

# Analysing Policy Alignment through Structural Topic Modelling: Ministerial Speeches and the Governance of the Sixth Labour Government of Aotearoa/New Zealand (2017-2023)<sup>1</sup>

Maiia Skrypnyk

March 8, 2024

1

There are very few things I aspired to do in politics that have a natural end point – poverty, inequality, ending environmental degradation – if you ever claim it’s job done on those issues, you set the bar too low.

---

— *Jacinda Ardern in her valedictory statement on April 5, 2023 (New Zealand Parliament, 2023)*

## 1 Context, Literature Review & Research Question

*‘New Zealand’s Election Had Been Predictable. Then ‘Jacindamania’ Hit’.* This is how the New York Times (Kwai, 2017), alongside many other media outlets worldwide (the Guardian, CNN, Bloomberg) (Shuttleworth, 2017; Griffiths, 2017; Withers, 2017), reacted to Jacinda Ardern becoming the youngest leader of the Labour Party of New Zealand (NZ) in August 2017. After this, the Labour Party had experienced an immediate leap in popular support, resulting in the number of their parliamentary seats being increased from 32 to 46 after the general elections in September 2017 (Electoral Commission, 2014, 2017). While the Labour Party still had had ten less seats than the sitting centre-right NZ National Party, in October 2017, the NZ First Party (9 seats) had reached a coalition agreement with the Labour Party (Coalition Agreement, 2017) and a confidence-and-supply agreement with the Green Party (Confidence and Supply Agreement, 2017). Thus, a minority coalition government was formed, with the overwhelming majority of ministers being Labour representatives, and Jacinda Ardern became the Prime Minister (Fletcher, 2018, p. 2).

---

<sup>1</sup>GitHub folder: <https://github.com/skrypnykm/QTANewZealandProject>

In October 2020, in the next general elections, the governing Labour Party achieved a landslide victory with 48% support, resulting in 19 additional seats in parliament (Electoral Commission, 2020), giving the Party a right to govern independently without a need for coalition. In January 2023, Jacinda Ardern stepped down from the prime ministership and was succeeded by the Labour Party leader Chris Hipkins (Radio New Zealand, 2023). In November 2023, ‘the Labour era’ had come to an end, with a centre-to-right conservative coalition of the National Party, NZ First and ACT Party forming the new government (Quinlivan, 2023).

Therefore, the Sixth Labour Government of New Zealand was incumbent for six years (2017-2023), and it is broadly known for its progressive social & civil rights, welfare, healthcare, labour, education policy agendas, focus on climate action and solving the housing crisis, as well as the highly appraised efficiency in handling the COVID-19 pandemic (Cooke, 2023). These issues, included but not limited to, have formed the pillars of the NZ Labour Party Policy Platform throughout the years. Budge (1993, p. 699, p. 710) believes that ‘no other party document is quite so comprehensive as the electoral programs’, and, supported by Budge and Laver (1993), Budge and Hofferbert (1990), argues that ‘...Parties’ policy commitments to their ideologically based priorities means that they are motivated to get the policies through [...]. They cannot abandon their priorities, because leaders identify and are identified with them; party bodies and activists support them; associated groups back them; and party voters expect them’. Ginsberg’s (1976) research has shown that ‘platform commitments *do* get carried through into general government action’.

This study aims to diverge from the more traditional – to political scientists (e.g., the Manifesto Research Group (1979-1989), Comparative Manifestos Project (CMP 1989-2009), the Manifesto Research on Political Representation (since 2009) (Manifesto Project, 2024) – analysis of the parties’ manifestos/programs/platforms themselves. The researcher’s goal is to computationally analyse a corpus of 850+ speeches delivered by the ministers of the Sixth Labour Government of NZ from 2017 to 2023 to understand **how and to what extent the Party priorities, as stated in the Labour Party Policy Platforms for the time period in question (published in 2017, 2019, and 2022), are addressed by the government representatives in their official speeches.** Such priorities are indicated as follows, with minor changes in wording in the 2022 edition:

- ***‘Labour’s values’*** (equality, opportunity, solidarity, sustainability, freedom);
- ***‘Tāngata Whenua’*** (respect for and support of the indigenous Māori population);
- ***‘Strengthening the economy’***;
- ***‘A Healthy Environment’***;
- ***‘Opportunity and fairness for all’*** (women, families, housing);
- ***‘A world-class education for all’***;
- ***‘Health — wellbeing, access, and fairer outcomes’***;
- ***‘Justice, human rights, and equality for all’***;

- *‘New Zealand’s identity and culture’* (arts, culture, heritage)
- *‘New Zealand’s place in a changing world’* (peacebuilding, sustainable development, relationship with the Pacific islands, defence, trade);
- *‘Effective and accountable government’*.

The author contemplates that there could be a significant alignment between the mentioned priorities and the topics mentioned in the speeches, indicating consistency and policy focus. Furthermore, the study aspires to examine the changes of given topics’ prevalence throughout the Government’s time in power. While exploring the temporal trends, it is important to note that this Government’s tenure had been influenced by several unexpected and thus pivotal tragic events in the history of New Zealand, such as the Christchurch mosque shootings and Whakaari/White Island volcano eruption, both 2019, and the world-scale COVID-19 pandemic, which might have influenced the congruence of Labour policies.

## 2 Data Acquisition & Corpus Description

The official website of the New Zealand Government, *beehive.govt.nz*, provides access to the ministerial speeches and diaries, as well as transcripts of governmental features (e.g., press conferences) and press releases. This information is available for 11 government compositions incumbent during the period 1993-2024. The corpus under analysis comprises all speeches delivered by the ministers of the Sixth Labour Government of New Zealand. On the Beehive website, these speeches could be found under two separate categories: “Labour-led government (2017-2020)” and “Labour (2020-2023)”. 465 speeches delivered from November 8, 2017, until November 5, 2020 fall under the former category, and 390 speeches delivered from November 11, 2020, until October 5, 2023, fall under the latter. Overall, 855 speeches were sourced from the above-mentioned website through web-scraping with the help of R.

**Rvest** was used as a key R package to scrape the data. Transcripts of speeches alongside the metadata like hyperlinks and dates were extracted by iterating through multiple (web)pages, combined and exported as a CSV file, which is available on GitHub together with the replication code. The corpus was pre-processed by removing numbers, punctuation marks, symbols, and stopwords. Collocations of two words were identified based on a minimum count of 10 times (frequency of occurring in the corpus) and a Z-score of 14. Then a Document Feature Matrix (DFM) was created.

Some of the corpus’ **textual statistics** are as follows: mean number of sentences = 79.91; mean number of types = 639.23; mean number of tokens = 1856. Therefore, the mean Type-Token Ratio (TTR), which is the ratio of unique words (types) to total words (tokens) in a text, equals 0.344, or approximately 34.4%. A lower TTR, like in our case, suggests more repetition of words within the text (Thomas, 2005). Mean lexical density (‘content-carrying’ words to total words) equals 0.68, and the mean Flesch-Kincaid Grade Level of 12.5 suggests that the texts are quite complex and expected to be understandable by an adult (12th grade student in the United States of America) (To et al., 2013). More information regarding the corpus’ keyness and feature frequencies of the words used would be provided in the ‘Summary results’ section below.

### 3 Research Methods

Together with the general textual statistics methods, **Structural Topic Modelling** was chosen as a primary method to analyse the aforementioned corpus.

In general, topic modelling is a suite of algorithms that aim to discover and annotate large archives of documents with the thematic information. Such algorithms are statistical methods that analyse the words of the original texts to discover the themes that run through them, how those themes are connected to each other, and how they change over time (Blei, 2012, p. 77).

Two main approaches for computational text analysis (based on the topic modelling technique) might be identified: a more ‘traditional’ approach called Latent Dirichlet Allocation (LDA), created in the early 2000s, and a more recently (mid-2010s) developed Structured Topical Modelling (STM) (Roberts et al., 2014). LDA is largely considered to be a ‘static’ or a ‘fixed-topic’ model, since topics are assumed to be constant across the entire text corpus (no inherent temporal component in the model – all the documents are believed to be generated in the same ‘time steps’) (del Gobbo et al., 2020, p. 672). Moreover, in the standard formulation of LDA, topics within a document are independent of one another (if Topic 1 is identified, the model will not give any information whether the same document has Topic 2 or 3), and the topics are modelled based entirely on the text of the document, excluding any external metadata (Blei et al., 2003).

As this research, among its other objectives, seeks to examine the topics’ prevalence evolution throughout a certain time period, the author considered that using a modernised STM method might be more beneficial. STM generally builds up on the LDA algorithms, though while in LDA the topic prevalence and content come from Dirichlet distributions with hyperparameters set in advance, in STM, they come from document metadata, e.g. date, time, and geographical location of production (Lebryk, 2021). Also, another advantage of STM is that it assumes the presence of multiple topics in a single document (Zhao et al., 2011, cited by Negara, 2019). STM tools are components of the `stm` R package.

### 4 Summary Results

First of all, some additional general textual statistics:

**Number of speeches per year** (Figure 1): peaking in 2018 at 197, with a minimum of 22 in 2017 (as the Government was only sworn in November); much fewer speeches (72) during the pandemic in 2020.

*Please see the next page for Figure 1*

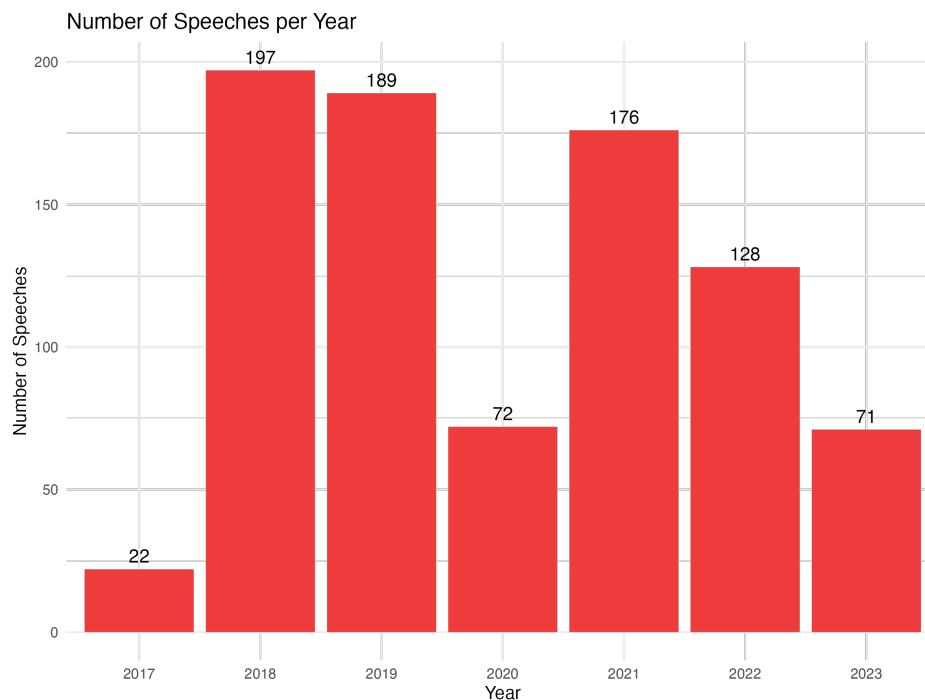


Figure 1: Number of Speeches per Year

**Feature frequency of the words in all speeches throughout the years** (Figure 2-3): while the frequency analysis does not give us substantially profound insights, it is still possible to distinguish that terms like 'economy', 'support', 'Māori', 'Pacific' are featured consistently throughout the whole tenure. It is also not surprising to see 'virus', 'response', 'work' and 'business' prevalent in 2020.

*Please see the next page for Figures 2-3*

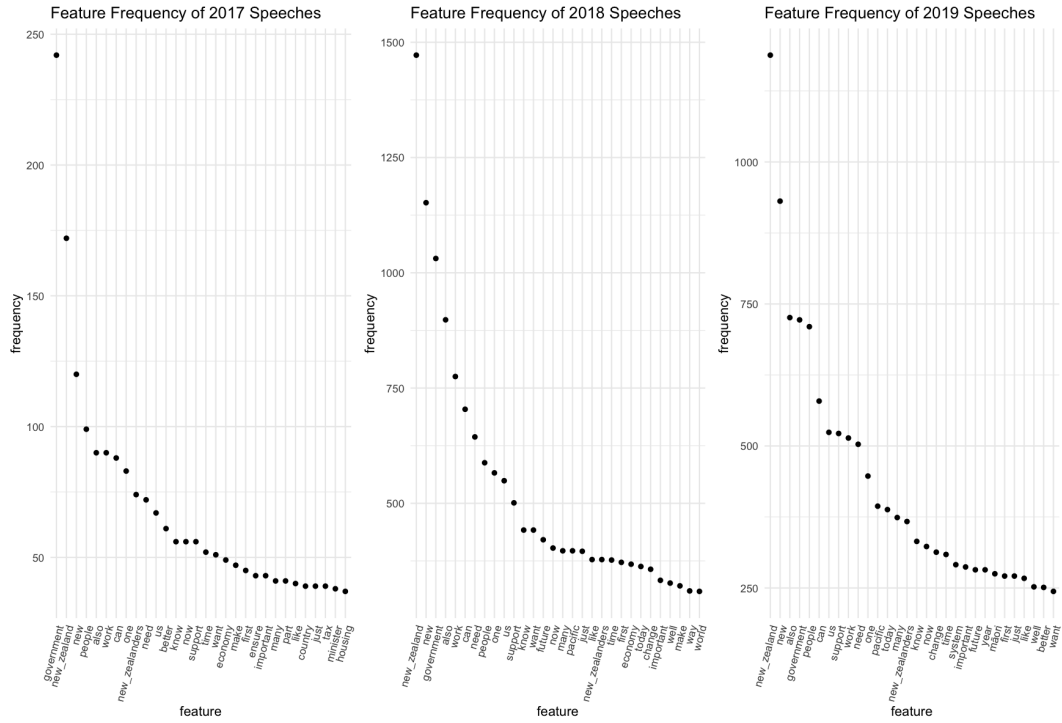


Figure 2: Feature Frequency of 2017-2019 Speeches

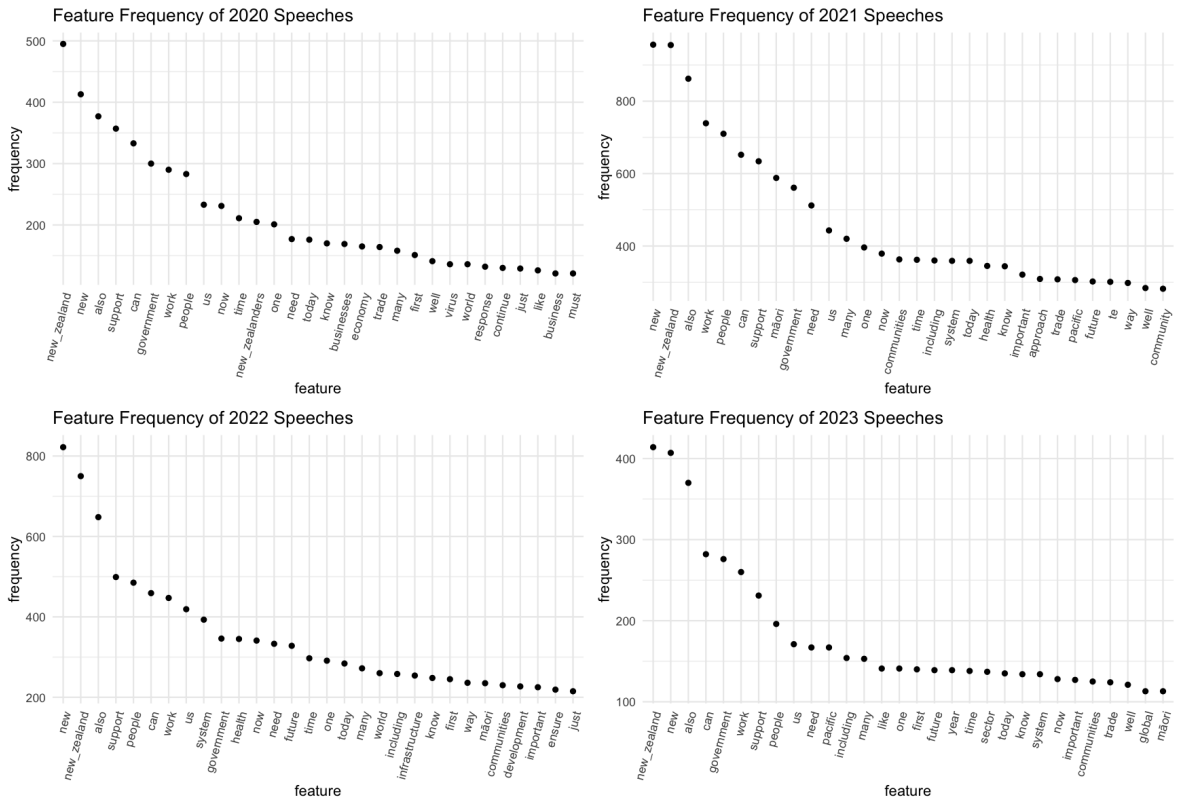


Figure 3: Feature Frequency of 2020-2023 Speeches

More insights on temporal developments of the speeches main themes could be derived from **relative frequency analysis (keyness)** (Figures 4-6). 'The purpose of a keyness analysis is to identify (lexical) items that are typical of a particular text variety' (Sönning, 2022) in comparison to other texts in the corpus.



Figure 4: Keyness for 2017-2019 speeches

It is important to notice how the tragic events of 2019, a volcano eruption and two terrorist attacks, are clearly displayed in 2019 speeches as compared to 2018 with its more peaceful agenda ('attack', 'victims', 'response', 'firearms' - just a month after the Christchurch attacks, New Zealand banned semi-automatic weapons and assault rifles nationwide (Kouchakji, 2019).

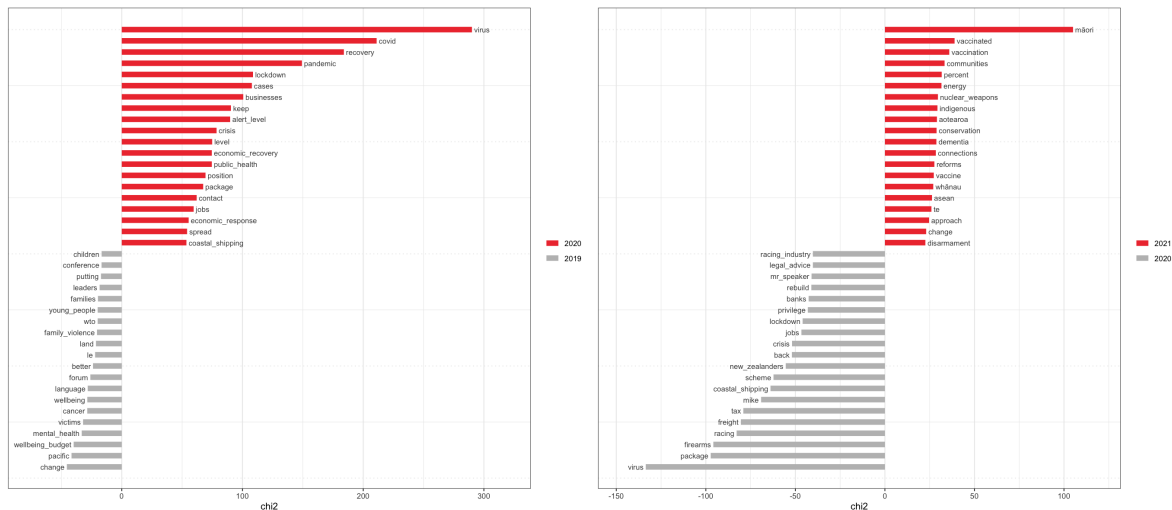


Figure 5: Keyness for 2020-2021 speeches

Clear domination of pandemic-related terms in 2020 and into 2021 ('vaccination'), but also 'nuclear weapons' - the Treaty on the Prohibition of Nuclear Weapons (TPNW)

entered in force in January 2021 (New Zealand Ministry of Foreign Affairs and Trade, 2021), as well as 'Māori', which may be connected to Nanaia Mahuta becoming the first Māori Minister of Foreign Affairs of NZ in 2020.

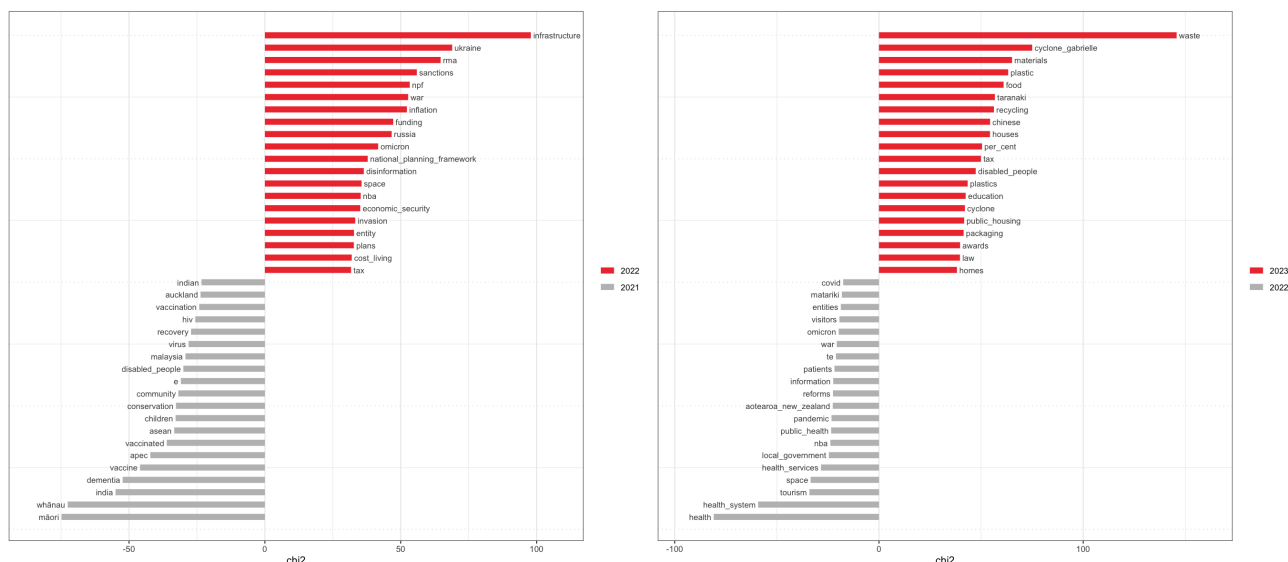


Figure 6: Keyness for 2022-2023 speeches

'Infrastructure' is the leader for 2022 (as the new thirty-year Infrastructure Strategy was adopted (New Zealand Infrastructure Commission, 2022), alongside the Russian invasion of Ukraine, 'rma' (replacement of the Resource Management Act 1991 (The Conversation, 2023), 'inflation', and 'omicron'.

In 2023, the Government focused on the environmental issues ('waste', 'plastic', 'recycling', opposition to mine the 'materials' from the 'Taranaki' island seabed (Stuff, 2023)), devastating Cyclone Gabrielle, and housing.

However, the primary analysis of this research lies within the realm of **Structural Topic Modelling**. The results of fitting an STM with  $k = 30$  with temporal metadata included are as follows (Figures 7-10):

*Please see the next page for Figures 7-10*



### Topic prevalence and top terms (FREX)

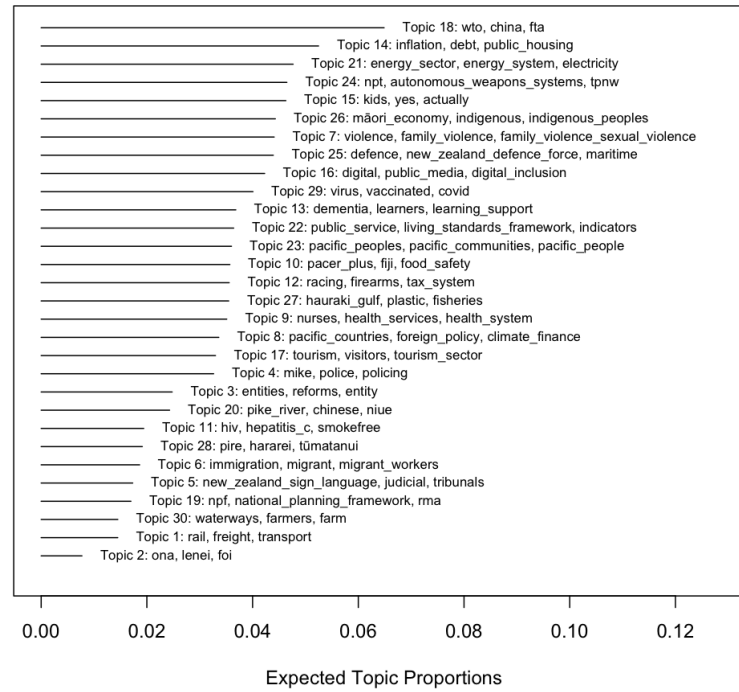


Figure 7: Topic prevalence and top terms (FREX)

### Topic prevalence and top terms (Highest Probability)

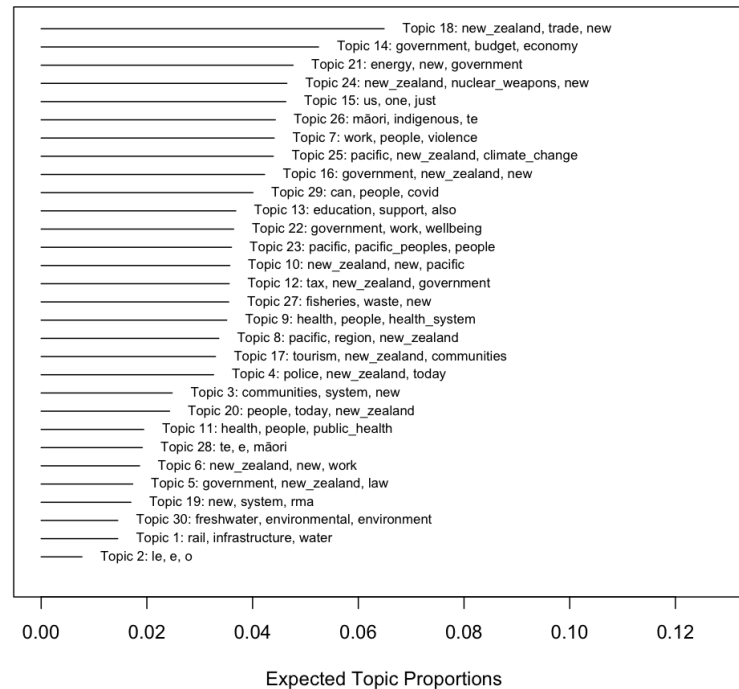


Figure 8: Topic prevalence and top terms (Highest Probability)

### Topic prevalence and top terms (Lift)

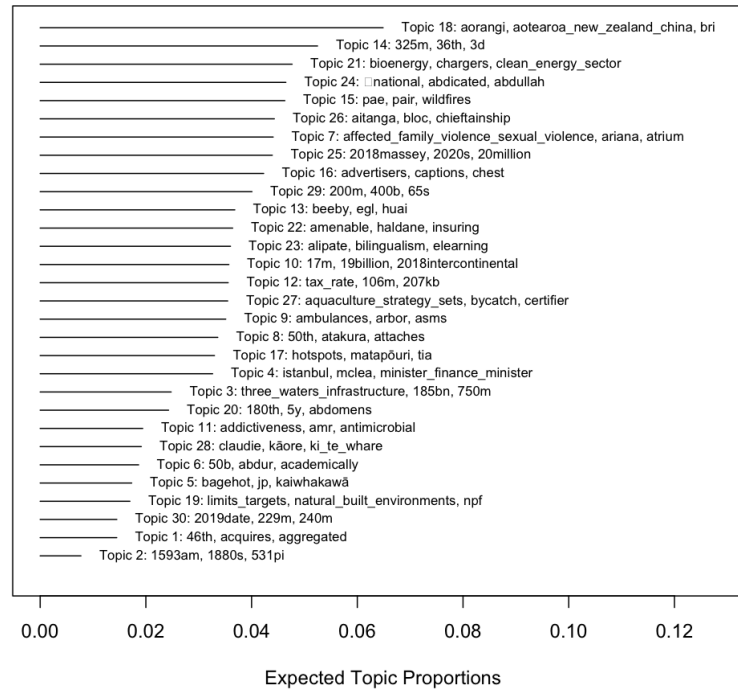


Figure 9: Topic prevalence and top terms (Lift)

### Topic prevalence and top terms (Score)

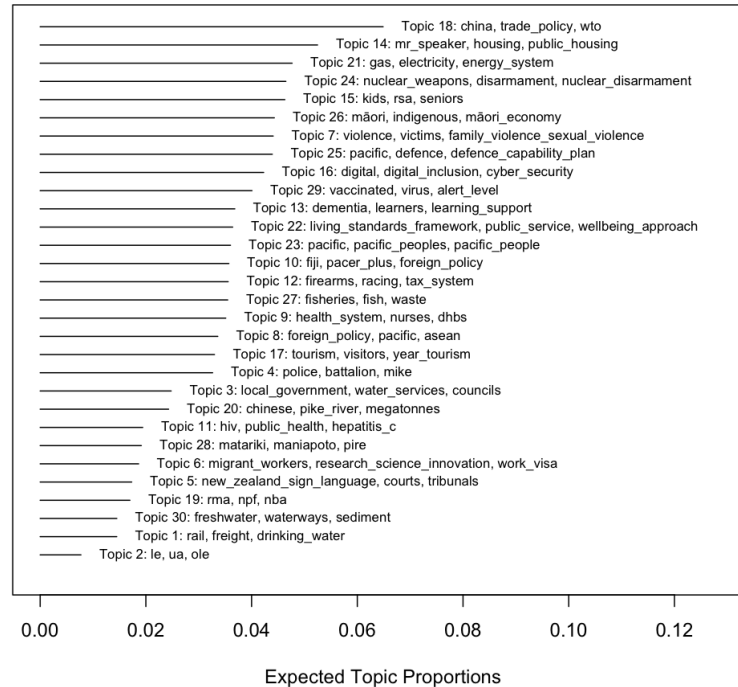


Figure 10: Topic prevalence and top terms (Score)

The results seem coherent by first sight. After checking several initial texts within each topic using the `findThoughts` function, the coherency of the majority of the modelled topics can be confirmed, and they could be labelled as (in ascending order as according to FREX):

- *Topic 18*: NZ-China trade relationship.
- *Topic 14*: Economy and Housing.
- *Topic 21*: Energy system.
- *Topic 24*: Nuclear weapons.
- *Topic 15, 13, 3, 20*: no clear association.
- *Topic 26*: Māori population.
- *Topic 7*: Family and sexual violence.
- *Topic 25*: NZ defence policy in the Pacific.
- *Topic 16*: Digital inclusion and Cyber security.
- *Topic 29*: COVID-19.
- *Topic 22*: Well-being and Living standards.
- *Topic 23, 10, 8*: NZ foreign policy regarding the states of the Pacific.
- *Topic 12*: Taxation.
- *Topic 27*: Fisheries and Environment.
- *Topic 9, 11*: Public Health.
- *Topic 17*: Tourism.
- *Topic 4*: Police.
- *Topic 28, 2*: Māori language words used in the speeches.
- *Topic 6*: Immigration.
- *Topic 5*: Law and Justice.
- *Topic 19*: Replacement of the Resource Management Act.
- *Topic 30*: Fresh drinking water and Environment.

The word clouds for the top-6 most prevalent topics are illustrated in the Figure 11:



Figure 11: Word clouds for Topic 18, 14, 21, 24, 26, 7

As it had been mentioned previously, the major advancement (and advantage) of the STM method is its ability to explore the prevalence of relevant topics over the time period of interest. Furthermore, it is also a chance for the Model to be tested for its predictive validity: it would be expected that certain topics would be more prevalent during relevant periods. For example, COVID-19 topic should peak in 2020, which is proved by the plot:

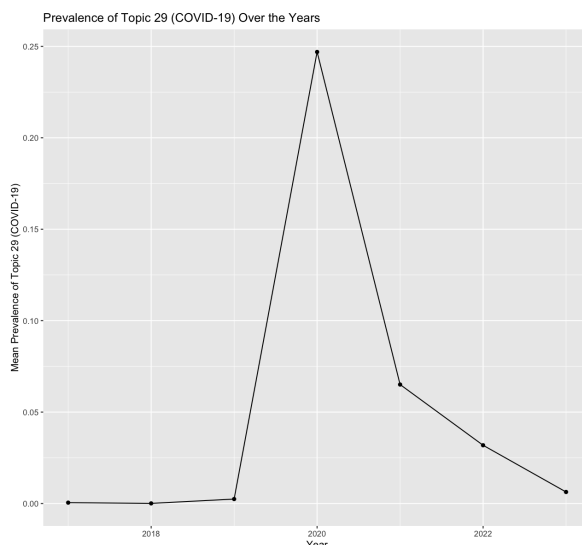


Figure 12: Prevalence of Topic 29 (COVID-19) over the years

It might be recalled that the keyness plot for 2021 showed a significant rise in frequency of usage of the term 'Māori'. This is also confirmed by the plot and thus supports the validity of the Model:

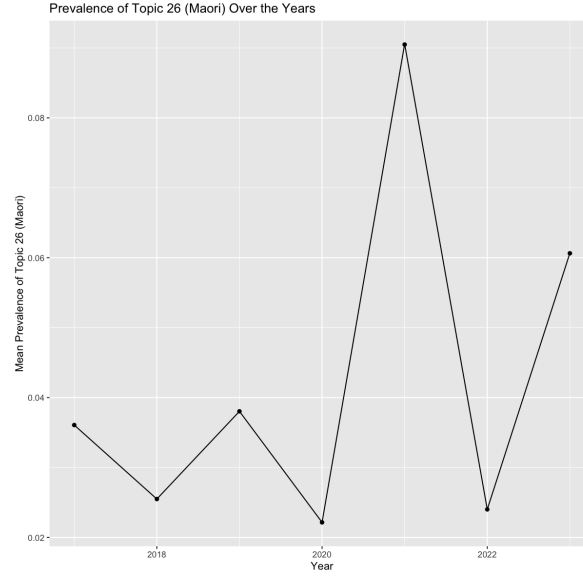


Figure 13: Prevalence of Topic 26 (Maori) over the years

## 5 Conclusion

Key findings of the corpus analysis support the author's assumption that the priorities stated in the Labour Party Policy Platform would constitute prevalent topics of the Government ministers. Among the 11 priorities, all but two are present in the Model: 'A world-class education for all' and 'Effective and accountable government' are the only priorities not explicitly represented as separate topics (however, the latter may underpin other topics, as 'Labour values' does). It is important to note that, quite insightfully, the most prevalent topic came to be NZ-China trade relationship, which alongside five other topics concerning nuclear weapons, defence strategy, and NZ foreign policy towards the Pacific islands, correspond to the priority 'New Zealand's place in a changing world'. Undoubtedly, COVID-19 pandemic and healthcare-related issues were vastly reflected in the speeches (and constitute two separate topics), while the terrorist attacks and environmental disasters were only distinguished through frequency analysis in the certain year, but they were not prevalent enough themes in the speeches to be modelled by STM throughout the whole period.

In conclusion, this research reveals the Sixth Labour Government's adherence to the Labour Party political pillars and its responsiveness to unexpected challenges in the suitable manner. This study enhances better understanding of the internal and external political dynamics of Aotearoa/New Zealand in 2017-2023, and contributes to the academic discourse both within political science and data science as a unique quantitative text analysis of New Zealand's highest officials' speeches for the given period.

## 6 Bibliography

1. Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84. <https://doi.org/10.1145/2133806.2133826>
2. Budge, I. (1993). Parties, Programs and Policies: A Comparative and Theoretical Perspective. *The American Review of Politics*, 14, 695–716. <https://doi.org/10.15763/issn.2374-7781.1993.14.0.695-716>
3. Budge, I., Hofferbert, R. (1990). Mandates and Policy Outputs: US Party Platforms and Federal Expenditures. *American Political Science Review*, 84, 111–31.
4. Budge, I., Laver, M.J. (1992). Coalition Theory and Government Policy. *British Journal of Political Science*, 23, 310–336.
5. Coalition Agreement (2017). New Zealand Labour Party & New Zealand First. 52nd Parliament. <https://www.mcguinnessinstitute.org/wp-content/uploads/2021/08/NZ-Labour-Party-NZ-First-2017.pdf>
6. Confidence and Supply Agreement. New Zealand Labour Party & Green Party of Aotearoa New Zealand. 52nd Parliament. [https://www.parliament.nz/media/4487/nzlp\\_\\_\\_gp\\_c\\_s\\_agreement.pdf](https://www.parliament.nz/media/4487/nzlp___gp_c_s_agreement.pdf)
7. Cooke, H. (2023, October 18). ‘How much of Labour’s legacy will survive? A comprehensive analysis’. *The Spinoff*. <https://thespinoff.co.nz/politics/18-10-2023/how-much-of-labours-legacy-will-survive-a-comprehensive-analysis>
8. Del Gobbo, E., Fontanella, S., Sarra, A., & Fontanella, L. (2021). Emerging Topics in Brexit Debate on Twitter Around the Deadlines: A Probabilistic Topic Modelling Approach. *Social Indicators Research*, 156(2–3), 669–688. <https://doi.org/10.1007/s11205-020-02442-4>
9. Electoral Commission (2014). 2014 General Election – Official Result. [https://www.electionresults.govt.nz/electionresults\\_2014/](https://www.electionresults.govt.nz/electionresults_2014/)
10. Electoral Commission (2017). 2017 General Election - Official Result. [https://www.electionresults.govt.nz/electionresults\\_2017/](https://www.electionresults.govt.nz/electionresults_2017/)
11. Electoral Commission (2020). 2020 General Election – Official Result. [https://www.electionresults.govt.nz/electionresults\\_2020/](https://www.electionresults.govt.nz/electionresults_2020/)
12. ‘Chris Hipkins sworn in as prime minister’ (2023, January 25). *Radio New Zealand*. <https://www.rnz.co.nz/news/political/483045/chris-hipkins-sworn-in-as-pri>
13. Fletcher, M. (2018). Developments in Social Legislation and Policy in New Zealand: ‘Jacindamania’ and a New Government. *Social Law Reports*, 2(2018). Max Plank Institute for Social Law and Social Policy. [https://www.mpisoc.mpg.de/fileadmin/user\\_upload/data/Sozialrecht/Publikationen/Schriftenreihen/Social\\_Law\\_Reports/SLR\\_No.2\\_2018\\_NZ\\_\\_final\\_.pdf](https://www.mpisoc.mpg.de/fileadmin/user_upload/data/Sozialrecht/Publikationen/Schriftenreihen/Social_Law_Reports/SLR_No.2_2018_NZ__final_.pdf)
14. Ginsberg, B. (1976). Elections and Public Policy. *American Political Science Review*, 70, 41–9.

15. Griffiths, J. (2017, September 1). 'All bets are off' in New Zealand vote as 'Jacindamania' boosts Labour'. CNN. <https://amp.cnn.com/cnn/2017/09/01/asia/new-zealand-labour-jacinda-ardern/index.html>
16. Kwai, I. (2017, September 5). 'New Zealand's Election Had Been Predictable. Then "Jacindamania" Hit.' The New York Times. <https://www.nytimes.com/2017/09/04/world/asia/jacinda-ardern-new-zealand.html>
17. Lebryk, T. (2021, April 18). Introduction to The Structural Topic Model (STM). <https://towardsdatascience.com/introduction-to-the-structural-topic-model-stm-34>
18. Manifesto Project (2024). <https://manifesto-project.wzb.eu/>
19. Negara, E. S., Triadi, D., & Andryani, R. (2019). Topic Modelling Twitter Data with Latent Dirichlet Allocation Method. 2019 International Conference on Electrical Engineering and Computer Science (ICECOS), 386–390. <https://doi.org/10.1109/ICECOS47637.2019.8984523>
20. New Zealand Government (2024). <https://www.beehive.govt.nz/>
21. New Zealand Labour Party Policy Platform (2016, March). [https://web.archive.org/web/20180628154413/https://d3n8a8pro7vhmx.cloudfront.net/nzlabour/pages/3956/attachments/original/1459459622/2016\\_Policy\\_Platform.pdf?1459459622](https://web.archive.org/web/20180628154413/https://d3n8a8pro7vhmx.cloudfront.net/nzlabour/pages/3956/attachments/original/1459459622/2016_Policy_Platform.pdf?1459459622)
22. New Zealand Labour Party Policy Platform (2019, December). <https://apo.org.au/sites/default/files/resource-files/2019-12/apo-nid307339.pdf>
23. New Zealand Labour Party Policy Platform (2022, December). <https://drive.google.com/file/d/1Ah2xlM0-qzemx76LP-dtmzQulahnU7U0/view>
24. New Zealand Parliament. (2023, April 5). Valedictory Statement — Rt Hon Jacinda Ardern. Hansard Debates. [https://www.parliament.nz/en/pb/hansard-debates/rhr/combined/HansDeb\\_20230405\\_20230405\\_44](https://www.parliament.nz/en/pb/hansard-debates/rhr/combined/HansDeb_20230405_20230405_44)
25. Quinlivan, M. (2023, November 24). 'Election 2023: National, ACT and NZ First's Coalition agreement'. Newshub. <https://www.newshub.co.nz/home/politics/2023/11/election-2023-national-act-and-nz-first-s-coalition-agreement.html>
26. Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural Topic Models for Open-Ended Survey Responses. American Journal of Political Science, 58(4), 1064–1082. <https://doi.org/10.1111/ajps.12103>
27. Shuttleworth, K. (2017, September 2). 'Jacindamania: rocketing rise of New Zealand Labour's fresh political hope'. The Guardian. <https://amp.theguardian.com/world/2017/sep/02/jacindamania-rocketing-rise-of-new-zealand-labours-fresh-p>

28. Thomas, D. (2005). Type-token Ratios in One Teacher's Classroom Talk: An Investigation of Lexical Complexity. University of Birmingham. <https://www.birmingham.ac.uk/documents/college-artslaw/cels/essays/languageteaching/daxthomas2005a.pdf>
29. To, V., Fan, S., Thomas, D. (2013). Lexical Density and Readability: A Case Study of English Textbooks. *The International Journal of Language, Society and Culture*.
30. Withers, T. (2017, September 19). 'How Jacinda-Mania Shook Up New Zealand's Election'. Bloomberg. <https://www.bloomberg.com/politics/articles/2017-09-19/how-jacinda-mania-shook-up-new-zealand-s-election-quicktake-q-a>
31. Zhao, W.X., Jiang, J., Weng, J., He, J., Lim, E.P., Yan, H. and Li, X., 2011, April. Comparing twitter and traditional media using topic models. In *European conference on information retrieval* (pp. 338-349). Springer, Berlin, Heidelberg.