate but Together*, „Studia Medioznawcze”, 2020, nr 4(83)
78. Goban-Klas T., *Zarys historii i rozwoju mediów. Od malowideł naskalnych do multimedialnych*, Wydawnictwo Naukowe AP, Kraków 2001
79. Gogołek W., *Potencjał informacyjny nowych mediów – rafinacja informacji sieciowej*, [w:] 350 lat polskiej prasy, mat. pokonf. Uniwersytet Warszawski, Warszawa 2011
80. Gogołek W., *Rafinacja dużej skali zasobów sieciowych — Big Data. Dziennikarskie źródło informacji*, 2016
81. Gogołek W., *Rafinacja informacji sieciowej*, [w:], Jastriebow A., Raczyńska M., *Informatyka w dobie XXI wieku, Nauka, Technika, Edukacja a nowoczesne technologie informacyjne*, Radom 2011
82. Gosek M., *Cyfrowy uczeń, Dylematy mediatyzacji*, Akapit, Toruń 2020
83. Gotkowski P., *Szeptanie w komunikacji marketingowej*, „Marketing w Praktyce”, 2009, nr 42
84. Grabias S., *O ekspresywności języka*, Lublin 1981
85. Grzegorzczkowska R., *Pojęcie językowego obrazu świata*, [w:] *Językowy obraz świata*, (red.) Bartmiński J., Lublin 1990
86. Grzegorzczkowska R., *Wprowadzenie do semantyki językoznawczej*, Warszawa 2001
87. Grzegorzczkowska R., *Wstęp do językoznawstwa*, PWN, Warszawa 2007
88. Grzenia J., *Komunikacja językowa w Internecie*, Warszawa 2006

89. Gurevitch M., Blumer J., *Political Communication Systems and Democratic Values*, [w:] Lichtenberg J. (red.), *Democracy and the Mass Media*, Cambridge University Press, Cambridge 1990
90. He B., Patel M., Zhang Z., Chang K.C.C., *Accessing the deep web*, *Communications of the ACM*, New York, USA, 2007, nr 50 (5)
91. Hepp A., *Transforming Communications. Media-related Changes in Times of Deep Mediatization*, „COMMUNICATIVE FIGURATIONS”, 2017, nr 16
92. Herman W., *Przesunięcie semantyczne w nowej leksyce radia i telewizji*, Semper, 2008
93. Hernes G., *Det mediavridde samfunn [The media-twisted society]*, in Hernes G. (red.) *Forhandlingsekonomi og blandningsadministrasjon*. Bergen: Universitetsforlaget, 1978
94. Hjarvard S., *The mediatization of religion: Theorising religion, media and social change*, „Northern Lights”, 2008, nr 6(1)
95. Hjarvard S., *The mediatization of Culture and Society*, Routledge, London and New York 2013
96. Hjarvard S., *The mediatization of Society. A Theory of the Media as Agents of Social and Cultural Change*, *Nordicom Review* 29, 2008, nr 2
97. Ho L., Xie M., *The use of ARIMA models for reliability forecasting and analysis*, “Computers & Industrial Engineering”, 1998, nr 35(1-2)
98. Huang S., Cai N., Penzuti P., Narandes S., Wang Y., Wayne X., *Applications of Support Vector Machine (SVM) Learning in Cancer Genomics*, *Cancer Genomics Proteomics*, 2018, nr 15(1)
99. Iyyer M., Enns P., Boyd-Graber J., P. Resnik, *Political Ideology Detection Using Recursive Neural Networks*, Computer Science, Linguistics, iSchool, and UMIACS University of Maryland, 2014
100. Iyyer M., Enns P., Boyd-Graber J., Resnik P., *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Baltimore, Maryland, USA, June 23-25, 2014
101. Jansson A., *The Mediatization of Consumption: Towards an Analytical Framework of Image Culture*, „Journal of Consumer Culture”, 2002, nr 2(1)
102. Jedliński R., *Językowy obraz świata wartości w wypowiedziach uczniów kończących szkołę podstawową*, Wydawnictwo Naukowe Akademii Pedagogicznej, Kraków 2000
103. Jenkins H., *Convergence Culture. Where Old and New Media Collide*, NEW YORK UNIVERSITY PRESS, 2006
104. Johnson S. C., *Hierarchical Clustering Schemes*, *Psychometrika*, 1967, nr 32(3)
105. Jungherr A., *Twitter in Politics: A Comprehensive Literature Review*, University of Konstanz - Deptment of Politics and Public Administration, 2014
106. Kaczmarek-Śliwińska M., *Hejt jako źródło sytuacji kryzysowych podmiotów w przestrzeni społecznej*, 2020, nr 8(2)
107. Kaczmarek-Śliwińska M., *Specyfika zarządzania sytuacją kryzysową w przestrzeni mediów społecznościowych w perspektywie typologii SCCT W. T. Coombsa*, „Studia Medioznawcze”, nr 20(4)
108. Kaliński J., Landau Z., *Gospodarka Polski w XX wieku*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003
109. Kam Ho T., *Random Decision Forests*, *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, Montreal, QC, 14–16 August, 1995
110. Kapela E., *Językowe środki wartościowania w wywiadach z politykami* [w:] Charciaek A., Fontański H., Lubocha-Kruglik J. (red.), *Jednostki języka w systemie i mowie*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2015
111. Kaplan A. M., Haenlein M., *Users of the world, unite! The challenges and opportunities of Social Media*, „Business Horizons”, nr 53(1), 2010
112. Katz E., Lazarsfeld P. F., *Personal Influence*, Free Press, Glencoe 1955
113. Kiousis S., *Interactivity: a concept explication*, *New Media & Society*, 2002

114. Kłosińska K., *Etyczny i pragmatyczny. Polskie dyskursy polityczne po 1989 roku*, Narodowe Centrum Kultury, Warszawa 2012
115. Kocks J. N., *Political Media Relations Online as an Elite Phenomenon*, Springer, Wiesbaden, 2016
116. Kolczyński M., *Strategia komunikacyjna Konfederacji w wyborach w 2019 roku – próba rekonstrukcji*, „Studia Politologiczne”, 2020, nr 55
117. Kołodko G., *Transformacja gospodarki polskiej. Sukces czy porażka?*, BGW, Warszawa 1992
118. Konik T., *Przypadki retorycznie istotnych zmian redakcyjnych w tłumaczeniach publicystyki prasowej o tematyce gospodarczej*, „Retoryka w komunikacji specjalistycznej”, (red.) Załęska M., Polskie Towarzystwo Retoryczne, Warszawa 2015
119. J., *Od Tele-Echa do Polskiego Zoo. Ewolucja Programu TVP*, Warszawa 2008
120. Kotsiantis S. B., *Decision trees: a recent overview*, „Artificial Intelligence Review”, 2013, nr 39
121. Kowalczyk K., *Typologia polskich partii politycznych według kryterium programowego*, 10.14746/ssp, nr 1, 2014
122. Kowalski M., *Językowy obraz świata i jego istnienie*, „The Peculiarity of Man”, 2018, nr 22
123. Krotz F., *Mediatisierung: Fallstudien zum Wandel von Kommunikation*, Wiesbaden: VS Verlag für Socialwissenschaften, 2007
124. Kuczma P., Gogołek W., *Web Information Potential on the Basis of the 2010 Presidential Elections*, Wyd. Skorpion, Warszawa 2011
125. Kudra B., *O komunikacji społecznej*, [w:] *Komunikowanie publiczne. Zagadnienia wybrane*, Kudra B., Olejniczak E. (red.), Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2014
126. Labrinidis A., Jagadish H. V., *Challenges and Opportunities with Big Data*, „Proceedings of the VLDB Endowment”, 2012, nr 5(12)
127. Landis J. R., Koch G.G., *The measurement of observer agreement for categorical data*, „Biometrics”, 1977, nr 33 (1)
128. Lantz B., *Machine learning with R. 2nd ed.* Birmingham, Packt Publishing, 2015
129. Lasswell H. D., *The structure and function of communication in society*, [w:] L. Brynson (red.), *The Communication of Ideas*, Harper & Brothers, New York 1948
130. Le T. M., Liaw S.Y., *Effects of Pros and Cons of Applying Big Data Analytics to Consumers Responses in an E-Commerce Context*, „MDPI”, 2017, nr 9(5)
131. Leśniczak R., *Komunikowanie polityczne Konferencji Episkopatu Polski po zakończonych wyborach prezydenckich 2015 r. Priming i framing – analiza tygodników opinii „Newsweek Polska” i „Polityka”*, Annales UMCS, 2017, nr 24(2)
132. Leśniczak R., *Komunikowanie polityczne Konferencji Episkopatu Polski podczas kampanii parlamentarnej 2019 r. w kontekście procesów mediatyzacji*, [w:] (red.) Habrajska G., *Teorie i praktyki komunikacji*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2020
133. Levinson P., *Nowe nowe media*, WAM, Kraków 2010
134. Levinson P., *Telefon komórkowy. Jak zmienił świat najbardziej mobilny ze środków komunikacji*, Muza, Warszawa 2006
135. Lindlof T., Schatzer J., *Media ethnography in virtual space: strategies, limits and possibilities*, Journal of Broadcasting and Electronic Media, 1998, nr 42 (2)
136. Lisowska-Magdziarz M., *Feniksy, labędzie, motyle. Media i kultura transformacji*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2012
137. Liszka P., *Język jako instrument i przedmiot teologii*, Papieski Wydział Teologiczny we Wrocławiu, 2005
138. Livingstone S., Lievrouw L., *The Handbook of New Media*, London 2002
139. Lu H., Karimireddy S. P., Ponomareva N., Mirrokni V., *Accelerating Gradient*

Boosting Machine, Cornell University, 2020

140. Lundby K., *Mediatization. Concept, changes, consequences*, Peter Lang, New York 2009
141. M. Kopińska, *Język jako narzędzie interpretacji rzeczywistości – językowy obraz świata*, 2011
142. Maciejewska-Mieszkowska K., *Telewizja jako aktor polityczny w ocenie społecznej Polaków*, Media Biznes Kultura, nr 2 (9), 2020
143. Maćkiewicz J., *Co to jest językowy obraz świata*, Rozprawy i Analizy, Lublin 1999
144. Maćkiewicz J., *Kategoryzacja a językowy obraz świata* [w:] *Językowy obraz świata* (red.) Bartmiński J., Lublin 1999
145. Maćkiewicz J., *Wyspa – językowy obraz wycinka rzeczywistości*, [w:] *Językowy obraz świata*, (red.) Bartmiński J., Lublin 1990
146. Maguś W., *Twitter jako narzędzie zarządzania wizerunkiem*, "Zarządzanie mediami", 2021, nr 9(3)
147. Manheim J. B., *The news shapers: Strategic communication as a third force in news making*, [w:] Graber D., McQuail D., & Norris P. (red.), *The politics of news: The news of politics* (pp. 94-109), Congressional Quarterly Press, Washington 1998
148. Mańczyk A., *Wspólnota językowa i jej obraz świata*, Krytyczne uwagi do teorii językowej Leo Weisgerbera, WSP, Zielona Góra 1982
149. Mazzoleni G., Schultz W., *Mediatization of Politics: A Challenge for Democracy?*, *Political Communication*, 1999, nr 16
150. McAfee A., Brynjolfsson E., *Big Data: The Management Revolution*, „Harvard Business Review”, 2012, nr 90(10)
151. McCulloch W., Pitts W., *A Logical Calculus of Ideas Immanent in Nervous Activity*, „Bulletin of Mathematical Biophysics”, 1943, 5 (4)
152. M. Szpunar, *Nowe stare medium. Internet między tworzeniem nowych modeli komunikacyjnych a reprodukowaniem schematów komunikowania masowego*, Warszawa 2012
153. McKeen C. J., Livingston F., *Physical and mental measurements of the students of Columbia University*, „The Psychological Review”, 1986, nr 6
154. McNair B., *Introduction to Political Communication*, Routledge, London 1998
155. M. Ulidis, *Blog a językowy obraz świata partii politycznej. Analiza kwantytatywno-korpusowa na przykładzie komunikacji Platformy Obywatelskiej*, „Nowe media studia i rozprawy”, 2013, nr. 4, <http://dx.doi.org/10.12775/NM.2013.005>, 2013
156. McQuail D., *Teoria komunikowania masowego*, Wydawnictwo Naukowe PWN, Warszawa 2008
157. McQuail D., Windahl S., *Communication Models for the Study of Mass Communications*, Routledge, 1993
158. Media Społecznościowe, sjp.pwn.pl/slowniki/media_spolnosciowe.html, 13.09.2021
159. Meyer-Schonberger V., Cukier K., *Big Data. Rewolucja, która zmieni nasze myślenie*, MT Biznes, 2014
160. Mikuła E., *Społeczny wymiar transformacji*, Uniwersytet Rzeszowski, 2004
161. Milewski T., *Językoznawstwo*, Warszawa 1965
162. Moles A.A., *Sociodynamique de la culture*, Paris, La Haye: Mouton &C, 1967
163. Molęda-Zdziech M., *Czas celebrytów. Mediatyzacja życia publicznego*, Difin, Warszawa 2013
164. Mueller H., Rauh C., *Reading Between the Lines: Prediction of Political Violence Using Newspaper Text*, Cambridge University Press, 2017
165. Nadkarni P. M., Ohno-Machado L., Chapman W. W., *Natural language processing: an introduction*, Journal of the American Medical Informatics Association, 2011, nr 18(5)
166. Natekin A., Knoll A., *Gradient boosting machines, a tutorial*, Front Neurorobot,

167. Negrine R., *Politics and the Mass Media in Britain*, Routledge, London 1994
168. Negrine R., *The Communication of Politics*, Sage, London 1996
169. Nickerson D. W., Rogers T., *Campaigns and Big Data*, „HKS Faculty Research Working Paper Series”, 2013
170. Nie K. S., Keea C. P., Ahmad A. L., *Mediatization: A Grand Concept or Contemporary Approach?*, „Procedia - Social and Behavioral Sciences”, 2014, nr 155
171. Niemelä K., *Religion in the Nordic media: introduction*. In *Nordic Journal of Religion and Society*, 2013, nr 26(1)
172. Niwattanakul S., Singthongchai J., Naenudornand E., Wanapu S., *Using of Jaccard Coefficient for Keywords Similarity*, *Proceedings of the International MultiConference of Engineers and Computer Scientists*, 2013
173. Nyzio A., *W słabości siła. Bilans ewolucji polskich partii politycznych i w latach 1989-2013*, „Polityka wewnętrzna i bezpieczeństwo”, 2014, nr 2
174. Oniszczyk Z., *Mediatyzacja polityki i polityzacja mediów. Dwa wymiary wzajemnych relacji*, „Studia Medioznawcze”, 2011, nr 4 (47)
175. Ożóg K., *Język współczesnej polskiej polityki*, „Stylistyka”, 2008, nr 17
176. Ożóg K., *O języku współczesnej polityki*, „Polityka i Społeczeństwo”, 2007, nr 4
177. Ożóg K., *Uwagi o języku polskiej polityki po roku 1989*, [w:] *Przekazy polityki*, (red.) Siewierska-Chmaj A., Konsorcjum Akademickie, Kraków–Rzeszów–Zamość 2009
178. Palczewski M., *Umberto Eco o mediach – w esejach, rozprawach i w powieści*, „Media, Kultura, Społeczeństwo”, 2014, nr 9-10
179. Pawelec R., Trysińska M., *Najnowsze słownictwo a współczesne media elektroniczne*, „Poznańskie Studia Polonistyczne Seria Językoznawcza”, Warszawa 2008, nr 17(37)
180. Peled A., *The Politics of Big Data A Three-Level Analysis*, European Consortium of Political Research (ECPR) General Conference, 2013
181. Perloff R. M., *Political Communication, Politics, Press and Public in America*, LEA, London 1997
182. Peszyński W., *Personalizacja politycznych preferencji*, „Preferencje Polityczne. Postawy – identyfikacje – zachowania”, 2012, nr 3
183. Pfetsch B., *Comparing Political Communication*, ACADEMIA, 2012
184. Piasecki A. K., *Wybory. Parlamentarne, samorządowe, prezydenckie 1989–2002*, Wydawnictwo Adam Marszałek, Toruń 2004
185. Pisarek W., *Słownik terminologii medialnej*, Towarzystwo autorów i wydawców prac naukowych Universitas, Kraków 2006
186. Pisarek W., *Wstęp do nauki o komunikowaniu*, Wydawnictwa akademickie i profesjonalne, 2008
187. Podemski K., *Badania polskiego dyskursu publicznego II RP, PRL i III RP Przegląd zagadnień*, „Kultura i Społeczeństwo: Polska Akademia Nauk Komitet Socjologii Instytut Studiów Politycznych”, 2013, nr 2
188. Predictive modeling, www.gartner.com/en/information-technology/glossary/predictive-modeling, 13.09.2021
189. Puzynina J., *Problemy wartościowania w języku i w tekście*, „Etnolingwistyka”, 2004, nr 16
190. Pye L.W., *Political communication*, [w:] Bogdanor V. (red.), *Tire Blackwell encyclopedia of political science* (2nd edition, pp. 442-445), “Cambridge, MA: Blackwell”, 2018, nr 68(2)
191. Quinlan J. R., *Simplifying decision trees*, „International Journal of Man-Machine Studies”, 1987, nr 27 (3)
192. R. Schumaker, H. Chen, *Textual Analysis of Stock Market Prediction Using*

- Financial News Articles*, „Americas Conference on Information Systems (AMCIS)”, 2006, nr 27(2)
193. Raszkowski A., *Przesłanki promocji miast i regionów*. „Zeszyty Naukowe Uniwersytetu Szczecińskiego. Ekonomiczne Problemy Usług”, 2010, nr 56
 194. Rheingold H., *The Virtual Community*, London, Seeker and Warburg, 1994
 195. Riley J. W., Riley M. W., *Mass Communication and the Social System*, [w:] Merton R. K., Broom L., Cotrell S. (red.), *Sociology Today*, Basic Books, New York 1959
 196. Riley J. W., Riley M.W., *Mass Communication and the Social System*, [w:] Merton R.K., Broom L., Cotrell S.(red.), *Sociology Today*, Basic Books, New York 1959
 197. Rozario R. B., *New Media and the Traditional Media Platforms: Introspection on the Differences in Technical and Ideological Factors and Audience-integration Patterns between New Media and Traditional Media*, doi.org/10.12724/ajss.26.3, 2013
 198. Rulka M., *Proporcjonalne formuły wyborcze i ich konsekwencje*, „Studia PRAWNICZE”, 2010, nr 2(184)
 199. Rush M., *Politics and Society*, An Introduction to Political Sociology, Prentice Hall, London 1992
 200. Sak H.,Senior A., Beaufays F., *Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling*, Google, USA, 2014
 201. Sapienza Z. S., Iyer N., Veenstra A. S., *Reading Lasswell's Model of Communication*, „Mass Communication and Society”, 2015, nr 18(5)
 202. Sartori G., *Homo videns. Telewizja i post-myślenie*, przeł. J. Uszyński, Telewizja Polska, Warszawa 2005
 203. Schmidhuber J., *Deep Learning in Neural Networks: An Overview*, „Neural Networks”, 2015, nr 61
 204. Schramm W., *How Communication Works*, [w:] Schramm W. (red.), *The Process and Effects of Mass Communication*, University of Illinois Press, Urbana 1954
 205. Schulz W., *Political communication*, [w:] Donsbach W. (red.), *The international encyclopedia of communication* (pp. 3671 -3682), MA: Blackwell, Cambridge 2008
 206. Schulz W., *Reconstructing Mediatization as an Analytical Concept*, „European Journal of Communication”, 2004, nr 19(1)
 207. Schwarz J., *Detecting political ideology in Youtube comments using machine learning*, „Seminar Social Media and Business Analytics”, 2009
 208. Sieradzka-Baziur B., *Język polityki - dotychczasowe prace i perspektywy badawcze*, „Horyzonty Polityki”,2011, nr 2(2)
 209. Skorowski H., *Polskie przemiany w perspektywie politologiczno-socjologicznej*, „MAZOWSZE Studia Regionalne”, 2014, nr 15.
 210. Sneath P. H. A., Sokal R.R., *Numerical Taxonomy: The Principles and Practice of Numerical Classification*, Freeman, San Francisco 1973
 211. Sojda S., *Słowacka i polska politolingwistyka*, [w:] *Z warsztatu współczesnego słowacysty. Studia słowacystyczne dedykowane pamięci profesor Marii Honowskiej*, (red.) Mieczkowska H., Hudymač A., Babik Z., Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2010
 212. Strömbäck J., *Four Phases of Mediatization: An Analysis of the Mediatization of Politics*, „The International Journal of Press/Politics”, 2008, nr 13(3)
 213. Szewczyk M., *Linguistic discrimination against the Roma in the polish-language web 2.0 discourse*, „ROCZNIKI NAUK SPOŁECZNYCH”, 2022, nr 50(2)
 214. Szpunar M., *Nowe stare medium. Internet między tworzeniem nowych modeli komunikacyjnych a reprodukowaniem schematów komunikowania masowego*, Warszawa 2012
 215. Tabakow M., Korczak J., Franczyk B., *Big data – definicje, wyzwania i technologie informatyczne*, Uniwersytet Ekonomiczny we Wrocławiu, Wrocław 2014
 216. Thompson J. B., *Ideology and Modern Culture*, Polity Press, Cambridge 1990

217. Thompson J. B., *The Media and Modernity: A Social Theory of the Media*, Polity Press, Cambridge 1995
218. TikTok User Statistics 2022, backlinko.com, 20.11.2022
219. W. Gogołek, i inni, *Z badań nad wykorzystaniem rafinacja informacji sieciowej. Wybory prezydenckie i parlamentarne 2015*, „Studia Medioznawcze”, Instytut Dziennikarstwa UW, Warszawa, 2015, nr 3(62), s. 31-43, http://studiamedioznawcze.pl/article.php?date=2015_3_62&content=gogolek&lang=pl, 10.08.2021
220. Trzcińska J., *Women and Women's Rights Narratives on Polish Twitter in Light of the Abortion Issue*, *Global Media Journal*, 2021, nr 11(2)
221. Ustawa o wykonywaniu mandatu posła i senatora, sejm.gov.pl, 19.11.2022
222. Välvirronen E., *From Mediation to Mediatization: The New Politics of Communicating Science and Biotechnology*, [w:] U. Kivikuru, T. Savolainen (red.) *The Politics of Public Issues*. Helsinki, Department of Communication, University of Helsinki, 2001
223. van Dijk J., *Models of Democracy and Concepts of Communication*, [w:] Hacker K., van Dijk J. (red.) *Digital Democracy: Issues of Theory and Practice*, Sage, London – Thousand Oaks – New Delhi 2004
224. van Dijk J., *Společne aspekty nowych mediów*, Wydawnictwo Naukowe PWN, Warszawa 2010
225. Walczak B., *Co to jest język polityki?*, 2017
226. Walker J., *Old Media and New Media: Like It or Not, They're Partners*, Reason, September 15, 2004
227. Ward linkage, www.datanovia.com, 28.12.2021
228. Weingart P., *Science and the Media*, „Research Policy”, 1998, nr 27(8)
229. Wełna Ź., *Polityk aktywny – polityk interaktywny. Obecność liderów politycznych w mediach społecznościowych na przykładzie polskiej kampanii parlamentarnej 2019*, „MEDIA I SPOŁECZEŃSTWO”, 2020, nr 12
230. Westley B. H., MacLean M., *A Conceptual Model for Mass Communication Research*, „Journalism Communication”, 1957, nr 34
231. Williams R., *Keywords*, London 1976
232. Worsowicz M., *Jak pisać do Internetu? Retoryczne aspekty poradnictwa warsztatowego*, „Retoryka w komunikacji specjalistycznej” (red.) Załęska M., Polskie Towarzystwo Retoryczne, 2015
233. Zablith F., Osman I. H., *ReviewModus: Text classification and sentiment prediction of unstructured reviews using a hybrid combination of machine learning and evaluation models*, „Applied Mathematical Modelling”, 2019, nr 71
234. Zamecki Ł., *Společne podstawy ładu politycznego*, Wydział Dziennikarstwa i Nauk Politycznych Uniwersytet Warszawski, 2011
235. Zhang Z., *Introduction to machine learning: k-nearest neighbors*, „Ann Transl Med”, 2016, nr 4(11)
236. Zhang Z., *Too much covariates in a multivariable model may cause the problem of overfitting*, *J Thorac Dis*, 2014 Sep, nr 6(9)
237. Żuk G., *Językowy obraz świata w polskiej lingwistyce przełomu wieków*, [w:] *Przeobrażenia w języku i komunikacji medialnej na przełomie XX i XXI wieku*, red. Karwatowska M., Siwiec A., Chełm 2010
238. Żurawski J., *Internet jako współczesny środek elektronicznej komunikacji wyborczej*, 2010

7. Streszczenie

Streszczenie

Badanie zmediatyzowanej komunikacji politycznej przy pomocy analizy statystycznej i algorytmów predykcyjnych sztucznej inteligencji

Głównym celem pracy było odpowiedź na pytanie: czy, w jaki sposób i w jakim stopniu analiza treści w komunikowaniu politycznym i narzędzia sztucznej inteligencji mogą służyć do budowy modeli predykcyjnych dopasowania do grupy politycznej.

Przeprowadzono szereg analiz statystycznych na danych politycznych pokazujących, że analiza treści komunikowania politycznego niesie ze sobą niezwykle korzyści w postaci zrozumienia kwestii w poruszanych treściach, ich ilości, powiązań, nacechowania, emocji i znalezieniu wzorców, które są charakterystyczne dla każdej z grup komunikujących się.

Zbudowano istotne statystycznie modele sztucznej inteligencji przewidujące poglądy polityczne na podstawie wypowiedzi i spełniające wyżej wymienione założenia ujęte w celu pracy z wydajnością prezentującą zadowalający poziom.

Study of mediatized political communication with the help of statistical analysis and artificial intelligence prediction algorithms

Summary

The main aim of the study was to answer the question whether, how and to what extent content analysis in political communication and artificial intelligence tools can be used to build predictive models of matching to a political group.

A series of statistical analyses on political data have been carried out showing that the aspect of analyzing the content of political communication has remarkable benefits in terms of understanding the issues in the discussed content, their quantity, relations, characterizations, emotions and finding patterns that are specific to each of the groups.

Statistically significant artificial intelligence models were built to meet predicting political views. They were based on the statements and fulfilling the assumptions included in order to work with an efficiency presenting a satisfactory level.

Rzeszów, dnia 27.02.2023 r.

Arkadiusz Gawel

OŚWIADCZENIE

Oświadczam, że złożona przeze mnie rozprawa doktorska „Badanie zmediatyzowanej komunikacji politycznej przy pomocy analizy statystycznej i algorytmów predykcyjnych sztucznej inteligencji” została napisana przeze mnie samodzielnie.

Oświadczam również, że przedstawiona rozprawa doktorska nie była wcześniej przedmiotem procedur związanych z uzyskaniem stopnia naukowego.

Jednocześnie wyrażam zgodę na jej publiczne udostępnianie.

.....

podpis