



## METHODOLOGY

### Authors' Note

The Situation Report methodology is an evolving and never-quite-complete document. It is anticipated that there will be month-over-month adjustments to existing measures, the addition of new measures, and refined language to better explain existing methods and measures.

**If you have any questions about the methodology document, would like additional clarification or spot an error, please reach out to the authors at [info@cdmrn.ca](mailto:info@cdmrn.ca).**

1. Introduction	3
2. Purpose	4
3. Research Approach	4
3.1. Social media approach and data	5
3.2. Survey approach and data	6
3.3. Media Monitoring approach and data	6
3.4. Research Limitations/Nature of the Research	7
4. Situation Report Measures	7
4.1. Ecosystem Snapshot	7
4.1.1. Network Visualization	8
4.2. Health of the information ecosystem	9
4.2.1. Vulnerabilities	9
4.2.1.1. Inequality	9
4.2.1.2. Polarization	10
4.2.1.2.1. Segmentation	11

## Canadian Situation Report Methodology – May 2024

4.2.1.2.2. Insularity	12
4.2.1.2.3. Division	13
4.2.1.3. Toxicity	15
4.2.1.3.1. Toxic Speech	15
4.2.1.3.2. Chilled speech	16
4.2.1.3.3. News avoidance	17
4.2.1.4. Trust	18
4.2.2. Threats	20
4.2.2.1. Misinformation	20
4.2.2.1.1. Concern about misinformation	20
4.2.2.1.2. Links to known misinformation websites	21
4.2.2.1.3. Discussion about misinformation	23
4.2.2.2. Foreign influence	23
4.2.2.2.1. Concern about foreign influence	23
4.2.2.2.2. Directed foreign influence	25
4.2.2.3. Artificial Intelligence	26
4.3. State of Engagement with News and Politics	27
4.3.1. News Engagement	27
4.3.1.1. News seeking:	28
4.3.1.2. News sharing:	28
4.3.1.3. News consumption:	29
4.3.2. News outlets	30
4.3.2.1. Inequality among news outlets:	30
4.3.2.2. Top 5 Canadian news outlets:	31
4.3.2.3. Local versus national news engagement:	31
4.3.2.4. Platforms used by news outlets:	31
4.3.3. Politicians	32
4.3.3.1. Top social media platforms	32
4.3.3.2. Engagement with federal political party leadership	33
4.3.3.4. Engagement with elected party members	33
4.3.3.4. Politicians vs. news engagement	33
4.4. Top Stories	34
5. Glossary	35

## 1. Introduction

Every month the [Canadian Digital Media Research Network \(CDMRN\)](#) publishes a Situation Report with the aim to evolve understanding of the overall health of the Canadian Information Ecosystem (CIE) as it relates to politics, media and broader democracy. We do so by monitoring and analyzing social media (Facebook, Instagram, Youtube, Twitter, Telegram, and TikTok) and survey data to characterize the current state of the Canadian CIE and how it evolves over time, and identify potential vulnerabilities and threats that may disrupt it.

Through Situation Reports, we specifically focus on engagement with news and politics online, perception of government and media, and the prevalence of misinformation and foreign influence. We also analyze the role prominent stories play in shaping the ecosystem, and provide insight on the information dynamics of major political and/or large scale events. We report on the same measures every month to provide insight about the evolving health of the CIE. These measures, as well as the scope of data on which they are based, is subject to change as we refine our ability to examine the full ecosystem.

This work is a part of the CDMRN's [Project on Infrastructure Ecosystem Resilience \(PIER\)](#), an initiative of the [Media Ecosystem Observatory \(MEO\)](#) at McGill University and the University of Toronto.

## 2. Purpose

This document provides an overview of the methodology used to collect, analyze and report on each of the measures included in the Situation Reports. We begin by describing our research approach, i.e. how we leverage a unique combination of digital trace (social media analysis), tracking survey and media monitoring to conduct our work. Second, we will explain the nature of our data set, specifically the scope, scale and composition of our trace and survey data. Third, we will share the assumptions made that influence the design and implementation of our research methods. Finally, we will describe the methodology used to calculate each of the measures included in our situation reports.

We hope that this document will assist readers of the Situation Reports learn how to accurately interpret the measures, as well as serve as methodological companion to other researchers interested in studying information ecosystems.

### **3. Research Approach**

Our information gathering methods rely on a multi-pronged approach that leverages empirical data from scientific research surveys, social media activity, and monitoring trends from both mainstream and online media sources. We then analyze and contextualize data gathered from each of these vectors to generate monthly situation reports that focus on three important areas. First, we establish the current state of and potential risks to the overall health of the Canadian information ecosystem, by identifying and describing key vulnerabilities (inequality, toxicity, polarization & insularity, and trust) and threats (misinformation, foreign influence). Second, we illuminate the ways in which people engage online with Canadian news and political content by identifying trends in citizens' appetite for news, engagement with news outlets, and online engagement with politicians.

Each facet of our approach to data collection and measurement is the product of a detailed set of research procedures. Our trace team identifies and adds new accounts to our trace database, analyzes a massive dataset of posts from political, media and influencers across social media platforms. Our survey team generates new questions and updates previous ones based on developing trends and the latest scientific research, ensuring that the most accurate, up-to-date version of our bilingual tracking survey launches successfully every month. Finally, our situation report team tracks current events and analyzes the flow of key events. The details of our three methods are described in detail in the remainder of this section. Some discussion of research limitations and the nature of social media and survey research is included at the end.

### 3.1. Social media approach and data

We collect social media data from platforms used for political discussion in Canada, specifically, Twitter, Facebook, Instagram, Youtube, TikTok, and Telegram. Our core data collection focuses on Canadian entities, including news outlets, elected federal and provincial politicians, influential journalists, commentators, non-profit and advocacy organizations, as well as other accounts that frequently engage with Canadian politics. We also collect data from a set of international entities, including media outlets, politicians, and influencers, to understand how out-of-country entities interact with the Canadian information ecosystem.

This list of politically influential accounts across platforms grows and evolves as we continue to add influential entities in the information ecosystem. It requires continuous pruning and maintenance to capture the bulk of the politically relevant conversation in Canada. As new accounts are added, percentages and numbers in the situation report are likely to fluctuate, but are not anticipated to radically change the overall portrait of the information ecosystem. However, we will note cases where many changes have been made to the list. The current seed list consists of 2,591 distinct entities with a total of 6,058 unique social handles. All entities are classified as a politician, news outlet, journalist, or influencer and we add additional details for each category such as province, whether a news outlet focuses on local or national news, etc.

We collect all posts produced by these politically influential accounts across all activity on the six platforms identified. We extract as much associated metadata as possible: likes, shares, and comment counts; embedded links; uploaded photos; hashtags; mentions; and the URL of the post. From March 26th to April 23rd, our seed list collectively produced 393, 951 posts from March 26th to April 23rd.

For our analysis, we report as if our social media dataset is comprehensive: all measures presuppose that the seed list contains all entities that fit our inclusion criteria and their social media accounts, and that we are not missing any influential members of the information ecosystem. A team of research assistants has manually identified each handle, but true comprehensiveness is elusive. Moreover, as more

accounts are being continuously added, some composition of the social media sample will vary month to month.

### **3.2. Survey approach and data**

To better understand the public opinions and attitudes of Canadians, our situation reports include results from monthly online surveys that we conduct with the help of a commercial polling panel provider. We ask a nationally representative sample of approximately 1,500 Canadians each month to respond to a core set of questions focused on the key topics and themes included in our situation reports. To ensure that the views of Canadians nationwide are reflected accurately in our results, we also weight our data according to 2021 census numbers across region (by province), age and gender.

### **3.3. Media Monitoring approach and data**

Over the course of each month, our Situation Report team regularly monitors the news and social media to identify and track trending true and false stories of the month. We do so to be able to contextualize our findings and also describe specific information trends over the month.

### **3.4. Research Limitations/Nature of the Research**

Our data collection and analysis methods are subject to change. In other words, measures may change slightly between monthly Situation Reports, as they draw on different datasets or analytical methods.

Data for this report is geared towards understanding the online information ecosystem in Canada. For simplicity, we use the term ‘Canadians’ when referring to who we are conducting our research on. However, for a limited number of social media entities tracked, we cannot verify that they are Canadian or located in Canada. Moreover, engagement across social media platforms is not disaggregated at the country level, so engagement measures capture the activity of the global population. All survey respondents are self-reported Canadian citizens.

## 4. Situation Report Measures

NOTE: All images provided in this section from May 2024. They are provided as examples of the types of visuals that are used for each measure. For details about the visual and finding for the month, refer to the [Situation Report](#).

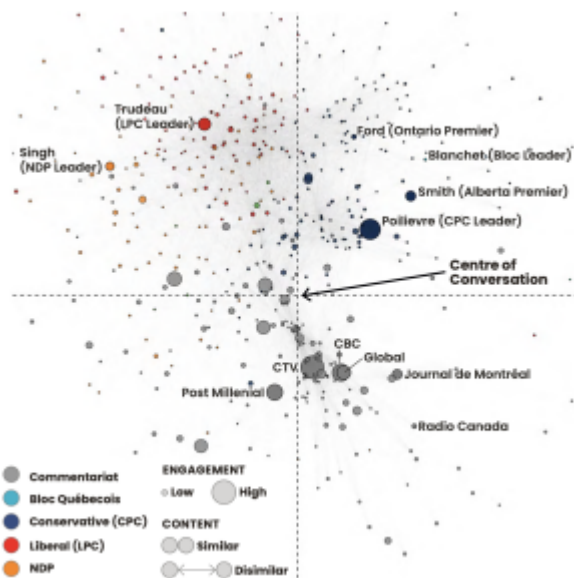
### 4.1. Ecosystem Snapshot

The ecosystem snapshot provides a depiction of the information ecosystem that month. Visualizations will vary from month to month covering the entire ecosystem to a specific story or trend. Infographics will include network maps, as well as many other types of infographics. We explain the methodology and interpretation of each type below as the types of visuals grow over time.

#### 4.1.1. Network Visualization

This ecosystem snapshot (network diagram) provides a 2-dimensional projection of the relationships between key actors in the CIE, with influential political and media entities identified.

This type of snapshot, a **network visualization**, features an information ecosystem network that is constructed using posting behaviour of politically influential entities. When two entities (nodes) share a web link, a hashtag, a mention, or share two posts that are very similar in subject and valence within 24 hours of each other (cosine difference for embeddings using [paraphrase-multilingual-MiniLM-L12-v2](#) of at least 0.75), we form a connection (edge) between them. These



connections are summed (undirected network) and weighted based on the overall number of connections coming from an entity. We then identified the 800 most influential entities, as measured by the strength of their connections to other entities, discarded the bottom 75% connection strengths, and projected the nodes and edges into a 2-dimensional space using a Fruchterman–Reingold layout. The origin point (where the x- and y-axes cross) is set at the overall weighted (by engagement) mean of all node's x and y placements in the layout. The visualization shows just those edges and nodes close to the origin.

## 4.2. Health of the information ecosystem

### 4.2.1. Vulnerabilities

*We evaluate the vulnerability of the information ecosystem by looking at four characteristics. First, **inequality** helps us understand the extent the online conversation in the CIE is skewed towards a minority of voices. Second, **polarization** shows how segregated information flows are within communities. Third, **toxicity** captures how harmful and uninviting the information environment can be to individuals, communities, and organizations. Fourth, we measure **trust** to gauge the extent people are confident information gatekeepers are acting in the best interests of Canadians. Together, these four characteristics can open or restrict information flows, diversify or narrow information sources and beliefs, and influence acceptance or denial of false or accurate information. Most measures are reported on a scale of 0 (low) to 1 (high).*



### 4.2.1.1. Inequality

We evaluate **inequality** using the gini coefficient (the extent influence within an information ecosystem deviates from a perfectly even distribution). The scale ranges from 0 (each entity has equal influence) to 1 (a single entity has all the influence). Fewer information sources (high value) severely restricts information flows, which can lead to faster spread and deeper penetration of misinformation.

Inequality in the information ecosystem refers to how concentrated engagement levels are across social media accounts and whether a small subset of accounts receive a disproportionately large amount of a platform's engagement. This concept highlights how much attention and influence are concentrated among a subset of entities, affecting the diversity and plurality of voices within the ecosystem.

High levels of inequality can lead to a scenario where a small number of accounts dominate the conversation, potentially skewing public perception and discourse. High inequality renders an ecosystem vulnerable to the spread of misinformation, when the validity of or the motives behind these sources is questionable.

Research on inequality often employs metrics such as the Gini coefficient, a widely accepted measure in economics, to quantify the distribution of income within and across countries. The Gini coefficient ranges from 0 to 1, where 0 indicates perfect equality (every entity has equal engagement) and 1 indicates perfect inequality (a single entity has all the engagement). Studies have shown that social media platforms often exhibit high Gini coefficients, reflecting significant disparities in

#### Inequality of the CIE

How skewed is the CIE towards a small set of accounts?



## Canadian Situation Report Methodology – May 2024

attention and influence among users.<sup>1</sup> These online inequality scores are much higher than economic gini values seen today across countries.<sup>2</sup>

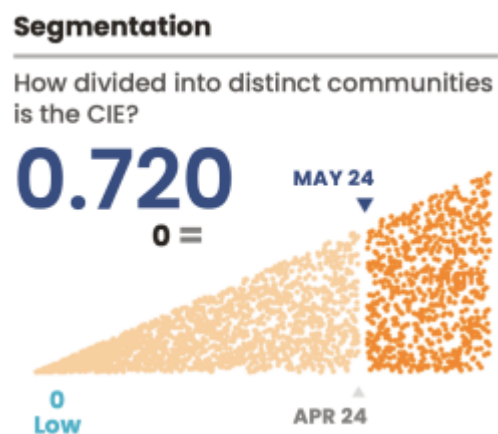
We use Gini coefficients to represent the inequality in the CIE and determine how (un)evenly distributed engagement is across the information ecosystem, specifically across the politically influential individuals and organizations we follow. Gini coefficients are measured on a scale of 0 to 1, where closer to 0 means that each entity has more equal engagement levels and 1 means that engagement levels are concentrated to a single entity.

#### 4.2.1.2. Polarization

We evaluate polarization through **segmentation** (the extent a network is divided into distinct communities, 0=low segmentation and 1=high segmentation), **insularity** (the extent political party families engage with others outside their political party, with a higher number representing a more insular party family) and **division** (how people feel about their own political party versus other parties, 0=low and 1=high).

##### 4.2.1.2.1. Segmentation

Segmentation in the information ecosystem refers to the degree of separation and distinctiveness between different groups or entities in a network, like communities that formed around different political ideologies. Segmentation affects the flow of information within and between different communities and the exposure to diverse and homogenous perspectives. High levels of segmentation can lead to echo chambers, where individuals are



<sup>1</sup> Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. In *Proceedings of the international AAAI conference on web and social media* (Vol. 8, No. 1, pp. 505-514).

<sup>2</sup> Zhu and Lerman, 'Attention Inequality in Social Media'

## Canadian Situation Report Methodology – May 2024

exposed predominantly to information that reinforces their existing beliefs. Meanwhile, low levels of segmentation can promote cross-cutting discussions and a more diverse exchange of ideas, both acting as precursors to democracy and system resilience.

Research on segmentation examines the structure and dynamics of social networks to understand how information flows within and between different groups. Studies frequently employ network analysis techniques to identify clusters or communities within the network. For example, researchers often divide communities across ideological differences, such as left and right communities, to see if different groups communicate distinctively.<sup>3</sup> Part of this distinctiveness in communication across communities is because news organizations are becoming sorted into ideological communities, a process sometimes referred to as political parallelism. These trends of distinct communities with distinct communication patterns extend to social media behaviour and are now a major point of study.<sup>4</sup>

We determine segmentation (or modularity) by generating and interpreting a network modularity score<sup>5</sup> on a scale of 0 to 1. Calculating this score involves four steps: 1) clustering entities (accounts) into communities using Louvain clustering<sup>6</sup>; 2) counting the edges (connections) between these communities; 3) estimating the number of edges entities within a community would have to each other if they were not part of a community, but instead behaved completely randomly; and 4) comparing the actual number of edges inside the communities to the expected number of edges if the network were truly random. 0 implies a complete lack of community structure (complete randomness of interaction) while 1 implies that every account is entirely distinct and never interacts with one another.

---

<sup>3</sup> Rupnik, Grčar, Leban, Petrović, Mozetič, Šubelj. 'Retweet Communities Reveal the Main Sources of Hate Speech'

<sup>4</sup>Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political analysis*, 23(1), 76–91.

<sup>5</sup> Brandes, U., Delling, D., Gaertler, M., Gorke, R., Hofer, M., Nikoloski, Z., & Wagner, D. (2007). On modularity clustering. *IEEE transactions on knowledge and data engineering*, 20(2), 172–188.

<sup>6</sup> Louvain clustering is a method of identifying communities in large networks by identifying small communities and expanding outwards through the network. See: Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10), P10008.

## Canadian Situation Report Methodology – May 2024

## 4.2.1.2.2. Insularity



In an information ecosystem, “insularity” refers to the degree to which communities interact (or don’t interact) with each other. We focus on political parties – specifically, how often politicians engage with politicians belonging to a different party. Analyzing political communities is a useful way of gauging how “antisocial” their members are (i.e. if they are willing to have online discussions with other communities), i.e. how entrenched political communities are, and which groups might be isolating themselves more than others.

Insularity is frequently studied with related concepts like “echo chambers”<sup>7</sup> or “filter bubbles.”<sup>8</sup> Echo chambers are communities within information ecosystems where members only encounter content and other users that agree with or confirm their beliefs. Researchers measure insularity as a way to better understand the formation and maintenance of online political communities and how they contribute to divisiveness and polarization, such as Giglietto et al.’s (2021) examination of partisanship online during the 2018 Italian election<sup>9</sup> and Lynch et al.’s (2017) use of insularity measures to analyze political sentiment after the 2013 military coup.<sup>10</sup>

To calculate, we employ our logic for the segmentation measure with the addition of party labels. We calculate insularity by computing the percentage of Canadian politicians’ connections across the network to other politicians in their same political

<sup>7</sup> Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9), e2023301118.

<sup>8</sup> Fletcher, R. (2020). The truth behind filter bubbles: Bursting some myths. Reuters Institute. <https://reutersinstitute.politics.ox.ac.uk/news/truth-behind-filter-bubbles-bursting-some-myths>

<sup>9</sup> Giglietto, F., Valeriani, A., Righetti, N., & Marino, G. (2021). Diverging patterns of interaction around news on social media: Insularity and partisanship during the 2018 Italian election campaign. In *Disinformation and Data Lockdown on Social Platforms* (pp. 80–99). Routledge.

<sup>10</sup> Lynch, M., Freelon, D., & Aday, S. (2017). Online clustering, fear and uncertainty in Egypt’s transition. *Democratization*, 24(6), 1159–1177. <https://doi.org/10.1080/13510347.2017.1289179>

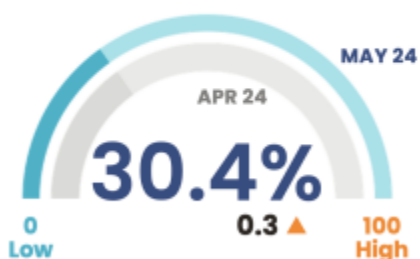
## Canadian Situation Report Methodology – May 2024

party for each of the three main political party families, normalized against the total number of politicians in that party. A higher score indicates that more of the connections for a given political party are to others in their same party, whereas a lower score means the inverse (more connections are to accounts outside their own party). We then report the insularity score for each party. These scores are best understood in relation to one another (i.e., how the parties compare in insularity). Parties that are less insular are closer to the core of the national conversation, have a less cohesive voice, and are more likely to engage with a variety of topics and individuals outside their immediate political group.

#### 4.2.1.2.3. Division

##### Division

How do people feel about their own political party versus other parties?



Polarization generally refers to the extent to which opinions, beliefs, and attitudes diverge sharply between different groups. Affective polarization, or what we call *division*, specifically measures the difference in how individuals feel about members of other political parties compared to their own. High levels of polarization can undermine social cohesion, disrupt constructive political discourse, and enable social unrest.

Research on polarization often focuses on understanding emotional and attitudinal divides within political landscapes. Studies frequently measure party favorability to capture the extent of animosity between supporters of different parties.<sup>11</sup> It is particularly challenging to assess in multiparty systems due to the complexity of inter-party relationships.<sup>12</sup>

We measure animosity by assessing the difference in feelings (sentiment) towards members of other political parties compared to one's own party. Using a method designed for multiparty systems, our calculation uses a questions that asks

<sup>11</sup> Merkley, E. (2023). Mass Polarization in Canada: What's Causing It? Why Should We Care? *Canadian Commission on Democratic Expression*. <https://www.mediatechdemocracy.com/all-work/mass-polarization-in-canada-whats-causing-it-why-should-we-care>

<sup>12</sup> Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral Studies*, 69, 102199.'

## Canadian Situation Report Methodology – May 2024

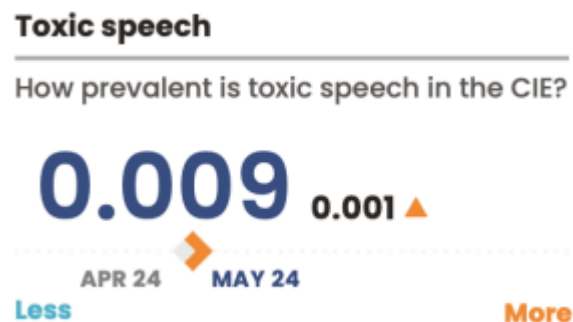
respondents to say how they feel about each political party in Canada (rated from 0 to 100), e.g. an individual demonstrating high favourability for the NDP party would rate them closer to 100 and an individual with low favourability for the NDP party would rate them closer to 0. We calculate average feeling towards political parties for each survey respondent and then calculate an individual animosity score based on the difference between their average score and each political party, with each political party weighted as per Wagner (2021).<sup>13</sup> We then average all respondent's differences to generate an overall score of animosity for the population. So societies where people rate parties closer together have lower scores. We account for the six major political parties: the Liberal Party, Conservative Party, New Democratic Party (NDP), Bloc Québécois, Green Party, and People's Party of Canada (PPC).

#### 4.2.1.3. Toxicity

We evaluate toxicity through **toxic speech** (presence of toxic speech among posts by political influencers, with lower values indicating lower toxicity), **chilled speech** (extent people avoid online political discussion and opinion sharing, 0=low to 1=high), and **news avoidance** (extent people avoid the news all together, 0=low to 1=high).

##### 4.2.1.3.1. Toxic Speech

Toxicity in the information ecosystem refers to the presence of offensive, insulting, threatening, or attacking language in online discourse. High levels of toxicity can contribute to a hostile and divisive online environment, discourage constructive dialogue, and negatively impact mental health and community cohesion. Monitoring toxicity enables us



<sup>13</sup> Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral Studies*, 69, 102199.

## Canadian Situation Report Methodology – May 2024

to gauge how hostile the internet is for users and the risk of downward decline and to develop interventions to promote more respectful and inclusive online interactions.

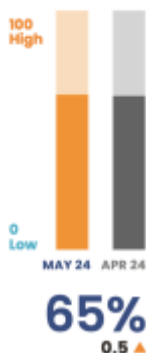
Research on toxicity often involves the use of automated text analysis tools to detect and measure harmful language in online content. For example, machine learning can detect online harassment through using what are called classifiers to identify toxic language.<sup>14</sup> Given the scope of online social media platforms, machine learning technologies are essential for large-scale analysis of social media data.

We measure toxicity in the CIE by assessing the extent “toxic” language is present that month. We do so by assigning each post in our dataset a “toxicity score” along a scale of 0 to 1. A low toxicity score (0) suggests a low frequency of offensive or insulting content and a high score (1) implies a high amount of toxic language is used. To do so, each post is processed by a [supervised toxicity classifier](#) called, *detoxify*, which evaluates the frequency of “toxic” (offensive, insulting, threatening, or attacking) language in that post. We then aggregate all the individual scores into one average measure along the same scale of 0 to 1. We calculate the mean toxicity scores for all posts to report a toxicity score for the whole CIE in that month.

#### 4.2.1.3.2. Chilled speech

##### Chilled Speech

To what extent do people avoid political discussion and opinion sharing?



Chilled speech in the information ecosystem refers to the extent to which individuals feel uncomfortable or reluctant to engage or participate in the online political dialogue. A high-level of chilled speech can suppress diverse viewpoints and curtail natural choke points for mis/disinformation, reduce political engagement, and undermine democratic discourse. It can ultimately lead to a less vibrant and inclusive public sphere.

Research on chilled speech often explores the factors that discourage people from expressing their views online, such as fear of harassment

<sup>14</sup> Wulczyn, 'Ex Machina: Personal Attacks Seen at Scale. Proceedings of the 26th International Conference on World Wide Web.'

## Canadian Situation Report Methodology – May 2024

or political backlash.<sup>15</sup> Studies have shown that environments perceived as hostile or overly contentious (high toxicity explained above) can lead to self-censorship and reduced willingness to engage in discussions. For example, perceived hostile media environments can lead to the silencing of minority opinions, highlighting the impact of a chilling effect on public discourse.<sup>16</sup>

We measure chilled speech by assessing how comfortable people say they are with participating in online conversations about their political opinions. Our measure combines survey questions that ask respondents how comfortable they are sharing their political views on a topic and how willing they are to engage with people who have differing views.

**Survey Question:** *Please indicate how comfortable you are expressing your political opinions online.*

- Not at all comfortable**
- Not very comfortable**
- Somewhat comfortable
- Comfortable
- Extremely comfortable

**Survey Question:** *Please indicate how comfortable you are directly responding to online political news content and social media posts that you disagree with.*

- Not at all comfortable**
- Not very comfortable**
- Somewhat comfortable
- Comfortable
- Extremely comfortable

---

<sup>15</sup> Enock, Stevens, Bright, Cross, Johansson, Wajcman, and Margetts. 'Understanding gender differences in experiences and concerns surrounding online harms: A short report on a nationally representative survey of UK adults'

<sup>16</sup> Hayes, Scheufele, and Huge, 'The Impact of Hostile Media Perceptions on Willingness to Speak Out: Revisiting the Spiral of Silence. Communication Research.'



## Canadian Situation Report Methodology – May 2024

For both chilled speech questions, we calculate the proportion who say “Not at all comfortable” or “Not very comfortable” and then average the proportions to generate an overall chilled speech measure.

#### 4.2.1.3.3. News avoidance

News avoidance refers to the tendency of individuals to deliberately avoid political & public affairs news stories. High news avoidance can lead to a less informed society, higher exposure to misinformation through unvetted sources, and a less resilient public discourse.

Understanding this measure enables us to identify barriers to news engagement and developing strategies to encourage more active participation in news consumption.

Research on news avoidance explores the reasons why individuals might choose to disengage from news, such as perceived negativity, information overload, or distrust in media sources.<sup>17 18</sup> Studies have shown that news avoidance can be a coping mechanism to reduce stress and anxiety associated with negative news.<sup>19</sup>

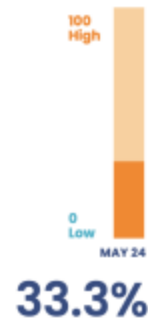
To measure news avoidance, we ask survey respondents:

**Survey Question:** *Whenever you are online, do you find yourself trying to avoid the news as much as possible or more interested in seeking it out?*

- Constantly avoiding**
- Often avoiding**
- Occasionally avoiding**
- I don't try to avoid it, but I don't seek it out either*
- Occasionally seeking out*

#### News Avoidance

To what extent do people avoid the news?



<sup>17</sup> Zhang, Akhter, Nassani, Haffar, 'Impact of News Overload on Social Media News Curation: Mediating Role of News Avoidance'.

<sup>18</sup> Rosenblum, Muirhead, Levendusky, Druckman, Fein, Leeper, Jamieson, Cappella, Stroud, Gentzkow, Shapiro, Prior, "I'm Not Sure What to Believe": Media Distrust and Opinion Formation during the COVID-19 Pandemic'

<sup>19</sup> Skovsgaard and Andersen, 'Conceptualizing and Measuring News Avoidance: Towards a Shared Understanding of Different Approaches and Terminologies. Journalism Studies.'

## Canadian Situation Report Methodology – May 2024

- Often seeking out
- Constantly seeking out

We report the proportion of respondents who say “Constantly avoiding”, “Often avoiding”, and “Constantly avoiding”.

#### 4.2.1.4. Trust

We evaluate trust by asking Canadians about their confidence in **information gatekeepers** (journalists, media organizations, politicians, and big tech).

Trust in the information ecosystem refers to the confidence that individuals have in the information they consume and in the sources providing that information. Trust is crucial because it affects how people perceive and interact with news and information. High levels of trust in the information ecosystem promote civic engagement.<sup>20</sup> Conversely, low levels of trust can lead to apathy and cynicism,<sup>21</sup> ultimately undermining the democratic process by eroding the public's ability to make well-informed decisions.

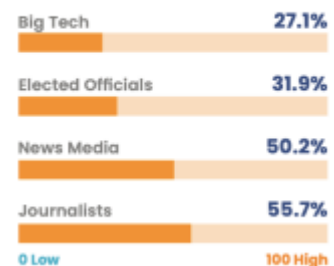
##### Trust in Information Gatekeepers

How much do people trust the information gatekeepers?



##### Trust

How much do people trust each type of gatekeeper?



<sup>20</sup> Putnam, R. D. (1994). What makes democracy work?. *Review-Institute of Public Affairs*, 47(1), 31.

<sup>21</sup> Pinkleton, B. E., Austin, E. W., Zhou, Y., Willoughby, J. F., & Reiser, M. (2012). Perceptions of news media, external efficacy, and public affairs apathy in political decision making and disaffection. *Journalism & mass communication quarterly*, 89(1), 23-39.

## Canadian Situation Report Methodology – May 2024

Studies show higher levels of political trust lead to greater civic engagement and better-informed citizens.<sup>22</sup> Researchers often rely on standardized survey instruments and scales to measure trust. For instance, the Reuters Institute Digital News Report frequently measures trust in news media using surveys across different countries, providing comparative insights.<sup>23</sup>

To measure trust, we ask survey respondents:

**Survey Question:** *How much trust, if any, do you have in [Journalists/News Media/Big tech/Elected officials] to act in the best interests of Canadians?\**

- No trust at all
- Very little trust
- A fair amount of trust**
- A lot of trust**

For each group, we report the proportion that indicate they have a fair or a lot of trust. We then average these values across the four groups to generate an overall score of trust in what we call the ‘Information Gatekeepers’.

### 4.2.2. Threats

*We assess threats to the CIE through evaluating the presence of and concern regarding **misinformation, foreign influence and AI-generated content**. Each of these threats can distort the conversation, mislead and manipulate Canadians, and decrease our collective ability to understand and respond to challenges and opportunities.*

<sup>22</sup> Torney-Purta, J., Henry Barber, C., & Richardson, W. K. (2004). Trust in government-related institutions and political engagement among adolescents in six countries. *Acta Politica*, 39(4), 380-406.’

<sup>23</sup> Reuters Institute. Digital News Report 2023.

<https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023/dnr-executive-summary>

### 4.2.2.1. Misinformation

We evaluate the threat of misinformation through measuring Canadians' **concern** about misinformation (0=low to 1=highly concerned), the extent **misinformation links** are shared across social media platforms by politically influential voices (0=little to 1=high sharing), and public **discussion about misinformation** and associated information phenomena (0=low to 1=high discussion).

#### 4.2.2.1.1. Concern about misinformation



We assess how concerned Canadians are about misinformation, particularly as it affects democracy. We are interested in the level of concern about misinformation for two reasons. First, this measure allows us to gauge capacity of the CIE to manage misinformation – if Canadians are highly concerned about misinformation, then they may be more likely to carefully review sources of information and look out for coordinated disinformation campaigns.

Second, although this measure doesn't tell us about the scale and nature of misinformation, it allows us to evaluate the perceived problem of misinformation in Canada. However, concern about misinformation may not always be a good thing, particularly in large doses. If Canadians are increasingly paranoid about the circulation of misinformation online, then they may begin to distrust otherwise reliable sources like reputable news organizations, or accuse true information of being false.

Many scholars use perceptions of misinformation prevalence to characterize the threat of misinformation. For example, Vegetti and Moncosu (2022) measure the perception of misinformation in 27 European countries to explore what drives

## Canadian Situation Report Methodology – May 2024

concerns about fake news.<sup>24</sup> Altay and Acerbi (2023), in their survey conducted in the UK and the US, find that people who strongly believe misinformation is a severe problem are more likely to share alarmist, misleading information about misinformation.<sup>25</sup>

To measure misinformation concern, we ask survey respondents:

**Survey Question:** *How concerned are you about online misinformation?*

- Not a problem*
- A minor problem*
- A moderate problem*
- A serious problem***
- A very serious problem***

We report the proportion of respondents who say “a serious problem” or “a very serious problem”.

#### 4.2.2.1.2. Links to known misinformation websites

We measure how much misinformation is circulated in the CIE by politically influential accounts. We do not measure how much misinformation *in total* there is in the CIE – that would require us to analyze every single post across multiple platforms in Canada. Rather, we focus on the volume of links to known misinformation sites shared by prominent accounts, providing insight into how prominent misinformation is online in Canada. By looking at how often the most prominent voices in the CIE share misinformation, and how much engagement is generated by that misinformation, we can begin to understand how much



<sup>24</sup> Vegetti, F., & Mancosu, M. (2022). Perceived Exposure and Concern for Misinformation in Different Political Contexts: Evidence From 27 European Countries. *American Behavioral Scientist*, 0(0). <https://doi-org.proxy3.library.mcgill.ca/10.1177/00027642221118255>

<sup>25</sup> Altay, S., & Acerbi, A. (2023). People believe misinformation is a threat because they assume others are gullible. *New Media & Society*, 0(0). <https://doi-org.proxy3.library.mcgill.ca/10.1177/14614448231153379>

## Canadian Situation Report Methodology – May 2024

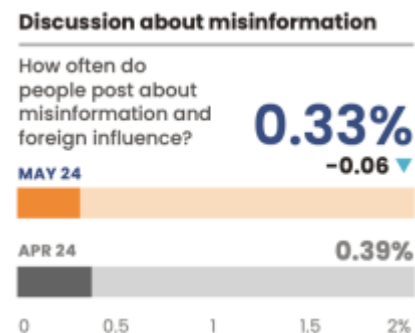
misinformation may be circulating in Canada, from whom, and how much attention that misinformation is getting.

The mechanisms that determine the circulation of misinformation online is a prominent topic among information scholars. Humprecht et al. (2021) have found that the factors that shape the likelihood that someone will spread misinformation are highly country-specific.<sup>26</sup> In other words, the dissemination of (mis)information in the Canadian information ecosystem is probably different from how (mis)information is spread among, for example, the Australian or Belgian information ecosystems. It is therefore important to look at prominent Canadian accounts to understand what drives misinformation here.

We compute this measure by looking at the prevalence of web linking to misinformation websites. We extract the web links for all posts, expand any that have been shortened by url shorteners, and clean all references to common websites are removed (e.g. youtube.com, twitter.com, facebook.com, instagram.com, t.me, google.com, whatsapp.com, apple.co). We then calculate prevalence by identifying all posts in our dataset that contain the URL of a [Global Disinformation Index](#) catalogue of known misinformation websites. We report this as a percentage of how many posts that month contained a link to a known misinformation site as well as the percentage of engagement with posts that contain any URL represented by posts with the URL of a known misinformation site. This is calculated by summing the total engagement received by posts with a misinformation site URL, then dividing this by the total sum of engagement received by posts with any URL.

#### 4.2.2.1.3. Discussion about misinformation

We analyze discussion about misinformation to assess the level of attention it is getting. Examining how often prominent accounts within the CIE provides an additional dimension to our analysis of concern



<sup>26</sup> Humprecht, E., Esser, F., Aelst, P. V., Staender, A., & Morosoli, S. (2023). The sharing of disinformation in cross-national comparison: analyzing patterns of resilience. *Information, Communication & Society*, 26(7), 1342–1362. <https://doi.org/10.1080/1369118X.2021.2006744>

## Canadian Situation Report Methodology – May 2024

about misinformation. Specifically, we measure the frequency of discussion about misinformation, disinformation, foreign interference, fake news, and deep fakes by counting how many times these and related terms, in English or French, appear in posts in the CIE during that month. We report this measure as a percentage of how many posts mention misinformation and/or related terms out of all posts from that month as an indicator of attentiveness to misinformation.

#### 4.2.2.2. Foreign influence

*We evaluate the threat of foreign influence through measuring Canadians' **concern about foreign influence**, and by evaluating the extent overt foreign influencers (China, Russia and India) are influencing the Canadian conversation (**directed foreign influence**).*

##### 4.2.2.2.1. Concern about foreign influence

We measure how concerned Canadians are about foreign interference in politics by asking them if they believe that foreign governments are influencing Canadian politics and how concerned they are about it. We can then track increasing or declining concern about foreign interference, allowing us to inform the national discussion about foreign influence.



Public concern about foreign interference can have various effects. Chow and Levin (2024) found that allegations of foreign interference in protest movements can significantly delegitimize the movement in the eyes of the public, even when there are no foreign actors involved.<sup>27</sup> This can be particularly harmful for protest movements and other activities that heavily rely on public perception and opinion – such as election campaigns. On the other hand, Manwaring and Holloway (2022), argue that public concern can actually increase

<sup>27</sup> Chow, W. M., & Levin, D. H. (2024). Muddying the Waters: How Perceived Foreign Interference Affects Public Opinion on Protest Movements. *American Political Science Review*, 1–18. doi:10.1017/S0003055424000327

## Canadian Situation Report Methodology – May 2024

resilience to foreign interference in democracies by increasing public awareness of disinformation and motivating citizens to expand their cyber skillset.<sup>28</sup>

To measure concern about foreign influence, we ask survey respondents:

**Survey Question:** *How concerned are you about foreign governments influencing Canadian media and politics?*

- Not a problem*
- A minor problem*
- A moderate problem*
- A serious problem***
- A very serious problem***

We report the proportion of respondents who say “a serious problem” or “a very serious problem.”

#### 4.2.2.2. Directed foreign influence

We also look at how connected prominent foreign accounts are to the CIE. We specifically focus on India, Russia, and China, as these countries are often invoked in conversations about foreign interference in Canada. Rather than interference, we focus our lens on foreign influence. We assess how ‘close’ foreign accounts are to Canadian accounts in regards to the content they share: we can understand these accounts to be in close proximity to the CIE if they frequently post similar links or discuss certain topics at the same time and in the same way. This measure, therefore, does not describe if or how foreign governments are influencing Canadian perception – just points out the overlapping parts of our digital information ecosystems. However, in the case that a foreign government does



<sup>28</sup> Manwaring, R., & Holloway, J. (2023). Resilience to cyber-enabled foreign interference: citizen understanding and threat perceptions. *Defence Studies*, 23(2), 334–357. <https://doi.org/10.1080/14702436.2022.2138349>



## Canadian Situation Report Methodology – May 2024

interfere in a Canadian election, we will hopefully be able to identify the campaign, inform the public, and defend Canadian electoral integrity by regularly analyzing the connection of accounts from potential adversaries to the CIE.

A common method for detecting instances of foreign interference or influence in information ecosystems is textual analysis. For instance, Lopez and Madhyastha (2021) examine foreign disinformation on Twitter during the 2016 US Presidential Election by analyzing word usage and other textual characteristics.<sup>29</sup> Smith et al. (2020) developed a system to automatically detect coordinated disinformation campaigns, relying on natural language processing, machine learning, graph analytics, and network causal inference.<sup>30</sup>

We examine the proximity of prominent foreign accounts to the COE. We track a core set of accounts associated with foreign entities (India, Russia, China) with a record of influencing Canadian politics, then integrate those accounts into the web linking, hashtag, mentions, and textual similarity network model described above in 4.1.1. We then compute the percentage of all connections between Canadian and these foreign entities. A higher percentage is indicative of common posting behaviour between Canadian entities and foreign ones.

#### 4.2.2.3. Artificial Intelligence

*We assess the threat of generative AI by measuring **concern** about AI generated content misleading the general public.*

We further investigate Canadians' perception of misinformation by asking survey respondents about how concerned they are about the threat of artificial

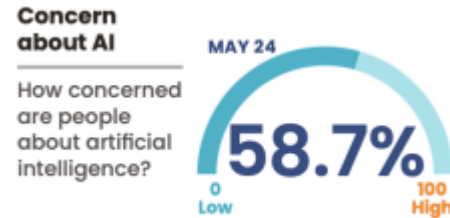
---

<sup>29</sup> López, J. A. D., & Madhyastha, P. (2021). A focused analysis of twitter-based disinformation from foreign influence operations. In *Proceedings of the 1st International Workshop on Knowledge Graphs for Online Discourse Analysis (KnOD 2021) co-located with the 30th The Web Conference (WWW 2021)* (Vol. 2877). CEUR Workshop Proceedings.

<sup>30</sup> Smith, S. T., Kao, E. K., Mackin, E. D., Shah, D. C., Simek, O., & Rubin, D. B. (2021). Automatic detection of influential actors in disinformation networks. *Proceedings of the National Academy of Sciences*, 118(4), e2011216118.

## Canadian Situation Report Methodology – May 2024

intelligence-generated content to Canadian democracy. Through the use of generative AI tools, people can create images and videos of whatever they like, including content featuring politicians or other public figures. These “deepfakes” often look real unless they are closely inspected, and can lead people to believe things about politicians that are not actually true but shape their political opinions nonetheless.



To measure concern about AI, we ask survey respondents:

**Survey Question:** *How concerned are you about AI-generated content misleading the public?*

- Not a problem*
- A minor problem*
- A moderate problem*
- A serious problem***
- A very serious problem***

We report the proportion of respondents who say “a serious problem” or “a very serious problem”

### 4.3. State of Engagement with News and Politics

We report on the state of engagement with news and politics to provide insight into social media preferences and behaviours for news and politics consumption and dialogue. We evaluate **news** by looking at news seeking and sharing to capture the appetite for news, regional preferences for news consumption, and engagement in political online conversation. Second, we characterize engagement with **news outlets**, by assessing inequality, top outlets and platform preferences to identify dominant news sources and the main mechanism for engagement. Finally, we examine online engagement with **politicians** by looking at relative engagement with party leaders and politicians, as well platform preferences. We do so to identify

## Canadian Situation Report Methodology – May 2024

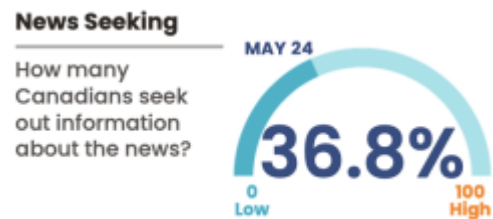
where most of the political discussion takes place (on which platform) and who leads this conversation. In a healthy ecosystem, we would expect to see a high level of news seeking, across a variety of sources, and the free flow of ideas and commentary on the political and media landscape.

### 4.3.1. News Engagement

We evaluate engagement with news using six indicators: 1) **news seeking** (the extent Canadians seek out news); 2) **news sharing** (the frequency of sharing news and opinions on social media); and 3-6) **news consumption types** (the extent Canadians consume news from social media, local, national and international news outlets.)

#### 4.3.1.1. News seeking:

To measure news seeking behaviour, we ask survey respondents:



**Survey Question:** Whenever you are online, do you find yourself trying to avoid the news as much as possible or more interested in seeking it out?

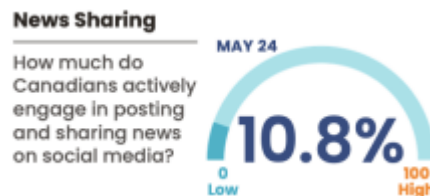
- Constantly avoiding
- Often avoiding
- Occasionally avoiding
- I don't try to avoid it, but I don't seek it out either
- Occasionally seeking out**
- Often seeking out**
- Constantly seeking out**

We report the proportion of respondents who say “Occasionally seeking out”, “Often seeking out”, and “Constantly seeking out”.

## Canadian Situation Report Methodology – May 2024

**4.3.1.2. News sharing:**

We measure news sharing through two separate questions and then combine the scores to generate a news sharing score.



We ask respondents:

**Survey Question:** *In the past month, how often have you shared a news story about politics on social media?*

- Never*
- Once or twice this month*
- Several times this month*
- Once or twice a week***
- Several times a week***
- Once or twice a day***
- Several times a day***

**Survey Question:** *In the past month, how often have you posted your own political views on social media?*

- Never*
- Once or twice this month*
- Several times this month*
- Once or twice a week***
- Several times a week***
- Once or twice a day***
- Several times a day***

For both news sharing questions, we calculate the proportion who say “Once or twice a week” or more and then average the proportions to generate an overall news sharing measure.

## Canadian Situation Report Methodology – May 2024

**4.3.1.3. News consumption:****News Consumption**

Where do Canadians get their news?



We measure the consumption rate of news across four different types of medium; national, international, and local news websites, as well as through social media. For each type, we ask if they use each of the types of online information sources. For people that indicate they do, we follow up by asking if they use these sources specifically for politics and current affairs.

To measure news consumption, we ask:

**Survey Question:** When you use [National news/International news/Local news/Social media websites], how often do you read, watch, or listen to politics and current affairs content, specifically?

- Never
- Once or twice this month
- Several times this month
- Once or twice a week**
- Several times a week**
- Once or twice a day**
- Several times a day**

We calculate the proportion who say “once or twice a week” or more to generate a news consumption score for each news medium.

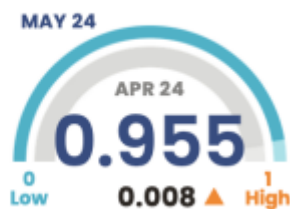
### 4.3.2. News outlets

We evaluate four characteristics of Canadian news outlets: 1) **inequality** of engagement with Canadian news outlets; 2) **top outlets** (the scale and nature of engagement with Canadian news outlets); 3) **national versus local news engagement** (preferences for local versus national news providers); and 4) **social media platform usage** trends by Canadian news outlets.

#### 4.3.2.1. Inequality among news outlets:

##### Inequality of news outlets

How skewed is the CIE to a small number of news outlets?



We measure the unequal distribution of engagement (as measured by the total number of likes, comments, and shares) among Canadian news outlets online by calculating their Gini coefficient. We do this in the same way as we calculate the Gini coefficient for the whole ecosystem (Section

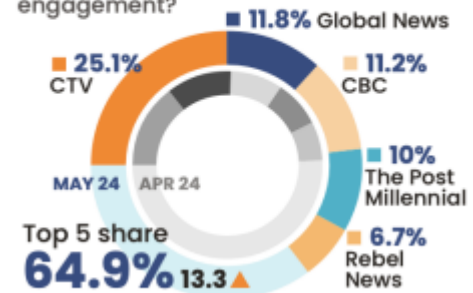
4.1.1.1), but only look at engagement with Canadian news outlet accounts. Similarly, a coefficient close to 1 implies that a small number of Canadian news outlet accounts are grabbing the majority of engagement generated by Canadian news online.

#### 4.3.2.2. Top 5 Canadian news outlets:

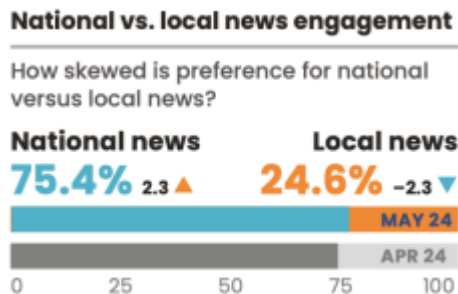
We report the percentage of Canadian news engagement received by only the 5 most popular outlets. We first identify the five outlets that received the highest amount of engagement (as measured by the total number of likes, comments, and shares on their posts), sum them, then divide their engagement by the total amount of engagement received by Canadian news outlets online and multiplying by 100. We are then able to see how much of online Canadian news engagement is generated by the biggest outlets.

##### Top-5 Canadian news outlets

What are the top 5 Canadian news outlets? What is their share of total engagement?



#### 4.3.2.3. Local versus national news engagement:

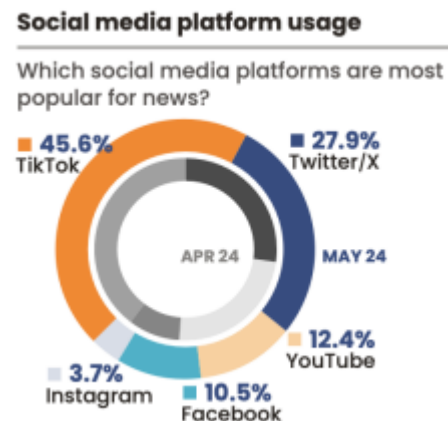


We report the percentage of all engagement with news outlet accounts gained by local news outlet accounts. All the accounts of Canadian news outlets in our dataset are categorized as either “national” or “local” level outlets, so we can compare trends among the two levels. We calculate the relative local news engagement by dividing the total amount of engagement generated by local news accounts across our

platforms by the total amount of engagement generated by all news accounts.

#### 4.3.2.4. Platforms used by news outlets:

We evaluate on which platforms Canadian news outlets receive the most engagement to better understand where Canadians seek their online news content. We calculate the platform with the most engagement with Canadian news outlets by summing the total number of likes on posts by Canadian news outlets per platform, and reporting the platform with the most likes. We also calculate the relative engagement generated by each platform by dividing the number of likes per platform by the total number of likes across all platforms. We report relative engagement as a percentage of engagement with news generated on each platform.



#### 4.3.3. Politicians

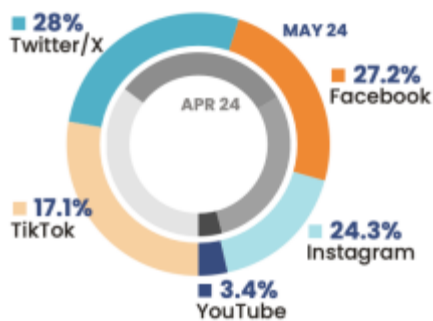
*We evaluate the role and characteristics of politicians in the information ecosystem using four measures: 1) the **top social media platforms** for engagement with Canadian politicians; 2) political engagement on social*

media (engagement with federal party leaders and Canadian party families on social media); and 3) **news versus politician engagement** (comparing engagement between politicians and news outlets).

#### 4.3.3.1. Top social media platforms

##### Top Social media platforms

Which social media platforms are most popular for Canadian politicians?



We identify the most popular platform for engagement with politicians by adding up the total number of likes on posts by Canadian politicians across our platforms for the whole month, then identifying which platform generated the most likes. We then report the percentage of all engagement received by Canadian politicians on the most popular platform to provide a sense of scale of how dominant that platform is in facilitating digital connection with politicians. This is calculated by totalling all the likes received by

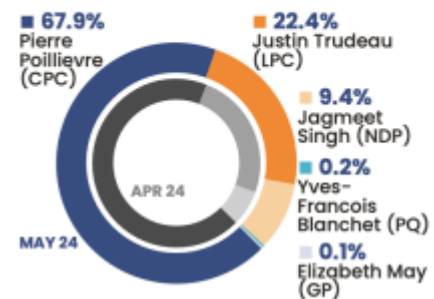
all Canadian politicians on the most popular platform, dividing it by the total number of likes received by all Canadian politicians on all our platforms of interest, then multiplying by 100.

#### 4.3.3.2. Engagement with federal political party leadership

We compare the engagement rates of the leaders of Canadian federal political parties which currently hold seats in the House of Commons. This is calculated by summing the total amount of likes each leader receives on their posts across our platforms of interest, then representing these likes as a percentage of all likes received by all party leaders.

##### Engagement with federal political party leadership

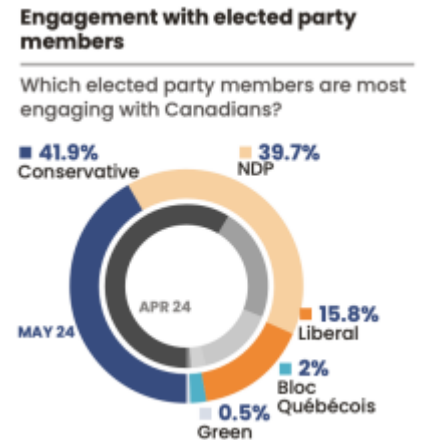
Which federal party leaders are most engaging with Canadians?





#### 4.3.3.4. Engagement with elected party members

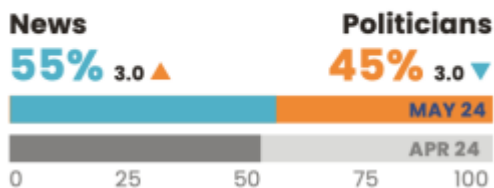
We compare the engagement rates of all Canadian party ‘families’ (Liberal, NDP, Conservative, Bloc Québécois, Green). We sum the total amount of engagement received by each party’s politicians, on both the federal and provincial level, then dividing that sum by the total amount of engagement received by Canadian politicians who are not party leaders. We report the percentage of all non-leader Canadian politician engagement received by each party.



#### 4.3.3.4. Politicians vs. news engagement

##### Politicians vs. news engagement

How skewed is engagement with politicians versus news outlets?



We report the percentage of engagement generated by Canadian politicians relative to Canadian news outlets to gain an idea of how prominent politicians are in the CIE. We calculate this percentage by dividing the total number of likes received by Canadian politicians by the total number of likes received by both news outlets and politicians.

## 4.4. Top Stories

Every month, we pick well-covered news stories and include them in our survey to gauge public awareness. If a respondent is aware of the story, then we follow up with a question asking survey respondents if they agree the story is true. We also include demonstrably false, but also circulated, stories and similarly ask if respondents saw and believed the story. These stories are mixed together and survey respondents are not given any indication which stories are real or fake, nor that some stories may be fake in the first place.

We determine these stories by collecting all the posts from that month up until the day we circulate the survey (approximately a week before we calculate and report all measures). We calculate the cosine difference for embeddings of all posts within

## Canadian Situation Report Methodology – May 2024

24 hours using [paraphrase-multilingual-MiniLM-L12-v2](#) with a threshold of 0.75. For matched texts, we form story “chains” linking earlier posts to later ones to make hundreds of chains in the month. We compute total engagement by chain and select the most-engaged-with chains to manually classify into topics and include a subset of them (typically the top-5 chains excluding sports, entertainment, etc.) in the survey. In months where this method does not give us clear misinformation stories, our team reviews known misinformation sources to supplement for inclusion in the story.

At the end of the month, we evaluate how and when these stories appeared in the CIE by measuring how often they are mentioned online, by whom, and on what platform. We depict this using violin plots: the ‘violin’ indicates the amount of engagement with the topic (wider when there is more engagement, thinner when there is less), and use shapes to indicate who posted about the topic and on what platform.