Proceedings of the Twentieth International AAAI Conference on Web and Social Media (ICWSM 2026)

Anatomy of an Election from a Gender Perspective

Zeynep Pehlivan, Esli Chan
McGill University
zeynep.pehlivan@mcgill.ca
esli.chan@mcgill.ca

Abstract

This paper investigates how gender shapes online political communication during an electoral campaign by analyzing the 2025 Canadian federal election. Drawing on a novel dataset of 193,620 social media posts from 921 candidates, we employ transformer-based topic modeling and a linguistic gender axis to quantify communicative patterns. Our analysis reveals three key findings. First, a statistically robust stylistic difference exists between the discourse of male and female candidates. Second, this gendered linguistic gap is not merely an artifact of ideology; it persists systematically across all major political parties, demonstrating its robustness across partisan contexts. Third, thematic analysis shows that women more frequently employ community-action framing across a range of policy-related discussions, while men place greater emphasis on economic and industrial issues during the electoral period. While the magnitudes of these effects are shaped by Canada's electoral environment, the mechanisms we identify offer a transferable framework for measuring gendered discourse in other political contexts.By documenting these patterns at scale, this study provides a high-resolution portrait of gender's persistent role in contemporary political communication and a replicable methodological toolkit for its study.

Introduction

In contemporary democratic societies, the digital transformation of political communication has fundamentally altered how gender dynamics manifest in electoral processes. While traditional analyses of political representation have focused on numerical parity and media coverage patterns, the rise of social media platforms has created new opportunities to examine how gender biases operate at multiple levels.

Canada provides a particularly compelling case. Despite its global reputation for progressive gender policies, disparities in political representation and discourse remain. Female politicians continue to face not only numerical underrepresentation across parties and regions, but also qualitative differences in how their communications are mediated by both algorithms and audiences.

The 2025 Canadian federal election offers a useful empirical setting for examining online political communication during an electoral period. The election was called on March 23 and concluded with voting on April 28, corresponding to the shortest campaign duration permitted under the Canada Elections Act. Rather than treating this compressed timeline as exceptional, we leverage it as a clearly bounded period that concentrates campaign communication and facilitates the observation of temporal shifts in discourse. To capture both pre-campaign activity and post-election dynamics, we analyze data from February 23 to May 28, 2025. This window spans Parliament's prorogation, the election call, the campaign period, and the immediate post-election phase, allowing us to trace how gendered narratives evolve across different stages of the electoral cycle. This design choice increases internal validity while naturally limiting external generalizability to an extent: our findings speak to the mechanisms through which gender shapes political discourse under heightened electoral pressure, not to the precise magnitudes observed in the Canadian case. By foregrounding mechanism over magnitude, the study provides an analytical framework that can be replicated in other electoral environments without assuming that the Canadian setting is exceptional.

Nevertheless, this case offers broader inferential insight into online political and electoral communication in advanced democracies. Canada's parliamentary electoral structures and highly digitized media environment—characterized by widespread internet penetration, high digital literacy, and the presence of politicians, legacy and alternative media, influencers, and citizens—closely resemble those of many consolidated democracies.

In this paper, we present a large-scale computational analysis of gender in electoral discourse. We draw on a novel dataset of 193K social media posts from 921 candidates across six political parties, spanning a diverse ecosystem of platforms (X, Instagram, YouTube, TikTok, and Bluesky), which enables us to analyze gendered communication patterns both in aggregate and across platform-specific affordances. Our goal is to move beyond anecdotal observations by quantifying the thematic and linguistic differences in how male and female candidates communicate in a high-stakes, compressed campaign. Specifically, this paper addresses the following question: How do male and female political candidates in Canada leverage social media discourse differently during an election campaign? Our analyses span multiple levels:

- Descriptive distributions of candidates by party and region.
- Exploratory analysis of posting activity.
- Topic modeling to uncover thematic divergences.
- Construction of a gender bias axis to measure linguistic styles.
- Statistical testing of gender, concept associations, and temporal shifts.

By integrating these computational approaches, we highlight both persistent gender inequalities in political discourse and methodological innovations for quantifying gender bias in dynamic electoral contexts.

The findings contribute to several scholarly domains: political communication research gains insights into digital-age gender dynamics; the measurement of gender representation in electoral politics and political institutions; computational social science benefits from a methodological framework for tracking temporal shifts in gender bias during high-intensity events. Practically, the results offer guidance for political practitioners, platform designers, and policymakers concerned with creating more equitable digital political spaces. Understanding how gender biases operate in computational systems processing political content is crucial for developing fairer, more representative digital democratic practices. The Canadian election provides an exceptional window into these dynamics, offering lessons relevant far beyond Canada's borders.

Related Work

Digital Gender Representation in Electoral Politics

Digital communication and representation is increasingly important in democratic practices, enabling connectivity and accountability between political representatives and the Canadian polity (Fountaine, Ross, and Comrie 2019). In electoral periods, digital communication and presence can impact the success of potential candidates. However, digital representation and communication, alongside financial barriers and access to resources, represent some of the many challenges women face in political spaces (Trimble 2018; Goodyear-Grant 2013a). The 'gender mediation thesis' describes the way that women are presented differently than men through gendered frames that can reinforce or contest gendered inequalities and stereotypes (Trimble 2018; Gidengil and Everitt 1999; Goodyear-Grant 2013a; Asr et al. 2021). For example, women political figures in the media are portrayed as confrontational as compared to men who appear to be assertive, demonstrating the 'unnatural' nature of women in politics (Gidengil and Everitt 1999). Women can also be portrayed alongside particular topics that are more feminine as compared to men who are more likely to be associated with substantive political and policy issues, such as foreign policy or finance (Goodyear-Grant 2013a). Gender mediation can occur through reproduction of skewed or unequal coverage that favour women or male candidates (Goodyear-Grant 2013a; Rohrbach 2024). Therefore, women candidates often have to consider how they perform their gender as political officials and in the online space (Sullivan 2023; Butler 1990; Dittmar 2020).

Social media presents an important opportunity for women political candidates to reclaim their narratives. Women candidates can directly engage with supporters, establish their priorities, construct their own image, and shape public discourse on their own terms (Kreiss, Lawrence, and Mcgregor 2017; Sullivan 2023). Social media can provide visibility for women candidates who may be excluded from institutional mainstream media coverage. However, women candidates may also be subject to direct receipt of harassment and misogyny as compared to their male counterparts (Rheault, Rayment, and Musulan 2019; Krook 2017). Moreover, online algorithm biases can promote content that is sensationalist or controversial, promoting stereotypical or harmful gendered representation of women candidates (Noble 2018; Vaidhyanathan 2018; Chan 2025).

Computational Approaches to Gender and Language

The quantification of gendered patterns in text has evolved significantly with advances in Natural Language Processing (NLP). While early approaches relied on dictionary-based methods, a paradigm shift occurred with distributional semantics. The seminal works (Bolukbasi et al. 2016; Kumar et al. 2020; Zhao et al. 2018) demonstrated that geometric relationships in static word embeddings (e.g., Word2Vec (Mikolov et al. 2013)) capture human-like stereotypes. The "gender axis" method, which defines gender as a vector in semantic space, emerged from this work and became a foundational technique.

This method's principles have been adapted and critically examined in the era of large language models (LLMs). Researchers now extend these analyses to contextual embeddings from models like BERT and sentence-transformers, allowing for more nuanced, context-dependent measurements of bias and style (Bartl, Ruppenhofer, and Scheible 2020; Bhardwaj, Majumder, and Poria 2020). For instance, recent surveys have moved beyond single bias axes to explore methodological challenges in transformer-based models (Nemani, Mittal et al. 2023). Others have proposed new sentence-level bias metrics, enabling the study of gender representation in multilingual embeddings (Dolci et al. 2023). Our work, while grounded in the robust logic of the original axis method, employs modern multilingual sentence-transformers. This allows us to apply these proven techniques to a contemporary, noisy, and multilingual dataset, contributing empirical findings to the ongoing study of gendered linguistic styles.

Computational Analysis of Political Discourse

The use of computational methods to analyze political communication on social media is a rapidly growing field. Researchers leverage large-scale datasets to move beyond anecdotal evidence and uncover systematic patterns in political strategy and public engagement. Topic modeling, for example, is widely used to identify the policy agendas of candidates, track shifts in party messaging over time, and understand how legislators allocate their attention across different issues (Theocharis and Jungherr 2020; Belcastro et al. 2022). While traditional methods like Latent Dirichlet Allocation (LDA) have been common, newer transformer-based models like BERTopic, which we employ, are increasingly preferred for their ability to capture semantic coherence in short, noisy social media texts.

Beyond topic analysis, computational studies of elections frequently examine candidate strategy through posting frequency, sentiment analysis, and network analysis of follower/retweet graphs. The high-volume, real-time nature of social media data makes it particularly well-suited for studying fast-moving events like elections. Some studies have explicitly framed elections as ”natural experiments” (Hale et al. 2014; Fujiwara, Müller, and Schwarz 2023), using the clear temporal boundaries of a campaign to measure the effects of specific events or strategies on public discourse and engagement.

Our study adopts this quasi-experimental framework, using Canadian federal election as an intervention point to analyze how gendered communication patterns evolve and adapt under heightened political pressure. By integrating these established computational techniques—topic modeling, quantitative linguistic analysis, and a quasi-experimental design—we situate our work at the intersection of these active research areas (Bright et al. 2020; Lorenz-Spreen et al. 2023).

Methodology

This section outlines the methodology used to construct and normalize the dataset across platforms and explains the methods used for analysis.

Data Collection and Annotation

We began by collecting the official list of candidates from government-maintained election registries. Each candidate was systematically annotated with party affiliation, geographic region, and gender. As gender was not provided in the official registries, we implemented a two-stage annotation procedure. First, we performed an initial automated classification using OpenAI's GPT-4.5 language model [1], drawing on publicly available candidate names and biographical details. Following this, we conducted a comprehensive manual verification where every automatically assigned label was cross-checked by research team members against authoritative sources, including official government websites, parliamentary directories, and public records. This hybrid approach combined computational efficiency with rigorous human oversight. A final audit of the verified dataset revealed a misclassification rate of only 0.016%, confirming a highly reliable set of gender annotations for our analysis.

The dataset covers a 95-day window from February 23 to May 28, 2025, which includes the official 37-day campaign period from March 23 to April 28. Unlike scheduled elections, this election was called unexpectedly, creating a unique natural experiment for studying rapid political mobilization. The sudden onset provides three methodological advantages: (1) clean temporal boundaries with March 23 serving as a clear intervention point, (2) elimination of gradual pre-election positioning that typically occurs in scheduled elections, and (3) observation of authentic political responses under sudden electoral pressure rather than prescripted strategies.

To build a comprehensive dataset, we employed platform-specific tools and workflows for collecting posts from X/Twitter, TikTok, Instagram, Bluesky, and YouTube, following the methodology described in (Pehlivan et al. 2025). For each platform, all available metadata was preserved in JSON format to ensure structural consistency and completeness. While media files (e.g., images, videos) were not collected directly, links to such content were retained in the metadata.

- **X/Twitter:** A custom scraping library queried the platform's advanced search interface to retrieve posts from candidates and news outlets. This enabled both bi-weekly data collection during the campaign and historical retrieval by specifying custom date ranges.
- **Instagram:** Posts were collected weekly using a custom scraper modeled after (Abrahams 2026).
- **Tiktok:** We employed the reverse-chronological scraper described in (Steel, Parker, and Ruths 2024), running weekly to balance coverage and engagement fidelity. TikTok's Research API does not include Canadian data, making scraping the only viable option.
- **Bluesky:** Candidate posts were retrieved through the official Bluesky API [2], capturing metadata and textual content.
- **YouTube:** Metadata for candidate posts was collected via the YouTube Data API [3], including titles, descriptions, publication timestamps, and engagement statistics (e.g., views, likes).

Data Normalization

Because each platform provides distinct data schemas, we implemented a normalization layer to create a unified textual and metadata representation. All textual content was merged into a single representation, and engagement metrics and structural attributes were standardized to allow comparability across platforms. This ensured consistency while preserving platform-specific richness, enabling seamless cross-platform analysis of political discourse.

The final dataset contains 193,620 posts from 921 unique candidates (610 male, 311 female), across six political parties, active on social media during the campaign period. Table 1 shows the distribution of candidate accounts and posts by platform. A candidate can have accounts on different platforms. Each post in the final dataset is annotated with its assigned topic label and topic model score as explained

[1]https://openai.com/fr-FR/index/introducing-gpt-4-5/

[2]https://docs.bsky.app/

[3]https://developers.google.com/youtube

Platform	# Cand.	# F Cand.	# M Cand.	# Posts	# F Posts	# M Posts
X/Twitter	630	185	445	132071	28545	103526
Instagram	732	269	463	45152	17702	27450
Tiktok	62	27	35	1500	661	839
Bluesky	176	74	102	13506	5362	8144
Youtube	100	36	64	1391	455	936
Total	**1700**	**591**	**1109**	**193620**	**52725**	**140895**

Table 1: Candidates: accounts, and posts across platforms.

in the following sections. To promote transparency and enable reproducibility while adhering to platform terms of service, a dataset containing the post IDs along with the generated annotations is made publicly available. Furthermore, to ensure the direct reproducibility of our Gender Axis Analysis, we also provide a file containing the aggregated data for each candidate: the sum of their post embeddings over the 95-day period and their total post count, which serves as the normalizer. Finally, the Jupyter Notebook containing the full analysis code used to generate the figures and statistical results in this paper is also provided[4].

Analysis Methods

Exploratory Data Analysis (EDA) Our analysis begins with an exploratory data analysis (EDA) to establish descriptive baselines of candidate distributions and posting behavior. Since the dataset is imbalanced—with substantially more male candidates than female, we use mean-normalized measures rather than raw counts to avoid skew. Our analysis includes temporal posting activity (before, during, and after the campaign) as well as engagement metric based on like. These descriptive baselines provide context for subsequent computational analyses of content in the posts.

We define a Gender Bias Score (GBS) to quantify relative posting intensity between male and female candidates within each topic. The score is computed as:

$$GBS = \frac{\text{\# Posts by Male}}{\text{\# Unique Male}} - \frac{\text{\# Posts by Female}}{\text{\# Unique Female}}$$

This formulation normalizes raw post counts by the number of active candidates of each gender, thereby controlling for the numerical imbalance between male and female candidates. Positive values indicate that, on average, male candidates post more frequently, while negative values indicate greater female posting intensity.

Topic Analysis

To examine thematic differences, we first apply BERTopic (Grootendorst 2022), a transformer-based topic modeling method that clusters semantically similar texts. Because the dataset includes content in both English and French, we use the multilingual sentence-transformer model[5] for embedding generation (Reimers and Gurevych 2019). This model supports over 50 languages and provides consistent embeddings across English and French, allowing us to capture cross-lingual topic structures.

It is important to note that a significant portion of posts (58%) were classified as outliers by the HDBSCAN(Campello, Moulavi, and Sander 2013) clustering algorithm within BERTopic. This is a common and expected outcome when analyzing short and noisy social media text. As a density-based algorithm, HDBSCAN is designed to identify and separate such noise from coherent, dense thematic clusters. Rather than forcing these thematically diverse posts into ill-fitting topics, which would compromise the integrity and interpretability of our topics, we treat the large outlier group as a substantive finding in itself, reflecting the fragmented nature of much of the online political discourse. Our subsequent analysis, therefore, focuses on the high-quality, interpretable topics that the model confidently identified.

In addition to thematic content, we analyze the sentiment of each post using a multilingual transformer-based sentiment classifier. Specifically, we employ the multilingual sentiment analysis model (tabularisai et al. 2025) [6] sentiment analysis, which is based on the DistilBERT multilingual model (Sanh et al. 2019) and supports sentiment classification across a wide range of languages, including English and French. The model assigns posts to five sentiment categories (very negative, negative, neutral, positive, very positive), providing a coarse-grained but robust measure of affective tone across languages.

Gender Axis Construction

To measure linguistic bias directly, we construct a gender bias axis in the embedding space. Following established methods in bias detection (Bolukbasi et al. 2016), we compute:

$$V_{\text{gender}} = V_{\text{female avg}} - V_{\text{male avg}}$$

where $V_{\text{female avg}}$ and $V_{\text{male avg}}$ represent the average embeddings of posts authored by female and male candidates, respectively. This axis provides a continuous dimension along which concepts can be projected, allowing us to assess whether political terms, topics, or key expressions align more closely with male- or female-associated discourse. Following prior work on embedding-based bias measurement, we interpret the gender axis as capturing aggregate linguistic tendencies encoded in sentence embeddings rather than isolating individual lexical or syntactic features (Bartl, Ruppenhofer, and Scheible 2020; Bolukbasi et al. 2016).

Statistical Testing

To assess whether observed gendered differences are systematic rather than artifacts of data imbalance, we incorporate statistical testing into our analysis pipeline. For gender axis validation, we compute separation metrics such as the area under the ROC curve (AUC) (Fawcett 2006) and Kolmogorov–Smirnov statistics (Siegel and Castellan

[4]https://github.com/ZeynepP/gendered-political-discourse

[5]https://huggingface.co/sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2

[6]https://huggingface.co/tabularisai/multilingual

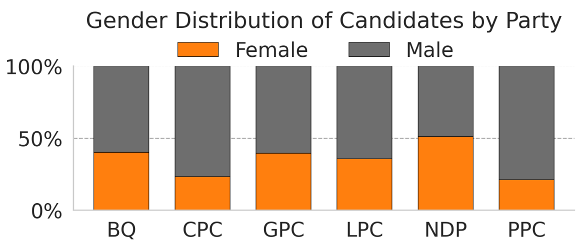

Figure 1: Proportion of candidates gender by party

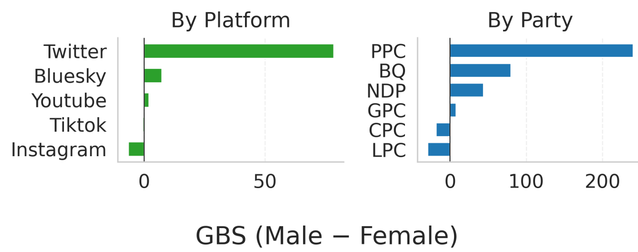

Figure 2: Gender Bias Score (GBS) by platform (left) and political party (right). A positive score indicates higher posting intensity for men.

1988) to evaluate whether male and female candidates' projections are distinguishable beyond chance. For topic-level comparisons, we test whether distributions of male- versus female-authored posts differ significantly using non-parametric Mann–Whitney U tests (Mann and Whitney 1947), which are robust to non-normal and skewed distributions common in social media data. Temporal robustness is evaluated through bootstrap-based permutation tests, which assess whether observed pre- versus campaign shifts in topic centroids could emerge under random partitioning of posts. Finally, when analyzing engagement metrics (e.g., likes), we again employ non-parametric tests suited to heavy-tailed distributions. Together, these procedures ensure that differences identified in our descriptive analyses are statistically robust and not artifacts of data imbalance or skewed measurement distributions.

Reproducibility

The dataset released on Github is organized into clearly documented JSONL files. Candidate-level information includes attributes such as gender and party affiliation, while post-level records contain platform, timestamp, candidate identifier, candidate party, candidate gender, inferred topic label, topic probability, and like count. The accompanying repository includes the complete analysis notebook and intermediate embedding state files required to replicate the gender axis construction and downstream analyses.

All experiments were conducted on the Digital Alliance of Canada research infrastructure (Fir cluster[7]) using a single NVIDIA H100 GPU. Topic modeling, including embedding extraction and clustering—took approximately 3 hours. Sentiment analysis using a multilingual transformer model required approximately 2 additional hours.

Results

Exploratory Data Analysis (EDA)

We begin with descriptive statistics of candidate distributions across political parties. Figure 1 illustrates candidate gender distribution across political parties. The New Democratic Party (NDP) fielded the highest proportion of female candidates, followed by the Greens (GPC), Bloc Quebécois (BQ) and Liberals (LPC), while the Conservative Party (CPC) and the People's Party of Canada (PPC) exhibited considerably lower female representation.

[7]https://docs.alliancecan.ca/wiki/Fir

When it comes to online presence, we note a significant gender bias where male candidates are much more likely to post on Twitter as shown in Figure 2. There is also a moderate gender bias where male candidates are more likely to use Bluesky, while women candidates are more likely to use Instagram.

It is notable that the gap in posting volume between male and female candidates becomes most pronounced during this peak campaign period, suggesting that the pressures of the election disproportionately amplified the posting activity of male candidates before receding in the election's aftermath. On average, women candidates consistently post less content online as compared to their male counterparts, and this difference is more prominent amongst Bloc Quebécois, NDP, and PPC candidates. We note one apparent deviation from this pattern for the Green Party, where female candidates exhibit a brief but pronounced surge in posting activity immediately prior to election day; as we show in the topic and temporal analyses below, this reflects coordinated, campaign-specific mobilization around community-oriented advocacy events (e.g., promotion of the "Vote Palestine" initiative) rather than a persistent difference in baseline communication behavior.

To understand engagement with candidate-generated content, we analyzed the number of likes on original posts (excluding retweets and replies), whose distributions by gender and party are shown in Figure 3, using Mann–Whitney U tests. This exclusion was methodologically necessary because likes on a retweet are attributed to the original post, meaning the like count in a retweet's metadata is typically zero. This refined analysis reveals that the overall median engagement for original content is nearly identical between genders, with male candidates having a slight but statistically significant edge (Male Mdn = 22 vs. Female Mdn = 21, $p < 10^{-4}$).

However, this aggregate parity conceals strong and divergent patterns at the party level. For Conservative, NDP, and PPC candidates, women's original content receives significantly higher median engagement than men's. This effect is most pronounced among Conservatives (Female Mdn = 58 vs. Male Mdn = 53) and the NDP (Female Mdn = 19 vs. Male Mdn = 13). Conversely, the trend reverses for Liberal and Bloc Québécois candidates, where men's original content receives significantly more engagement than women's. This is particularly notable for Liberal candidates (Male

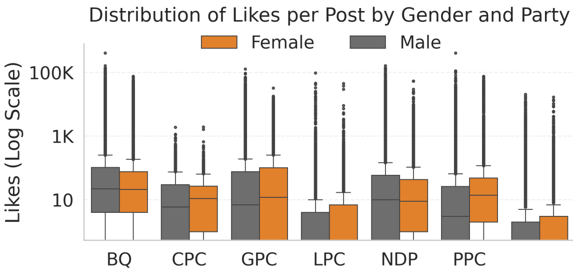

Figure 3: Distribution of likes for original, candidate-generated content by gender and party, plotted on a logarithmic scale. Each box shows the interquartile range (IQR), with the central line indicating the median.

Mdn = 23 vs. Female Mdn = 17). The Green party showed a significant distributional difference in favor of women, despite equal medians.

These findings demonstrate that the dynamics of gendered audience engagement are not monolithic; they are strongly moderated by party affiliation, with different partisan environments appearing to foster greater reception for the original content of either male or female candidates.

Further disaggregating engagement by platform for original, candidate-generated content reveals a consistent pattern. We performed Mann-Whitney U tests on the number of likes per original post and found a statistically significant pro-male engagement gap on four of the five platforms. This gap was most extreme on YouTube, where the median number of likes for a male candidate's post was nearly five times that of a female candidate's (Male Mdn = 73.00 vs. Female Mdn = 15.00, $p < 0.0001$). A more modest, though still significant, gap favouring male candidates was also observed on Twitter (Male Mdn = 15.00 vs. Female Mdn = 12.00), Instagram (Male Mdn = 36.00 vs. Female Mdn = 34.00), and Bluesky (Male Mdn = 7.00 vs. Female Mdn = 5.00).

In stark contrast, Tiktok was the only platform where we found no statistically significant difference in median engagement for original posts (p = 0.0825). This suggests that while a pro-male engagement gap for candidate-generated content may be the norm across most social media, the specific architecture and audience norms of a platform like Tiktok can foster a more equitable dynamic for audience reception.

Topic Analysis: Gendered Thematic Emphasis

To investigate how thematic engagement differs by gender, we examined topic clusters through a dominance framework that controls for data imbalance. Rather than relying on raw post counts—which would overrepresent male candidates due to their numerical advantage—we computed per-candidate posting rates within each topic. Topic coherence scores and the top 10 keywords for each topic are provided as a part of dataset to demonstrate the validity and interpretability of our topics.

The relatively large proportion of posts classified as outliers reflects structural characteristics of electoral social media communication rather than a methodological failure.

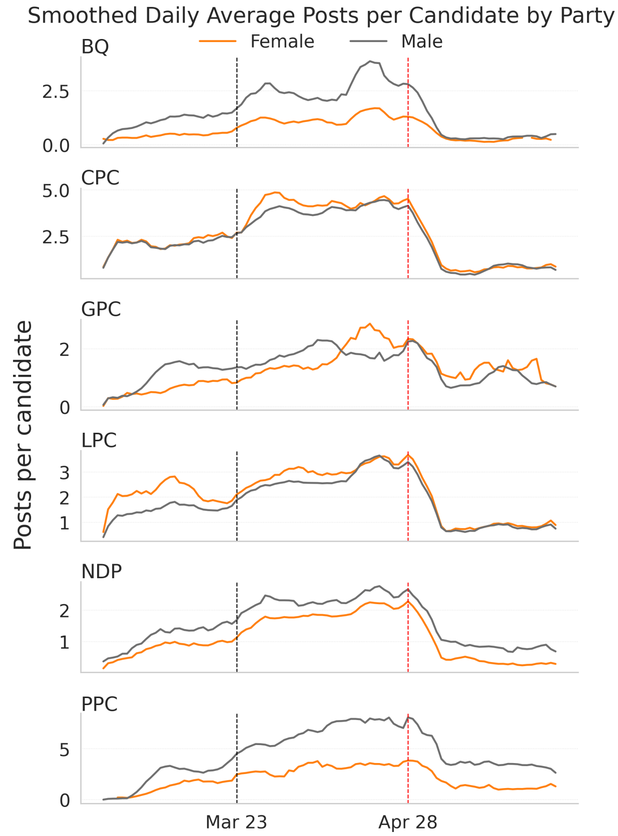

Figure 4: Normalized daily posting activity by gender for each major political party, smoothed with a 7-day rolling average. The y-axis shows the average number of posts per candidate. The vertical dashed lines indicate the start (March 23) and end (April 28) of the official campaign period.

Prior work shows that campaign discourse is often episodic, reactive, and event-driven, consisting of slogans, mobilization cues, and platform-specific interactions rather than sustained thematic discussion (Theocharis and Jungherr 2020; Belcastro et al. 2022). In such settings, density-based clustering methods prioritize the identification of semantically coherent topical cores over exhaustive coverage (Grootendorst 2022). Treating outlier detection as a mechanism for preserving topic coherence therefore aligns with established practice in computational analyzes of electoral discourse. Importantly, the prevalence of outliers does not imply missing linguistic signals, as these posts continue to contribute to embedding-based analyzes that operate independently of topical density (Bolukbasi et al. 2016; Bartl, Ruppenhofer, and Scheible 2020).

Figure 5 visualizes the gender dominance of each candidate topic, sorted by the share of female engagement. To control for the numerical imbalance where male candidates are more numerous, this visualization is not based on raw post counts. Instead, for each topic, we first compute the per-candidate posting rate for each gender. The stacked bars in the figure represent the percentage share of this total normalized activity attributed to women (orange) and men (gray).

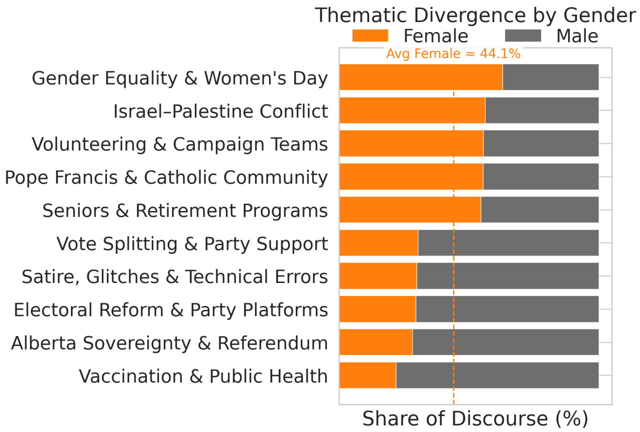

Figure 5: Top 5 topics with the highest female participation and Top 5 topics with the highest male participation, selected from the 50 most active topics. For each topic, gendered participation is computed using average posting rates per candidate within each gender, rather than raw post counts, to account for the male-heavy candidate pool. The dashed vertical line indicates the average female share across the top 50 topics under this per-candidate.

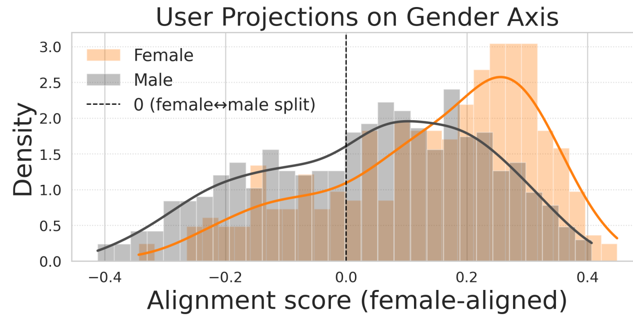

Figure 6: Distributions of aggregate alignment scores on the gender axis for male (gray) and female (orange) candidates.

Female-dominant topics (orange) include areas such as volunteering & campaign teams, dining & local food culture, vaccination & public health, and seniors & retirement programs. In contrast, male-dominant topics (gray) are concentrated around budget and industrial policy, oil, energy & climate, tariffs & trade disputes, and labor unions & worker rights.

This divergence suggests that women candidates engaged more actively in community-centered and social policy discussions, while men disproportionately emphasized economic and industrial issues. While topics like budget & industrial policy were male-dominant, issues such as public health and retirement programs were strongly female-dominant. Gendered topic preferences also vary systematically by party, reflecting the interaction between partisan agendas and gendered communication styles. While the specific issues emphasized differ across parties, the overall pattern of women engaging more in community-oriented discourse and men emphasizing economic and industrial topics remains consistent. The classification of the Israel-Palestine conflict as a female-dominant topic is particularly revealing. The temporal analysis of this topic's linguistic style (Figure 10) provides the clearest evidence that this framing was a campaign-specific strategy. As the figure illustrates, the female-aligned, mobilization-focused discourse was a phenomenon that existed only during the campaign period. Immediately following election day, the topic's linguistic alignment abruptly flips, becoming consistently male-dominant. This sharp reversal suggests the community-action framing was tied to electoral incentives, likely shifting to more abstract policy discussion once the campaign concluded.

These findings align with existing scholarship which views women candidates as mobilizers and strategic communicators (Dittmar 2015; Carroll and Fox 2014). A focus on public health, and seniors demonstrates a relational and care-oriented approach to fostering civic participation. Women may seek to strategically differentiate their campaigns by leveraging the association of women's interests in 'soft politics' for social and gendered issues (Bauer 2015). However, further research is needed to interrogate whether women choose to pursue these topics as a point of community engagement or if they are pushed into advancing these discussions by party officials or due to gendered expectations. Women also discuss the actual election process, such as campaigning and volunteering, more often than men. This enables women to act as mobilizers in connecting with their supporters and community while also asserting legitimacy in their presence in the political space.

By applying a normalized, per-candidate metric rather than raw counts, this analysis highlights structural differences in agenda-setting that would otherwise be obscured by the skewed gender distribution of candidates.

Gender Axis Construction and Descriptive Separation

To operationalize the gender axis, we projected each candidate's aggregate linguistic embedding onto it, yielding a scalar alignment score. The descriptive statistics reveal a clear separation between genders as illustrated in Figure 6. The median alignment score for female candidates (Mdn = 0.18) was substantially higher than for male candidates (Mdn = 0.06). Notably, the 75th percentile for male candidates (0.18) was nearly identical to the median for female candidates, indicating that three-quarters of male politicians used language less aligned with the female-coded pole of the axis than the median female politician.

This separation is statistically significant (Kolmogorov–Smirnov $D = 0.26$, $p < 10^{-12}$). The axis demonstrated fair predictive power, achieving an AUC of 0.67 in distinguishing between male and female candidates. This result is highly robust; a permutation test where gender labels were randomly shuffled yielded a mean AUC of 0.50 (SD = 0.02), confirming that the observed separation is driven by gender-associated linguistic patterns rather than chance. While the distributions show considerable overlap, as expected in a professional context, these results provide strong quantitative evidence that our axis captures meaningful gendered differences in political discourse.

Gender Axis Validation via Sentiment

To further assess the interpretability of the gender axis, we examine its relationship with post-level sentiment. As shown in Figure 7, sentiment varies systematically across quartiles of the gender axis. Posts in the highest quartile (Q4; most female-aligned) exhibit higher sentiment scores than those in the lowest quartile (Q1; most male-aligned), with intermediate quartiles showing a gradual transition rather than a sharp boundary. This pattern indicates a distributional shift in affective tone across the axis, providing external validation that the gender axis captures interpretable stylistic and gendered dimensions of political communication rather than reflecting purely abstract embedding geometry. It demonstrates that gender is consistently aligned with affective tone in political communication, linking gendered norms of performance (Butler 1990) directly with styles of political messaging. This aligns with theoretical accounts of gender as a flexible and interpretative performance, reinforcing prior findings that embedding-based bias axes capture aggregates of correlated linguistic signals rather than direct proxies for specific lexical or grammatical features (Bolukbasi et al. 2016; Bartl, Ruppenhofer, and Scheible 2020; Dolci et al. 2023).

Importantly, the relationship between gender-axis alignment and affective tone is not limited to post-level distributions. Figure 8 demonstrates that the association persists under candidate-level aggregation. Each point represents a candidate's mean gender-axis alignment and mean sentiment across all authored posts, with marker size proportional to posting volume. At the candidate level, sentiment is summarized as the mean of ordinal sentiment scores, following common practice in large-scale computational analyses, and associations are evaluated using rank-based statistics to avoid interval-scale assumptions. The positive association remains clearly visible even after aggregation, indicating that the observed pattern is not driven by a small subset of highly active, emotionally expressive, or viral posts. Rather, it reflects a stable, candidate-level communicative style that systematically links female-aligned linguistic framing with higher affective tone across the corpus.

Party-Level Conditioning of the Gender Axis

In the context of electoral communication, this aggregation is particularly relevant. Campaign discourse often blends policy references with mobilization cues, emotional appeals, and relational framing, making it difficult to isolate individual stylistic dimensions. Accordingly, the gender axis in our analysis should be understood as capturing systematic differences in how candidates frame political communication—such as emphasis on community engagement, mobilization, and interpersonal orientation versus more impersonal or policy-centric framing—rather than as reflecting a single linguistic attribute. This interpretation is consistent with prior computational studies of gendered political language that emphasize framing and communicative style over

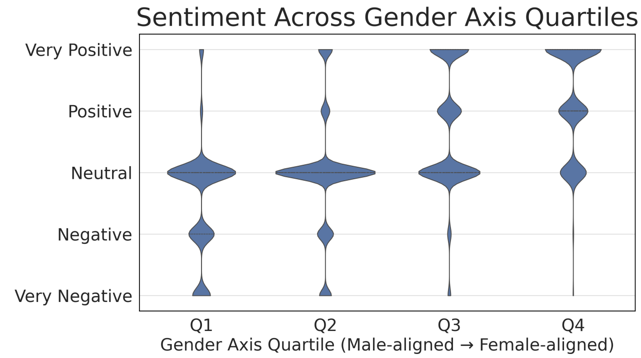

Figure 7: Distribution of post-level sentiment scores across quartiles of the gender axis. Posts more strongly aligned with the female-coded pole of the axis (Q4) exhibit systematically higher sentiment than those aligned with the male-coded pole (Q1).

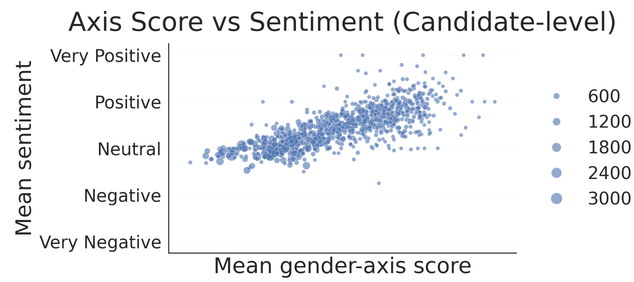

Figure 8: Relationship between candidates' average alignment on the gender axis and their mean sentiment across posts. Each point represents a candidate, with size proportional to posting activity. Positive values indicate more female-aligned linguistic styles. A strong positive association is observed (Spearman $\rho = 0.76, p < 0.001$).

narrowly defined lexical markers (Meeks 2016b; Theocharis and Jungherr 2020).

To investigate the interplay between gender and party affiliation, we disaggregated the alignment scores by political party (Figure 9). The results reveal a remarkably consistent pattern: within every major political party, female candidates exhibit a systematically higher alignment with the female-coded linguistic pole than their male counterparts (AUC $\approx$ 0.70–0.77). This finding demonstrates the robustness of the gendered linguistic gap, which persists across diverse ideological contexts.

Beyond this consistent intra-party gender gap, the analysis also uncovers significant party-level differences in baseline communication styles. For instance, candidates from the NDP and Bloc Québécois, regardless of gender, scored higher on the axis, suggesting their party-wide discourse aligns more closely with female-associated linguistic patterns. Conversely, the People's Party of Canada (PPC) showed a distinct shift toward the male-aligned pole. The fact that the gender gap remains stable even as these party baselines shift confirms that gendered linguistic styles and

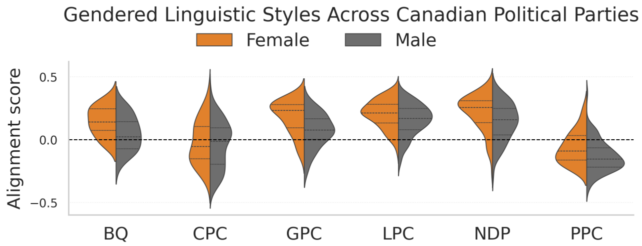

Figure 9: Gendered linguistic styles across Canadian political parties, showing alignment scores for male (gray) and female (orange) candidates.

party ideology operate as two distinct, interacting forces shaping political communication.

Further, political candidates operate in highly professionalized communication environments, where party messaging and media training could plausibly erase gender-based linguistic differences (Malloy 2003; Kam 2009). The fact that the gender gap persists systematically within every party—even under strong ideological and organizational constraints—demonstrates the robustness of gender performance as an independent factor in political communication. This affirms existing literature that differentiates women candidate's communication style from men, where women employ more personable and engaging language whereas men adopt formal and impersonal tones (Meeks 2016a).

A natural question is whether the observed gender axis primarily reflects ideological or topical differences rather than gendered communication styles. While political ideology clearly shapes baseline discourse across parties, our results indicate that ideology alone cannot account for the observed patterns. Across all major parties, female candidates consistently exhibit higher alignment with the female-coded pole of the gender axis than their male counterparts, a within-party separation that persists even under strong ideological discipline. At the same time, party-level shifts in baseline alignment—such as higher overall alignment scores for candidates from the NDP and Bloc Québécois regardless of gender—indicate that ideology and gender function as distinct but interacting forces shaping political communication. This pattern aligns with prior work in political communication, which shows that gendered styles of political expression persist within ideologically constrained environments and cannot be reduced to left–right policy positioning alone (Dittmar 2020; Meeks 2016a).

Temporal Dynamics of Gendered Topics

To analyze how the linguistic style of specific topics evolved, we plotted the daily gender alignment of each topic's centroid over time (Figure 10). This temporal analysis reveals several key patterns. First, some topics exhibit a stable gender alignment throughout the entire period. For instance, discourse around "Volunteering & Campaign Teams" remains consistently and strongly female-aligned, while "Oil, Energy & Climate" remains consistently male-aligned, demonstrating the durability of these linguistic associations regardless of the electoral cycle.

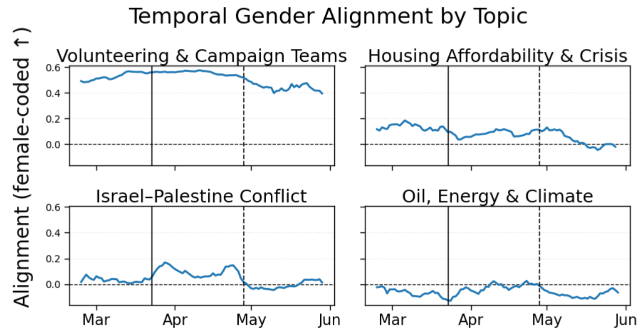

Figure 10: Temporal trajectories of topic-level alignment on the gender axis. Each panel shows the daily centroid alignment of posts within a topic, computed by projecting the average daily embedding onto the gender axis and smoothed using a 7-day rolling mean. Positive values indicate stronger alignment with the female-coded pole of the axis. Solid vertical lines denote the election call (March 23), and dashed vertical lines denote election day (April 28).

Second, the analysis shows that the linguistic framing of certain key issues shifted dynamically during the campaign. The discourse around "Housing Affordability & Crisis," for example, illustrates a complex dynamic. While starting as linguistically neutral, the topic remains so for the initial phase of the campaign before shifting decisively toward the female-aligned pole in early April, a framing that persists until election day before returning to a neutral baseline.

Limitations

Our study has several limitations that define the scope of our findings and point to directions for future research. First, our analysis relies on a binary operationalization of gender. While our hybrid annotation pipeline achieves high accuracy, this framework necessarily excludes non-binary and gender-diverse candidates. Second, our study focuses on the 59% of candidates who were active on social media; accordingly, our findings characterize the discourse of *digitally active* candidates and may not generalize to those who do not use these platforms.

The dataset also exhibits a numerical imbalance between male and female candidates, reflecting the underlying composition of the candidate pool. A potential concern is that such imbalance could artificially exaggerate separation in embedding-based analyses by increasing variance in the majority group. We mitigate this risk in several ways. All analyses are conducted at the candidate level using normalized per-candidate representations rather than raw post counts, preventing highly active candidates from dominating the embedding space. Moreover, the observed gender separation persists consistently within political parties, where sample sizes are smaller and gender distributions more balanced, suggesting that the effect is not driven by global sample size asymmetries. Finally, the gender axis reveals substantial overlap between male and female candidates rather than a bimodal split, indicating that the observed separation reflects systematic stylistic tendencies rather than a mechanical artifact of class imbalance.

Finally, our aggregate analysis smooths over platform-specific dynamics that may shape political communication differently across social media environments. While our analysis aggregates across platforms, differences in platform affordances and audience norms (e.g., TikTok's emphasis on audiovisual engagement versus X's textual discourse) may shape how gendered communication is performed, an avenue we leave for future platform-specific investigation.

Conclusion

This paper sought to understand how male and female candidates use social media differently in a high-stakes election. Our large-scale computational analysis demonstrates that gender plays a persistent and independent role in shaping online political communication during the 2025 federal election. Our findings are threefold. First, we reiterate that there is a statistically robust and measurable difference in linguistic style between male and female candidates. Gendered linguistic styles persist even within strong party discipline and provide powerful empirical support for Butler's theory of gender performativity. Our second key finding builds upon this, highlighting the most powerful insight: the linguistic gap persists systematically within every political party. This provides strong evidence that gender is a robust factor that persists across diverse ideological contexts. Third, we find that thematic divergences exist between women and male candidates - while women focus more on community approaches and social-oriented policies, men focus on economic and industrial issues. Whether this thematic focus reflects a strategic choice to appeal to certain voters, constraints imposed by gendered expectations, or a genuine difference in political priorities is a critical question for future research.

Ultimately, our findings reveal that even in the highly structured world of professional politics, gender remains a fundamental and persistent axis of linguistic differentiation, shaping the very language of democratic discourse in the digital age. This extends prior research which differentiates women candidates' communication styles (Meeks 2016a; Trimble 2017; Goodyear-Grant 2013b) by evaluating the relationship between gendered political communication, party affiliation, and social media engagement.

Building on this study, our future research agenda proceeds in two main directions. First, we intend to extend our analysis to non-electoral periods. By analyzing a multi-year window of everyday political communication, we can establish a more robust baseline against which the specific, high-pressure dynamics of a campaign can be systematically compared. Such a design would enable us to distinguish between campaign-specific effects and more durable, structural patterns in gendered communication.

Our findings document clear gendered patterns, but observational data of this nature cannot definitively disentangle their causal mechanisms. As we posited, these divergences could reflect strategic choice or genuine differences in policy priorities. The temporal findings—where the framing of topics like the Israel-Palestine conflict flips immediately after the election—provide strong support for the role of strategic choice tied to electoral incentives. The consistent thematic focus on social and community policy by women could also reflect genuine differences in political priorities. Disentangling these intertwined forces is a critical task for future research, for which this study provides a necessary empirical foundation.

Ethical Statement

The proposed dataset adheres to the FAIR principles of Findability, Accessibility, Interoperability, and Reusability. The dataset is hosted on Github, where it is assigned a unique identifier to ensure findability and accessibility. The data is accessible in a standardized and well-documented format, enabling interoperability across research tools and platforms.

The dataset was carefully designed with ethical considerations in mind throughout the data collection and processing phases. We only collected publicly available data from X/Twitter, Bluesky, Instagram, TikTok, and YouTube from candidates who are public-facing individuals.

To reduce the potential for misuse of the dataset, we have uploaded it to Github under a license that restricts usage to non-commercial, academic, and research purposes. All of the software tools we've used were used in compliance with their respective licenses.

References

Abrahams, A. 2026. *Social Media Exposed: Track Disinformation, Spot Bots, and Find Out What's Really Happening*. No Starch Press. ISBN 9781718503649.

Asr, F. T.; Mazraeh, M.; Lopes, A.; Gautam, V.; Gonzales, J.; Rao, P.; and Taboada, M. 2021. The Gender Gap Tracker: Using Natural Language Processing to measure gender bias in media. *PLOS ONE*, 16(1): 1–28.

Bartl, M.; Ruppenhofer, J.; and Scheible, S. 2020. Unmasking Contextual Stereotypes: Measuring and Mitigating BERT's Gender Bias. In *Proceedings of the Second Workshop on Gender Bias in Natural Language Processing*, 1–16.

Bauer, N. M. 2015. The effects of counter-stereotypical gender strategies on candidate evaluations. *Political Psychology*, 36(2): 185–201.

Belcastro, L.; Branda, F.; Cantini, R.; Marozzo, F.; Talia, D.; and Trunfio, P. 2022. Analyzing voter behavior on social media during the 2020 US presidential election campaign. *Social Network Analysis and Mining*, 12(1): 83.

Bhardwaj, R.; Majumder, N.; and Poria, S. 2020. Investigating Gender Bias in BERT. *arXiv preprint arXiv:2009.05021*.

Bolukbasi, T.; Chang, K.-W.; Zou, J.; Saligrama, V.; and Kalai, A. 2016. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. arXiv:1607.06520.

Bright, J.; Hale, S.; Ganesh, B.; Bulovsky, A.; Margetts, H.; and Howard, P. 2020. Does Campaigning on Social Media

Make a Difference? Evidence From Candidate Use of Twitter During the 2015 and 2017 U.K. Elections. *Communication Research*, 47(7): 988–1009.

Butler, J. 1990. *Gender Trouble: Feminism and the Subversion of Identity*. Routledge.

Campello, R. J. G. B.; Moulavi, D.; and Sander, J. 2013. Density-Based Clustering Based on Hierarchical Density Estimates. In *Advances in Knowledge Discovery and Data Mining*, 160–172. Springer Berlin Heidelberg.

Carroll, S. J.; and Fox, R. L., eds. 2014. *Gender and Elections: Shaping the Future of American Politics*. Cambridge University Press, third edition.

Chan, E. 2025. *Toxic Monetization and Amplification in Digital Platform Design: The Case of Online Gendered Hate*, 145–171. Cham: Springer Nature Switzerland. ISBN 978-3-031-92670-9.

Dittmar, K. 2015. *Navigating Gendered Terrain: Stereotypes and Strategy in Political Campaigns*. Philadelphia, PA: Temple University Press. ISBN 978-1-4399-1149-5.

Dittmar, K. 2020. Gendered Aspects of Political Persuasion in Campaigns. In *The Oxford Handbook of Electoral Persuasion*. Oxford University Press. ISBN 9780190860806.

Dolci, R.; Colombo, P.; Piantanida, P.; Staerman, G.; et al. 2023. Measuring Gender Bias in Sentence Embeddings. *International Journal of Data Science and Analytics*.

Fawcett, T. 2006. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8): 861–874.

FORCE11. 2020. The FAIR Data principles. https://force11.org/info/the-fair-data-principles/.

Fountaine, S.; Ross, K.; and Comrie, M. 2019. Across the great divide: gender, Twitter, and elections in New Zealand and the United Kingdom. *Communication Research and Practice*, 5(3): 226–240.

Fujiwara, T.; Müller, K.; and Schwarz, C. 2023. The Effect of Social Media on Elections: Evidence from The United States. *Journal of the European Economic Association*, 22(3): 1495–1539.

Gebru, T.; Morgenstern, J.; Vecchione, B.; Vaughan, J. W.; Wallach, H.; Iii, H. D.; and Crawford, K. 2021. Datasheets for datasets. *Communications of the ACM*, 64(12): 86–92.

Gidengil, E.; and Everitt, J. 1999. Metaphors and Misrepresentation: Gendered Mediation in News Coverage of the 1993 Canadian Leaders' Debates. *Harvard International Journal of Press/Politics*, 4(1): 48–65.

Goodyear-Grant, E. 2013a. *Gendered News: Media Coverage and Electoral Politics in Canada*. ISBN 9780774826259.

Goodyear-Grant, E. 2013b. *Gendered news: Media coverage and electoral politics in Canada*. UBC Press.

Grootendorst, M. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*.

Hale, S. A.; John, P.; Margetts, H.; and Yasseri, T. 2014. Investigating Political Participation and Social Information Using Big Data and a Natural Experiment. In *arXiv preprint arXiv:1408.3562*.

Kam, C. J. 2009. *Party Discipline and Parliamentary Politics*. Cambridge University Press.

Kreiss, D.; Lawrence, R.; and Mcgregor, S. 2017. In Their Own Words: Political Practitioner Accounts of Candidates, Audiences, Affordances, Genres, and Timing in Strategic Social Media Use. *Political Communication*, 35: 1–24.

Krook, M. L. 2017. Violence against women in politics. *Journal of Democracy*, 28(1): 74–88.

Kumar, V.; Bhotia, T. S.; Kumar, V.; and Chakraborty, T. 2020. Nurse is closer to woman than surgeon? mitigating gender-biased proximities in word embeddings. *Transactions of the Association for Computational Linguistics*, 8: 486–503.

Lorenz-Spreen, P.; Oswald, L.; Lewandowsky, S.; et al. 2023. A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature Human Behaviour*, 7: 74–101.

Malloy, J. 2003. High Discipline, Low Cohesion? The Uncertain Patterns of Canadian Parliamentary Party Groups. *The Journal of Legislative Studies*, 9: 116–129.

Mann, H. B.; and Whitney, D. R. 1947. On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 18(1): 50–60.

Meeks, L. 2016a. Gender style in the U.S. House: The case of the 112th Congress. *Politics & Gender*, 12(3): 548–570.

Meeks, L. 2016b. Gendered Styles, Gendered Differences: Candidates' Use of Personalization and Interactivity on Twitter. *Journal of Information Technology & Politics*, 13.

Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.; and Dean, J. 2013. Distributed Representations of Words and Phrases and their Compositionality. arXiv:1310.4546.

Nemani, K.; Mittal, A.; et al. 2023. Measuring Gender Bias in Transformer Models: A Survey. *arXiv preprint arXiv:2306.10530*.

Noble, S. U. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press. ISBN 9781479849949.

Pehlivan, Z.; Park, S.; Abrahams, A. S.; Desblancs-Patel, M.; Steel, B. D.; and Bridgman, A. 2025. Building a Media Ecosystem Observatory from Scratch: Infrastructure, Methodology, and Insights. arXiv:2506.10942.

Reimers, N.; and Gurevych, I. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Rheault, L.; Rayment, E.; and Musulan, A. 2019. Politicians in the line of fire: Incivility and the treatment of women on social media. *Research & Politics*, 6: 205316801881622.

Rohrbach, T. 2024. Media-based mechanisms of gendered evaluations of politicians (Dissertation Summary). *Studies in Communication Sciences*, 24: 265–276.

Sanh, V.; Debut, L.; Chaumond, J.; and Wolf, T. 2019. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *ArXiv*, abs/1910.01108.

Siegel, S.; and Castellan, N. J., Jr. 1988. *Nonparametric statistics for the behavioral sciences*. McGraw-Hill, 2nd edition.

Steel, B.; Parker, S.; and Ruths, D. 2024. The Invasion of Ukraine Viewed through Large-Scale Analysis of TikTok.

Sullivan, K. V. R. 2023. "Don't Put Color in Your Hair, Don't Do This, Don't Do That": Canadian Mayors' Mixed Gender Performance on Social Media. *Politics & Gender*, 19(3): 867–890.

tabularisai; Gyamfi, S.; Borisov, V.; and Schreiber, R. H. 2025. multilingual-sentiment-analysis (Revision 69afb83).

Theocharis, Y.; and Jungherr, A. 2020. Computational Social Science and the Study of Political Communication. *Political Communication*, 38(1): 1–22.

Trimble, L. 2017. Still launching a thousand trolls: The enduring gender mediation thesis. In Tremblay, M.; and Praud, J., eds., *Doing politics differently? Women in the Canadian Parliament*, 106–122. UBC Press.

Trimble, L. 2018. *Ms. Prime Minister: Gender, Media, and Leadership*. ISBN 9781442662964.

Vaidhyanathan, S. 2018. *Antisocial Media: How Facebook Disconnects Us and Undermines Democracy*. USA: Oxford University Press, Inc. ISBN 0190841168.

Zhao, J.; Zhou, Y.; Li, Z.; Wang, W.; and Chang, K.-W. 2018. Learning gender-neutral word embeddings. *arXiv preprint arXiv:1809.01496*.

Paper Checklist

1. For most authors...
   (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? Yes
   (b) Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? Yes
   (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? Yes
   (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? Yes
   (e) Did you describe the limitations of your work? Yes
   (f) Did you discuss any potential negative societal impacts of your work? Yes
   (g) Did you discuss any potential misuse of your work? Not explicitly
   (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? Yes
   (i) Have you read the ethics review guidelines and ensured that your paper conforms to them? Yes
2. Additionally, if your study involves hypotheses testing...
   (a) Did you clearly state the assumptions underlying all theoretical results? Yes
   (b) Have you provided justifications for all theoretical results? NA
   (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? Yes
   (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? Yes
   (e) Did you address potential biases or limitations in your theoretical framework? Yes
   (f) Have you related your theoretical results to the existing literature in social science? Yes
   (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? Yes
3. Additionally, if you are including theoretical proofs...
   (a) Did you state the full set of assumptions of all theoretical results? NA
   (b) Did you include complete proofs of all theoretical results? NA
4. Additionally, if you ran machine learning experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? Yes
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? NA
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? NA
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? Yes
   (e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? Yes
   (f) Do you discuss what is "the cost" of misclassification and fault (in)tolerance? Yes
5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets, **without compromising anonymity**...
   (a) If your work uses existing assets, did you cite the creators? Yes
   (b) Did you mention the license of the assets? No
   (c) Did you include any new assets in the supplemental material or as a URL? Yes
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? Yes
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? Partially. The data contains PII (candidate names), which is central to the analysis. The

paper does not explicitly discuss the potential for offensive content, which may be present in retweets or quotes from the public.

(f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR (see FORCE11 (2020))? Yes

(g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset (see Gebru et al. (2021))? No.

6. Additionally, if you used crowdsourcing or conducted research with human subjects, **without compromising anonymity**...

(a) Did you include the full text of instructions given to participants and screenshots? NA

(b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? NA

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? NA

(d) Did you discuss how data is stored, shared, and deidentified? NA