The Representation of Political Parties
A (Network) Analysis of Mastodon and the Dutch House of Representatives elections

Elena Ibi[1][2643914], Bas de Greeuw[1][2794502], Danny de Vries[1][2794008], and Abhilash M. Abhilash[1][2795647]

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

Abstract. Prior to elections viewpoints of parties of various users are often expressed on Social Networking Sites (SNS) since expressing opinion is one of the core characteristics of social media. Recently, (on November 2023) elections for the Dutch House of Representatives took place. This study focuses on on a data-driven social web approach aiming to answer the research question: *to what extent is the relatively new SNS 'Mastodon' representative of the election voting of the dutch population?* In the analysis phase of this research, processed data from several election-related resources such as political parties and topics from election manifestos are cross-referenced and statistically analysed with Mastodon instance-wide (e.g. servers) and Streaming API (e.g. toots and mentions). The results indicated that there is a significant increase in activity related to the Dutch election on the Mastodon platform and political parties and candidates are increasingly creating user accounts and instances. Privacy and ethical considerations when accessing Mastodon platform data are discussed and in the future work section, the study addresses several further enhancements that can be performed to automate more of the processing methods and to further expand the election data sources used.

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 [6]. 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) [11]. 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, have emerged new Social Networking Sites, the most publicly known and somewhat successful being the general-purpose social network [2] Mastodon[1]. Mastodon is a decentralized social network with microblogging features similar to X[2], colloquially and formerly known as Twitter. Using Mastodon to analyse the adoption of a new social media network for political purposes is interesting for mainly two important reasons. (1) Since its release, especially in the last two years, Mastodon has seen a massive increase in users and activity. Their analytics publication suggests a growth from around 3,500,000 in October 2022 to 8,100,000 users in October 2023 [1]. A large influence on this growth is the acquisition of Twitter by Elon Musk [9]. Many users, at least temporarily, transitioned from X/Twitter to Mastodon. (2) Elections for the Dutch House of Representatives occur every 4 years. Moreover, we have even seen Dutch political parties create Mastodon instances for their party members. For example, the servers of Bij1[3] and Piratenpartij[4]. 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 and, is 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.

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 discussed on the platform align with popular voting guides and results?*

To answer the research question and the related sub-questions, this paper 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 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.

---

[1] https://joinmastodon.org/

[2] https://twitter.com/

[3] https://social.bij1.org/about

[4] https://mastodon.social/@Piratenpartij@social.globalpirates.net

2   Related Work

Social networking sites and other popular online platforms have been used as a means to express political opinions and show support and/or dissent for particular political parties or ideologies quite regularly during election periods since the rise of social media. Citizens are able, thanks to social media, to follow politicians, political commentators, and political consultants as a way to spread messages of endorsement and opposition through the sharing and posting of digital content; this phenomenon known as 'political advertising' took place even before the rise of social media and allowed for properly structured campaign advertisements to take place on digital platforms (https://blog.oup.com/2017/06/history-political-social-media/) As a matter of fact, as well as political opinions and views provided by individual users or voters, political parties use these platforms and the underlying channels of communication to run their political campaigns and thus, draw further attention for political debate and discourse.

Since the officially recorded use of social media for political campaigns, with Barack Obama's electoral campaign run in 2008, until current times, social media has been one of the primary political ideology battlefields across the globe, optimally leveraging the growing number of social media users and consequently that of people eligible to vote. (https://journals.sagepub.com/doi/10.1177/20563051211063461) This incremental evolution saw a deeper and more analytical organization of campaigns and political discourse on social media platforms effectively making these platforms key players in the domain of digital political journalism. (Kreiss D. (2012). *Taking our country back: The crafting of networked politics from Howard Dean to Barack Obama.* Oxford University Press.)

Consequently, online political discourse through digital platforms became

3   Methodology

3.1   Data collection & datasets

To gather social web data from Mastodon, the official public Mastodon API [5] is interfaced with, by the use of an unofficial wrapper library known as; Mastodon.py[6]. Mastodon is an ActivityPub-based[7] 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, *joingmastodon.org*. The account is created on the general and largest public server (provider) *mastodon.social* operated by the Mastodon GmbH non-profit. However, the public API allows for exploring other public servers, which are integrated in the methods provided by the API wrapper.

---

[5] https://docs.joinmastodon.org/client/intro/

[6] https://mastodonpy.readthedocs.io/en/stable/

[7] https://www.w3.org/TR/activitypub/

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:** fetch 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:** 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, and 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[8]. This gives a comprehensive insight into voting behavior from the recent election.
- **Government Open Data (overheid.nl)**: specifically the datasets from The Dutch Electoral Council[9] (Dutch: Kiesraad), the government body that is responsible for counting the votes and publishing the results[5]. This source is used as the official results of parties and candidates for the recent elections.
- **ProDemos voting guide (stemwijzer)**: a voting guide called Stemwijzer[10] with pre-defined topics. By answering positive, negative, or neutral to 30 statements, electors can compare their positions with those of electable political parties. Many of these voting guides exist, however, ProDemos is colloquially known as the "official" voter guide, and is partly funded by the Dutch government [10]. This source provides insight into important topics from political parties for the recent elections, which subsequently get primed by the parties that are included.
- **Electoral Council (kiesraad)**: the Kiesraad[11] is a central electoral committee, an advisory body, and acts as a central polling station during the Dutch House of Representatives election. For this report, we used the published

---

[8] https://nos.nl/
[9] https://www.kiesraad.nl/
[10] https://home.stemwijzer.nl/
[11] https://english.kiesraad.nl/

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 the 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 the next elections were expected 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. A subset of political parties has been constructed, 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 1* displays an overview of parties used for this research. For the analysis of topics, a manually created aggregated dataset is used, as shown in *Table 2* of main topics, and potential subtopics, based on the statements provided by the Stemwijzer and the analysis of the manifestos made by the CPB.

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.

---

[12] https://www.cpb.nl/en/charted-choices-2025-2028

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 focuses on (Network) Data Analysis which is performed using the Python programming language (latest stable version 3.12.0) following the Jupyter[13] notebook standard which combines code, documentation, data, and visualizations. Besides the Mastodon.py API wrapper to access the Mastodon data, the notebooks use NetworkX[14] 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 create visualizations, a combination of Matplotlib[15] and Seaborn[16] are used.

4   Results

The results section is divided into three subsections, each corresponding to the individual research questions as described in the *introduction* section. (R1) Relates to the activity of political parties on the platform, (R2) relates to topics from election manifestos and voting guides, and (R3) relates to 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.

A clear increase in general activity surrounding the Dutch elections on the platform is found. When querying toots with general terms on Dutch elections, for example, "verkiezingen", "dutch elections", or "tweede kamer", the results have very clearly spiked in the last period, as shown in figure 1. A total of 18,230 toots were retrieved, which included these election-related terms, of which around 16,000 were in the last two years. Therefore, Mastodon has been more widely adopted for the most recent election year, 2023, as opposed to previous elections, 2017 and 2021, which showed almost no activity on these general queries.

---

[13] https://jupyter.org/
[14] https://networkx.org/
[15] https://matplotlib.org/
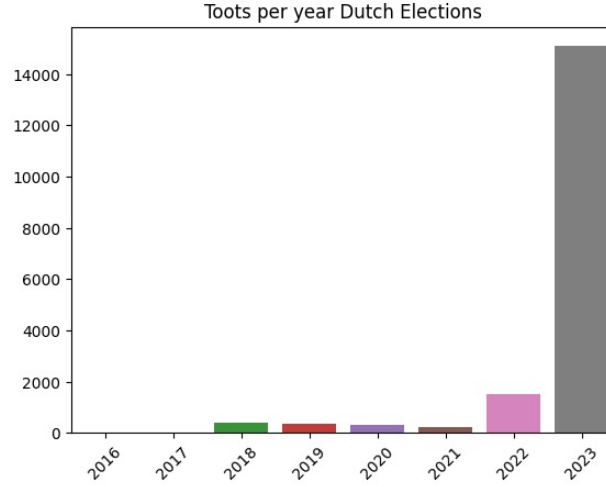[16] https://seaborn.pydata.org/

Fig. 1: Bar chart of query word results, using general terms surrounding Dutch elections

4.1   Activity of Political Parties (R1)

With the list of terms extracted from the voting guide, the platform is queried for these terms and cross-referenced with party names, resulting in a dataset reflecting what parties get mentioned most topically on Mastodon. The resulting graph in figure 2a makes it clear that the Dutch socialist party (SP) is mentioned in conjunction with the topical terms the most. The candlestick graph, where the wick represents the deviation of terms, shows that the SP also has the widest variety of topics mentioned. One possible explanation is that their party initials are found in Dutch spelling quite commonly. However, when isolation their initials, the results are the same, although this could still be the result of errors in our search endpoint syntax. It is safe to assume the SP is an outlier, considering the second-highest activity party, is active on the platform itself and has a separate Mastodon instance

Activity around parties seems to focus on a small set of parties. The second-highest mention, which seems like a more realistic activity hotspot, is the Piratenpartij. Piratenpartij has its own Mastodon instance and multiple accounts, which also have the most activity itself, when compared to other party accounts on the platform, more of this is shown in the subsection for (R3). Their deviation in figure 2b also shows they are talked about in conjunction with a wide variety of topics. Interestingly, the GL-PvdA party comes very close to the Piratenpartij, even though they are not active on the platform themselves. An easy explanation would be, that this party has a lot of young voters and was mentioned in passing quite often because of their fusion (they used to be two separate parties). Moreover, a large group of their constituents is known to be active in online

(a) Bar chart showing individual political party mentions

(b) Candlestick chart showing parties being mentioned based on topical query words. The wick represents the spread of unique topics
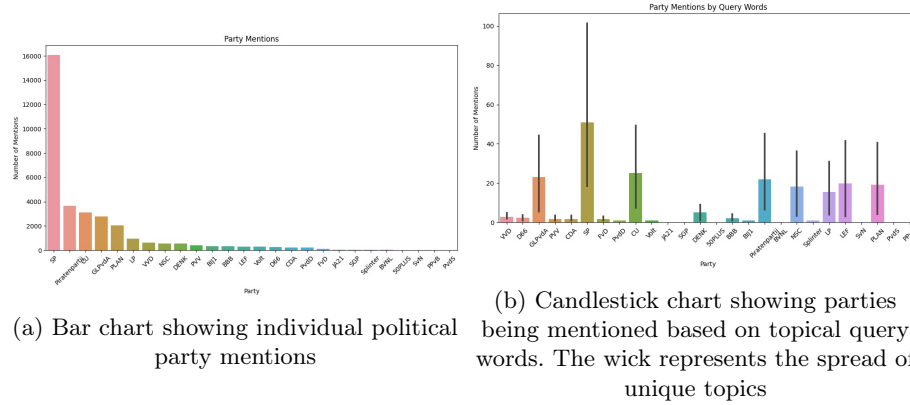
Fig. 2: Graphs visualizing party activity based on party name mentions and query words based on topics

discourse on other social media as well, as well as the sister party of the SP aimed at teenagers.

Finding M1: *Very few parties get mentioned, however, the parties that are mentioned are actively discussed. The cross-referencing might be influenced by the prevalence of the initials of parties in the Dutch language.*

4.2   Election-related topics and query words (R2)

The related topics mentioned per party mention show an interesting spread when plotted. A hot topic, reasonably, is climate (and climate change) and "economy", which in hindsight, might have been too general of a term, that etymologically contains most other terms as well. These topics get mentioned almost exclusively by left-leaning parties, although it must be noted that almost all parties that have a reasonable amount of activity are left-leaning. When comparing this chart to the activity around parties, it becomes clear certain queries result in actual political discourse toots and queries that don't have that much activity. Most mentions group around the active parties, while other mentions are so sparse and distributed, that it seems more like the discourse is somewhat evenly spread across topics when the party and topic itself get mentioned by users. There is no clear distinction in the actual difference of discourse between parties.

Finding M2: *Overall discourse seems to be on similar topics, only showing up in searches surrounding parties that have activity on the platform*

4.3   Analysis of Party accounts and servers (R3)

To more easily convey what parties have adopted the network, multiple queries have been performed, and analysed to find accounts connected to parties. A list of all parties, including their nicknames and different spellings of abbreviations

Fig. 3: Party mentions related to topics

or those nicknames is looped through and queried in conjunction with different terms like "official", "party", "House of Representatives", or "elections", albeit in Dutch. These queries resulted in the tree of different party accounts, as seen in figure 4.

Noticeably, some parties are not active on the main server but instead have made their own server. The Piratenpartij and Bij1 both have their own server, where the Piratenpartij server is quite active, at least compared to Bij1's server. This also does not surprise from a political point of view, as the Piratenpartij, was founded to legislate internet laws and protect net neutrality. The self-proclaimed radical left Bij1 party also occasionally mentions their need to operate in decentralized structures and hierarchies in their news and SNS outings. The distinction between servers is more easily seen in figure 5 which branches to parties starting from the server as opposed to branching off from the party.

Fig. 4: Tree of party accounts, branching from their district

Fig. 5: Tree of party accounts, branching from their respective server

(a) This is a subcaption

(b) This is a subcaption

Fig. 6: Graphs visualizing parties and follower counts

Finding M3: *Out of all parties, 7 parties are present on Mastodon and 2 have their own instances.*

5   Discussion

Privacy & ethical considerations mainly for our research project on Mastodon

Besides these considerations specifically for this research, in general, it is becoming increasingly more difficult to investigate the social web and user-generated content and metadata due to several privacy scandals related to social networking platforms. Big Data Graph API's access is being heavily restricted and only available after extensive authorisation (e.g. Facebook Graph API after the Cambridge Analytica scandal [3]) and real-time streaming API's access is

being rated limited, paywalled, or blocked entirely due to fear of AI scraping (e.g. Twitter rate limited their API [4], developers need to pay for Reddit API usage [7]).

6   Future work

In future work, it will be important to consider several key factors to enhance the effectiveness and enrich the datasets used in this research which are described in more detail below.

Currently, this research uses a subset of metadata available for each political party. Mainly the party names and election results. To further cross-reference the Mastodon we could enrich the metadata for each party to incorporate political spectrum data (e.g. official grouping from ProDemos or how parties voted based on the StemmenTracker [17]) and 'weigh' each party based on left-wing, neutral or right-wing. With this, the research could incorporate a 'popularity weighting' of each party since current visualizations treats each party the same not taking into account the number of seats in the House of representatives or the number of members.

The (Network) analysis currently also focuses on political parties and thus instances or servers related to the political parties but analysis of individual user accounts of faction leaders is not yet performed. To further analyse the network and get an overview of connections analysis of personal accounts needs to be performed to plot further relations of the network based on who faction leaders follow or who are following them (e.g. handshake lemma) [12].

Query words and a list of election-related topics in this research are manually labeled. This can be further expanded by automating this task, for example, using Natural Language Processing (NLP) to scan through the table of contents of election manifestos or determine topics automatically based on keywords in quotes from voting guides. This also further enhances the reliability of the data since topics displayed in voting guides are what parties mostly disagree on. This doesn't necessarily mean these topics are the most popular in society among voters. For future research, it would also be interesting 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 voting guides.

7   Conclusion

The research presented in this paper delved into the dynamics of political activity on the Mastodon platform, focusing on the Dutch House of Representatives elections. The results, analyzed through three research questions, provided valuable insights into the adoption and engagement of political parties on this decentralized social network.

---

[17] https://home.stemmentracker.nl/

Firstly, the overall activity surrounding the Dutch elections on Mastodon exhibited a significant increase, particularly in the most recent election year, 2023 contrasted with minimal activity observed in the preceding elections of 2017 and 2021. The examination of the activity of political parties revealed that while a few parties dominated the discussion, the nature of their engagement varied. The Socialist Party (SP) emerged as a notable outlier, prominently featured in discussions, possibly influenced by the prevalence of their initials in the Dutch language. Interestingly, parties like the Piratenpartij demonstrated active engagement, with their own Mastodon instance, while others, like the GL-PvdA party, garnered attention despite not being directly active on the platform.

Furthermore, the analysis of election-related topics and query words highlighted the prevalence of certain themes, such as climate and the economy, predominantly associated with left-leaning parties. However, no distinct differences in discourse were observed between parties, emphasizing the even spread of discussions across various topics when parties with activity on the platform were mentioned.

Finally, the exploration of party accounts and servers revealed that a limited number of parties actively participated in Mastodon. Notably, some parties established their own instances, reflecting a commitment to decentralized structures and specialized platforms. The Piratenpartij, with its focus on internet laws and net neutrality, and Bij1, advocating for decentralized structures, exemplified this trend.

In summary, Mastodon's role in political discourse around Dutch elections has evolved, with increased activity and nuanced engagement by political parties. The findings contribute to our understanding of the dynamics of political discussions in decentralized social networks, shedding light on the platforms and strategies adopted by parties to navigate the digital social media networking landscape. With this work, we invite researchers, journalists, and practitioners alike to further investigate Mastodon in relation to the Dutch House of Representatives elections and explore any other new and upcoming Social Networking sites using a similar methodology.

8   Acknowledgements

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 one of the authors' usernames, a code repository with several subfolders is available:

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

References

1. API, M.: Mastodon analytics (2023), https://mastodon-analytics.com/, accessed: 2023-11-09
2. boyd, d.m., Ellison, N.B.: Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication **13**(1), 210–230 (10 2007). https://doi.org/10.1111/j.1083-6101.2007.00393.x, https://doi.org/10.1111/j.1083-6101.2007.00393.x
3. Chang, A.: The facebook and cambridge analytica scandal, explained with a simple diagram (2018), https://www.vox.com/policy-and-politics/2018/3/23/17151916/facebook-cambridge-analytica-trump-diagram, accessed: 2023-12-20
4. Davis, W., Lawler, R.: Elon musk blames data scraping by ai startups for his new paywalls on reading tweets (2023), https://www.theverge.com/2023/7/1/23781198/twitter-daily-reading-limit-elon-musk-verified-paywall, accessed: 2023-12-20
5. Kiesraad: Databank verkiezingsuitslagen (2023), https://www.kiesraad.nl/verkiezingen/verkiezingsuitslagen, accessed: 2023-11-09
6. Kiesraad: Kiesraad stelt uitslag tweede kamerverkiezing 22 november 2023 vast (2023), https://www.kiesraad.nl/actueel/nieuws/2023/12/01/kiesraad-stelt-uitslag-tweede-kamerverkiezing-22-november-2023-vast, accessed: 2023-11-09
7. Lutkevich, B.: Reddit pricing: Api charge explained (2023), https://www.techtarget.com/whatis/feature/Reddit-pricing-API-charge-explained, accessed: 2023-12-20
8. NOS: Nos verkiezingsuitslagen (2023), https://app.nos.nl/nieuws/tk2023/, accessed: 2023-10-09
9. Perez, S.: Decentralized social network mastodon grows to 655k users in wake of elon musk's twitter takeover (2022), https://techcrunch.com/2022/11/03/decentralized-social-network-mastodon-grows-to-655k-users-in-wake-of-elon-musks-twitter-takeover/?guccounter=1, accessed: 2023-11-09
10. ProDemos: Over de stemwijzer (2023), https://home.stemwijzer.nl/over-stemwijzer/, accessed: 2023-11-09

11. Richter, A., Koch, M.: Functions of social networking services. From CSCW to Web 2.0: European Developments in Collaborative Design Selected Papers from COOP08 (2008)
12. Yang, X., Guo, Y., Liu, Y., Steck, H.: A survey of collaborative filtering based social recommender systems. Computer Communications **41**, 1–10 (2014). https://doi.org/https://doi.org/10.1016/j.comcom.2013.06.009, https://www.sciencedirect.com/science/article/pii/S0140366413001722