

The Representation of Political Parties

A Network Analysis of Mastodon and the Dutch House of Representatives elections

Elena Ibi¹[2643914], Bas de Greeuw¹[2794502], Danny de Vries¹[2794008], and
Abhilash M. Abhilash¹[2795647]

Vrije Universiteit van Amsterdam, Amsterdam, The Netherlands
<https://vu.nl/en/about-vu/departments/computer-science>
`{firstname.lastname}@student.vu.nl`

Abstract. The abstract should briefly summarize the contents of the paper in 15–250 words summarizing the research question, method and main findings.

Keywords: Social Web · Social Network · Network Analysis · Mastodon · Dutch Elections · Political Parties · User-generated Content.

1 Introduction

On the 22nd of November, 2023, the dutch population voted for the national Dutch House of Representatives elections [?]. Of the 13,473,750 eligible voters, 10,475,139 voted, resulting in a turnout of around 77.7%. Of the 26 parties that participated in the election, 15 parties received enough votes for a seat in the House of Representatives. Prior to elections, viewpoints and topics of particular parties are usually widely discussed on Social Networking Sites (SNS). As an example, users might post their support, or opposition, for a political party, discuss topics that are mentioned in party manifestos, and discuss candidates that are on the electable candidates list.

Since the last elections, there have emerged new Social Networking Sites, the most publicly known and somewhat succesful being Mastodon¹. Mastodon is a decentralized social network with microblogging features similar to X², colloquially and formerly known as Twitter. Using Mastodon to analyze the adoption of a new social media network for political purposes is interesting for mainly two important reasons. (1) Since it's release, especially in the last two years, Mastodon has seen a massive increase in users and activity. Their analytics publication suggest a growth from around 3.500.000 in october 2022 to 8.100.000 users in october 2023 [?]. A large influence for this growth is the acquisition of Twitter by Elon Musk [?]. Many users, at least temporarily, transition from X/Twitter to Mastodon. (2) Elections for the Dutch house of representatives

¹ <https://joinmastodon.org/>

² <https://twitter.com/>

occur every 4 years. Moreover, we’ve even seen dutch political parties create Mastodon instances for their party members. For example the servers of Bij1³ and Piratenpartij⁴. Therefore, we can assume Mastodon is becoming increasingly more representative of the dutch voting population or its political environment. (*) Besides these reasons, it is important to note that the use of Mastodon has a data collection advantage. Their API is mostly public mostly secured with the use of simple rate limiting. Therefore, it is easy to collect data from the platform and utilize it for case study purposes.

In order to investigate this phenomenon, this report aims to explore the following research question: ***"To what extent is the relatively new Social Networking Site Mastodon representative of the election voting of the dutch population?"***. To approach this research question in-depth, the following sub-questions are formulated:

- **R1:** *What’s the distribution of political parties on the platform and do they align with the outcome of the election?*
- **R2:** *What political topics are discussed in posts and are they representative of the election manifesto of political parties?*
- **R3:** *Do the topics that are discussed on the platform align with popular voting guides and results?*

In order to answer the research questions, this papers starts with an examination of prior research on Mastodon as a platform and literature using related methods. For the aforementioned approaches, the focus lies on network analysis, and semantical data analysis. The report will briefly explain the data gathering and pre-processing, using the Mastodon API data and other, election related, datasets. The data collection and analysis, and subsequently comparing it to government published or endorsed sources will attempt to conceptualize the adoption of this relatively new SNS. Finally, the report will conclude with a discussion of the results and a reflection on the research question, attempting to find out whether the SNS is representative, or conversely the official sources.

2 Related Work

Literature section with a short overview of other papers discussion related questions or using related methods or data

3 Methodology

3.1 Data collection datasets

To gather social web data from Mastodon, the official public Mastodon API⁵ is interfaced with, by the use of an unofficial wrapper library known as;

³ <https://social.bij1.org/about>

⁴ <https://mastodon.social/@Piratenpartij@social.globalpirates.net>

⁵ <https://docs.joinmastodon.org/client/intro/>

Mastodon.py⁶. Mastodon is an ActivityPub-based⁷ Twitter-like federated social network node. The API wrapper is feature complete for Mastodon API version 3.5.5. First a user account is created on the platform by completing the sign-up for an account flow on the Mastodon official website *joiningmastodon.org*. The account is created on the general and largest public server (provider) *mastodon.social* operated by the Mastodon gGmbH non-profit. However, the public API allows for exploring other public servers, which is integrated in the methods provided by the API wrapper.

To interact with the Mastodon API, an application registration is performed, which gives a client key and client secret to allow logging in and accessing API data using access tokens. For this research we mainly used API methods for:

- **Accounts, relationships and lists:** allows for getting information about accounts and associated data as well as update personal entries of this type.
- **Instance-wide data and search:** fetch information associated with the current instance, as well as data from the instance-wide profile directory. The API wrapper, uses this endpoint to combine results for certain search methods.
- **Streaming:** allow access to the streaming API for public, local and hashtag streams. The API wrapper utilizes streaming and a built-in rate limit tracker to optimize speed and data integrity.

Arguments and parameters used in queries are related to the Dutch elections, e.g. names of political candidates, popular topics from parties. However, this is further expanded upon in the data preprocessing and results section of this research. To check, validate, and cross-reference, the data is complemented with five additional "official" or "officially endorsed" election related data sources:

- **Institut Public de Sondage d’Opinion Secteur (IPSOS) exitpoll:** a market research company which, commissioned by the 'Nederlandse Omroep Stichting'⁸ (NOS; English: Dutch Broadcasting Foundation). They publish research surrounding the elections, attempting to capture data such as what demographic of voters switch between parties. For example, constituents of which municipalities have switched the most between parties[?]. This gives a comprehensive insight of voting behaviour from the recent election.
- **Government Open Data (overheid.nl):** specifically the datasets from The Dutch Electoral Council⁹ (Dutch: Kiesraad), the government body that is responsible for counting the votes and publishing the results[?]. This source is used as the official results of parties and candidates for the recent elections.
- **ProDemos voting guide (stemwijzer):** a voting guide called Stemwijzer¹⁰ with pre-defined topics. By answering positive, negative or neutral to 30 statements, electors can compare their positions with those of electable

⁶ <https://mastodonpy.readthedocs.io/en/stable/>

⁷ <https://www.w3.org/TR/activitypub/>

⁸ <https://nos.nl/>

⁹ <https://www.kiesraad.nl/>

¹⁰ <https://home.stemwijzer.nl/>

political parties. Many of these voting guides exist, however, ProDemos is colloquial known as the "official" voter guide, and partly funded by the dutch government [?]. This source provides insight in important topics from political parties for the recent elections, which subsequently get primed by the parties that are included.

- **Electoral Council (kiesraad)**: the Kiesraad¹¹ is a central electoral committee, an advisory body, and acts as a central polling station during the dutch house of representatives election. For this research we used the published Candidacy for the House of Representatives election list and the Political Party Registrar.
- **Netherlands Bureau for Economic Policy Analysis (cpb)**: the dutch economics bureau (CPB)¹² performs election manifestos analysis to determine how feasible manifestos of political parties are. This gives an overview of topics that are in the election manifestos of political parties.

3.2 Data preprocessing (scope)

With these gathered sources surrounding the dutch-election, a relevant scope for this report is determined. For the starting point of our temporal data, we constructed a timeline starting at Mastodon's initial release on 16 March 2016 (7 years ago). This means we analyse three House of Representatives elections. In 2017 with a turnout rate of 81,57% of dutch population, in 2021 with a turnout rate of 78,71% of dutch population and the most recent, as of writing this paper in 2023, with a turnout rate of 77,75% of dutch constituents. Usually dutch elections take place every four years, thus we expected the next elections to be held in 2025. However a snap election was called after the fourth term of Rutte and his coalition collapsed due to immigration policy disagreements between the coalition parties.

During this timeframe, a total of 42 unique parties were electable during the elections. 28 parties in 2017, 37 parties in 2021 and 26 parties in 2023. For this report we created a subset of political parties based on incumbent parties ('zittende partijen' in dutch) as a result of the most recent and previous elections. This gives a total of 20 parties. This research also has taken synonyms and abbreviations of parties into account, for example, Forum voor Democratie is often abbreviated to 'FvD' or GL-PvdA is still often mentioned with full names after the merger in 2022 as Groenlinks and Partij voor de Arbied. *Table ??* displays an overview of parties used for this research. For the analysis of topics, we manually created an aggregated dataset, as shown in *Table ??* of main topics, and potential subtopics, based on the statements provided by the stemwijzer and the analysis of the manifestos made by the CPB.

¹¹ <https://english.kiesraad.nl/>

¹² <https://www.cpb.nl/en/charted-choices-2025-2028>

Overview of all 'sitting' Political Parties			
VVD	D66	PVV	CDA
GL-PvdA	SP	FvD	PvD
ChristenUnie	Volt	Ja21	SGP
Denk	50Plus	OPNL	BBB
Bij1	Piratenpartij	Splinter	BVNL

Table 1: All parties based on the results of the dutch elections of 2017, 2021 and 2023.

Overview of all topics (query words)			
Immigration	Economics	Climate	Agriculture
Taxes	Healthcare	Rural areas	Public transport
Defence	Education	Housing	Religion
European Union	Income	Sentences	

Table 2: All topics based on the election manifestos summaries and voting guide.

3.3 Data analysis

This research focusses on Network Analysis which is performed using the Python programming language (latest stable version 3.12.0) following the Jupyter¹³ notebook standard which combines code, documentation, data, and visualizations. Besides the Mastodon.py API wrapper to access the Mastodon data the notebooks use NetworkX¹⁴ to perform Network Analysis. This library allows for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks and has built-in functions for common analysis measurements and displaying graphs. To display data plot distributions, and creating visualizations a combination of Matplotlib¹⁵ and Seaborn¹⁶ are used.

4 Results

The results section is divided into three subsections eachs corresponding to the individual research questions as described in the *introduction* section. The first (R1) research question related to activity of political parties on the platform, the second (R2) relates to topics from election manifestos and voting guides and the third (R3) is a network analysis of instances and servers of political parties on the network. Each with accompanying descriptions of the methods performed and graphs to further understand the data and the result of the analysis.

Before describing the individual research questions what's interesting to note related to the the general increase of activity in terms of the dutch elections

¹³ <https://jupyter.org/>

¹⁴ <https://networkx.org/>

¹⁵ <https://matplotlib.org/>

¹⁶ <https://seaborn.pydata.org/>

on the whole Mastodon platform is that if a query is performed on all toots related to dutch elections (includes terms such as 'dutch elections', 'verkiezingen', 'tweedekamer', 'tweede kamer') and plot them based on year we get the result as shown in figure ?? . This results in a total of 18.230 toots which included election related terms of which around 16.000 are in the last two years. A clear spike in activity is seen in the most recent election year, 2023, as opposed to previous elections, 2017 and 2021, which show almost no activity on general election query words.

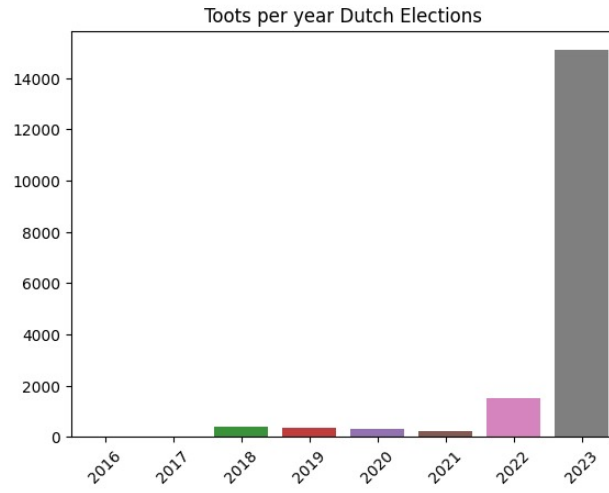


Fig. 1: Bar chart of query words related to dutch elections

4.1 Activity of Political Parties (R1)

Finding M1: *Out of all parties x parties are present on Mastodon and have instances.*

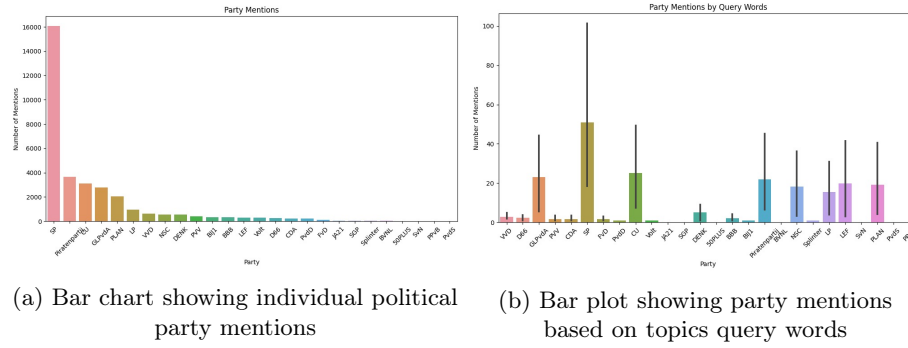


Fig. 2: Graphs visualizing party activity based on party name mentions and query words based on topics

4.2 Election-related topics and query words (R2)



Fig. 3: Party mentions related to topics

Finding M2: Out of all parties x parties are present on Mastodon and have instances.

4.3 Network Analysis of Parties instances and servers (R3)

Finding M3: Out of all parties x parties are present on Mastodon and have instances.

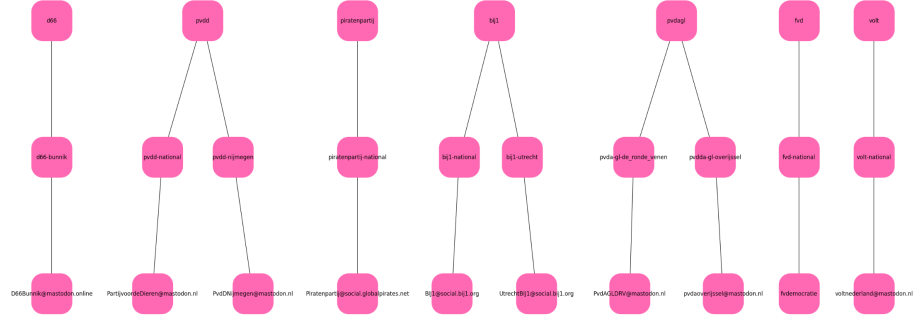


Fig. 4: Overview of parties and corresponding sub instances

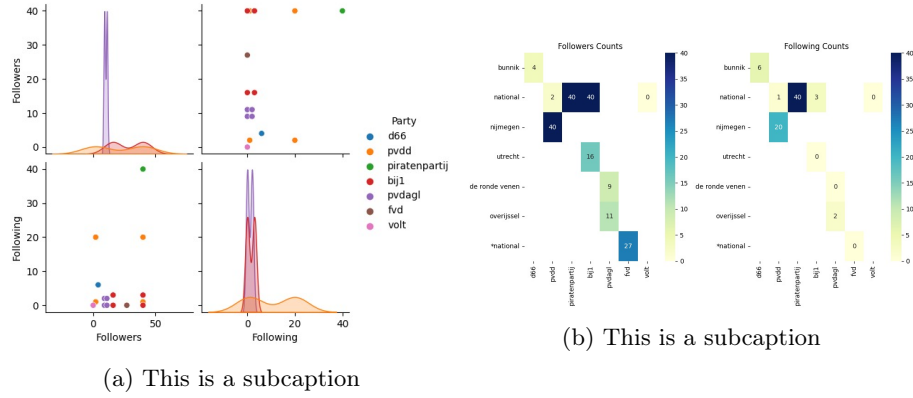


Fig. 5: Graphs visualizing parties and follower counts

5 Discussion

Privacy & ethical considerations

6 Future work

Each of the parties is placed on a political spectrum (left, lean left, center, lean right, right). Quote a source. There is probably an 'official' list for this. Based on what they voted (maybe stemmentracker)? If they are left-wing, right-wing. How many zetels etc. this could be part of our weighting. We now say each party is equal in our graphs. But some parties are 'larger' in terms of ledennummers etc.

List of topics could probably be larger. We now manually made a list of topics based on summaries and manually scanning through table of content of manifestos. This could be automated in future work by having NLP scan table of contents from election manifestos, 'moties' or determine the topic using classification based on voting guides (stemwijzers).

The network analysis focusses on political parties but most party have one 'fractievoorzitter'. To further analyse the network we could also switch to individual user accounts of the fractievoorzitter's of the kamer and see based on who they follow or their follower account what other politicians of the party have accounts on Mastodon.

The topics displayed in the voting guides are the topics parties mostly disagree on and are picked based on a larger list of topics. That doesn't necessarily mean these topics are the most popular topics in society. For future research we would like to cross-reference topics from the voting guides with topics discussed on social networks to see if topics that gain traction on social networking sites align with topics from the stemwijzer.

7 Conclusion

With this work, we invite researchers, journalists and practitioners alike to further investigate Mastodon in relation to the Dutch House of Representatives elections or explore any other new and upcoming Social Networking Site using similar methodology.

8 Acknowledgements

We thank coordinator Dr. Davide Ceolin (Vrije Universiteit Amsterdam) for providing guidance and assistance during the project and Dr. Emmanuelle Beauxis-Aussalet (Vrije Universiteit Amsterdam) who provided valuable feedback and answered our questions during the seminars which helped us further expand our research.

9 Conflicts of Interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The author(s) has no affiliation with any of the companies and organizations mentioned in this article and this work has not been supported by any funding agency, private organization, or political party.

10 Appendix

In the spirit of open research in order to support reproducibility and enable future work in this problem space the datasets and Python notebooks in this work are publicly available on GitHub using the MIT License. Under the *dandevri* username (one of the authors) we have several a code repository with several subfolders:

1. **Notebooks:** Source Code for the Python Jupyter Notebooks for data scraping and processing. <https://github.com/dandevri/vu-social-web-data/notebooks>
2. **Datasets:** The processed and transformed datasets used in the notebooks. <https://github.com/dandevri/vu-social-web-data/datasets>