



Misinformation is not about bad facts: an analysis of the production and consumption of fringe content

JooYoung Lee^{1*}, Emily Booth¹, Hany Farid² and Marian-Andrei Rizoio¹

Handling Editor: Diogo Pacheco

*Correspondence:

jooyoung.lee@uts.edu.au

¹Data Science Institute, University of Technology Sydney, 15 Broadway, Ultimo, 2007, New South Wales, Australia

Full list of author information is available at the end of the article

Abstract

What if misinformation is not an information problem? To understand the role of news publishers in potentially unintentionally propagating misinformation, we examine how far-right and fringe online groups share and leverage established legacy news media articles to advance their narratives. Our findings suggest that online fringe ideologies spread through the use of content that is consensus-based and “factually correct”. We found that Australian news publishers with both moderate and far-right political leanings contain comparable levels of information completeness and quality; and furthermore, that far-right Twitter users often share from moderate sources. However, a stark difference emerges when we consider two additional factors: 1) the narrow topic selection of articles by far-right users, suggesting that they selectively share news articles that engage with their preexisting worldviews and specific topics of concern, and 2) the difference between moderate and far-right publishers when we examine the writing style of their articles. Furthermore, we can identify users prone to sharing misinformation based on their communication style. These findings have important implications for countering online misinformation, as they highlight the powerful role that personal biases towards specific topics and publishers’ writing styles have in amplifying fringe ideologies online.

Keywords: Misinformation; Online social media; Linguistic style

1 Introduction

Misinformation has historically been understood as an information or factual accuracy issue, where concerns arise when inaccurate narratives emerge despite the existence of a social or expert consensus on the topic. This has led to a focus on interventions that seek to remedy inaccuracies, like fact-checking, presuming that people will willingly change their perspective on an issue when presented with an alternative, authoritative form of information. These approaches, however, overlook the human side of the problem, such as the everyday anxieties and grievances that can motivate belief in misinformation [1]. Broader understandings of what constitutes misinformation have been proposed by scholars, to go beyond factual inaccuracy and deliberately fabricated content, but also include

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

conspiracy theories, manipulated images, rumors, unverified information, and misleading content, including inflammatory content spread to generate revenue such as through advertising [2, 3]. The ubiquity of digital technologies and social media means the reach and consequences of misinformation are increasingly amplified [4–6].

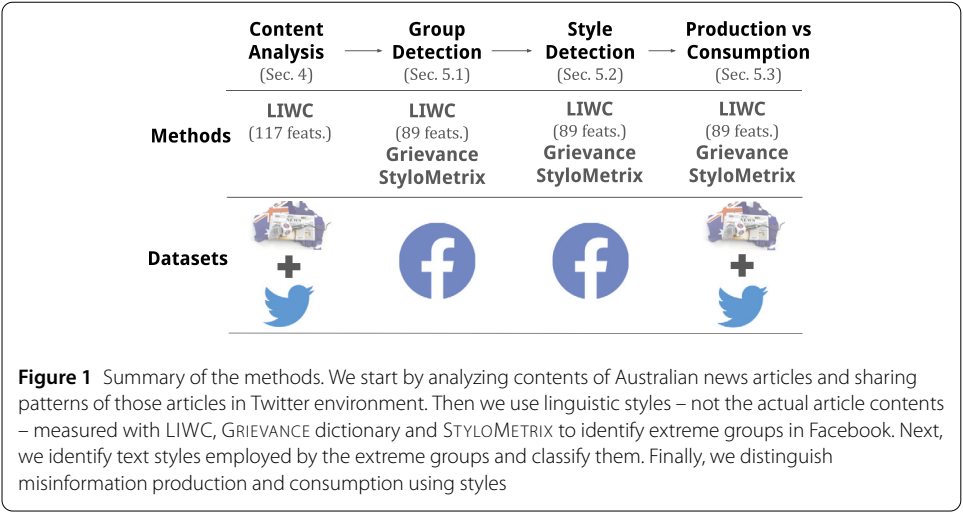
Technological interventions by social media platforms have thus far proved to be limited in controlling the spread of misinformation [7]. As a result, it follows that we must turn our attention to the heart of the issue: that humans have uncertainties and worries, and seek solace through answers. Thus, when considering misinformation online, it is crucial to consider what is being shared not as an abstract issue of factual inaccuracy, but in its context: who is sharing it, who are the likely consumers, and how is it being shared to reach audiences effectively. These audiences can include both those known to hold extremist views, like far-right groups, and fringe groups. By “fringe groups”, we mean those which adopt stances that are opposed to a consensus opinion, such as vaccine safety, but which do not necessarily draw on an overt political ideology to justify their stance. These topics can still overlap with those that extreme political groups are interested in, and they may draw on similar sources or arguments to support their views, making them interesting related cohorts to study together. A far-right group is therefore a fringe group, but not all fringe groups are far-right in nature.

To understand the role of news publishers in potentially unintentionally propagating misinformation, we examine how far-right and fringe online groups share and leverage established legacy news media articles to advance their narratives. We also demonstrate that far-right and moderate news articles in our sample differ not in the information completeness of the articles, but in their writing style, thus providing a new perspective on misinformation as an issue of style. Given their perceived authority and reach, mainstream publishers play a critical role in shaping the informational landscape and are often co-opted—deliberately or not—by actors seeking to legitimise fringe narratives. From this, we consider how style might be leveraged by moderate and consensus-based news sources to better counter the influence of online misinformation.

Research Questions We use content- and style-based measures to show that misinformation, like any other content, can be styled to target particular online cohorts. In particular, we answer the following research questions (RQs):

- *RQ1: What role does the style and content of news media play in enabling the spread of misinformation?*
- *RQ2 Can we differentiate fringe groups based on linguistic styles?*
- *RQ3 Can we differentiate contents from producers and consumers based on linguistic styles?*

The graphic in Fig. 1 depicts the steps of our analysis, the datasets used and the methods. We begin by analyzing the content completeness of news articles based on their political ideology. We compare 1) publishers of different ideologies and 2) production and consumption by far-right users. Based on the findings that content completeness of news articles does not significantly differ based on publishers' ideology, we leverage the text style (instead of its content) to investigate differences and distinguish between publishers of different ideologies and fringe online communities. Lastly, we show that far-right production and consumption can be distinguished based on the employed writing style. We expand on the details below.



We conducted a comprehensive analysis of Australian news media consumption, focusing on several key aspects. Firstly, we utilized the Trust Index [8] – a metric measuring information completeness – to compare news articles from legacy news media publishers to those with distinct ideological stances. Secondly, we examined the differences in writing style across these sources. Additionally, we delved into the sharing patterns of news articles by highly partisan online users. This differs from much of the prior work that has started from the assumption that misinformation is an information problem, and studied the connection between misinformation and the publishers’ political leaning [9, 10].

Our analysis revealed a distinct difference between what articles were published by far-right news sources and what articles were shared by the extreme (here, also far-right) ideological cohort in our sample, particularly in terms of linguistic patterns. However, we notably found less distinction between far-right and moderate articles. Even in cases where differences were statistically significant, the effect sizes were small, indicating that the sharing of extreme content is less tied to the source’s perceived ideological stance and more about how users interpret the value of an article as a tool to advance their own worldview. This is to say that users will individually select articles to share on the basis of their own views, regardless of what source published it.

This leads us to the conclusion that misinformation is not strictly a matter of factually-inaccurate content being spread by far-right sources, but that mostly-accurate content can be selected to fuel an existing fringe worldview. We note here that the sharing of content does not necessarily equate to the consumption of content, as people may share news articles based on the headlines alone. However, given the high usage of social media use to consume news media that is reported today– 54% of the American population [11] and 49% of the Australian population, as of 2024 [12]–it is likely that sharing is at minimum indicative of the type of news consumed by these groups. Furthermore, those who are deeply invested in fringe communities can often isolate themselves from news sources that express views counter to their own [1]. To explore this further, we investigate how such groups express their worldviews through different linguistic styles.

This study uses stylistic metrics from established dictionaries (such as LIWC and the GRIEVANCE dictionary) to construct stylistic classifiers. Using these stylistic classifiers, we can accurately distinguish consumers who may be vulnerable to misinformation (like anti-

vaccination and far-right sentiment) from regular online social media users. We further show that we can detect extreme ideological users' writing styles, as categorized by our manual labeling. Lastly, we show that we can successfully distinguish far-right sources from far-right consumers using a stylistic classifier.

The summary of our contributions are as follows:

1. An analysis that comparatively examines the ideological landscape of Australian news and how users selectively share news to further their own ends regardless of the source (*RQ1*).
2. A classifier to identify linguistic styles within extreme online groups as well as styles that are commonly used in the online misinformation space (*RQ2*).
3. We show that the production and consumption of misinformation exhibit patterns that can be differentiated by styles (*RQ3*).

Ethical considerations This project was approved by the Human Ethics Committee of our institution. We provide detailed discussion of ethical considerations in Sect. B.

2 Related work

In this section, we investigate content-based misinformation detection methods, their associated limitations, and style-based approaches. We acknowledge that style and content may exist on a continuum in how both factors influence word choice, and can at times appear inseparable. However, several studies have attempted to isolate stylistic features using computational tools that categorize words based on their lexical properties.

Content-based methods Traditionally, misinformation detection techniques have relied on content-based information, such as encoded texts using language models. Alkhodair et al. [13] proposed a recurrent neural network model for detecting rumors by utilizing Word2Vec [14] representation. They showed that their model outperforms the state-of-the-art sequential classifier from Zubiaga et al. [15]. They then applied their model to emerging breaking news in a real-time Twitter stream. The F1 scores for the two case studies were 0.757 and 0.791. Our style-based classifier achieved comparable results when distinguishing far-right production and far-right consumption. Similarly, Horne and Adali [16] integrated three types of features, including stylistic elements, to classify fake news. Their findings revealed significant distinctions between fake and real news content. Meanwhile, Sarnovský et al. [17] identified fake news within the Slovak online sphere. They applied deep learning models to Word2Vec representations of the texts. While the models demonstrated remarkable performance (best model reaching the accuracy of 98.93%), it is worth noting that the dataset predominantly comprised articles centered around COVID-19.

A shortcoming of content-based methods lies in the dynamic nature of misinformation topics. Models trained on predefined subjects may struggle to adapt to emerging themes. Furthermore, training deep-learning models requires substantial data to mitigate the risk of over-fitting. To address these limitations, Raza and Ding [18] introduced a context-based model focusing on social aspects to identify fake news. This model incorporates users' social interactions, such as comments on news articles, posts, and replies, as well as upvotes and downvotes. This approach can serve as a valuable complement, especially when engagement data is readily available.

Style-based methods Whitehouse et al. [19] showed that general-purpose content-based classifiers tend to overfit to specific text sources. In response, the authors proposed a style-based classification approach. The proposed stylometric features, however, leverage the categories in the General Inquirer (GI) dictionary, encompassing content-specific words related to religion and politics. To de-emphasize content we omit such categories when employing the Linguistic Inquiry and Word Count (LIWC).

More recently, Kumarage et al. [20] demonstrated the effective detection of AI-generated texts in Twitter timelines using various stylistic signals. Their study showed that classifiers employing the proposed stylometric features outperformed Bag of Words [21] and Word2Vec embeddings [14]. Notably, they mentioned that among the stylometric features, punctuation and phraseology features proved to be the most significant. While these findings are motivating, their research primarily focuses on distinguishing between human and AI-authored content within a given Twitter timeline. In contrast, we aim to discriminate between different writing styles.

In another attempt to verify authorship using style representations content-controlled style representations have been proposed [22]. The authors demonstrated that performance varies when controlling for different levels of contents, e.g., authorship verification within texts from the same conversation or the same domain. They used a clustering algorithm on text samples and manually inspected the resulting clusters of texts to find out what styles were learned, such as ‘punctuation.’

Khalid and Srinivasan [23] showed that online communities have representative and distinctive style features by predicting a community membership using style and content separately.

In our proposed work, we attempt to learn and distinguish styles used by extreme groups since we observed that these groups strategically adopted certain styles to reach vulnerable demographics in online misinformation space. This is different from learning to represent innate styles of individuals or groups which maybe more subtle.

3 Datasets and measures

We present the datasets and measures used in this work.

3.1 News and social media datasets

Australian news by Google News To evaluate the potential role of Australian news publications in facilitating misinformation dissemination, we collected Australian news articles sourced from Google News via The Daily Edit (TDE) platform. TDE aggregates news articles from Google News, encompassing 14 distinct topics: ‘Climate Change,’ ‘Sport,’ ‘Human Migration,’ ‘World,’ ‘Finance,’ ‘Technology,’ ‘Taiwan,’ ‘Top Stories,’ ‘Entertainment,’ ‘Australia,’ ‘Business,’ ‘Health,’ ‘Science,’ and ‘China.’ This is the exhaustive list of topics available when the region is set to Australia (AU) in Google News. The published period for the news articles spans from March 1, 2014, to January 12, 2023. The dataset contains 214,754 articles published by 9929 news publishers.

TDE computes a *Trust Index* score for each article as follows. Articles about the same event are grouped into *stories*. Articles are represented as sequences of sentences; all similar sentences within a story are clustered together, and clusters are interpreted as *details*. Consequently, each article has a set of supporting details (common narrative elements across multiple sources in a story). TDE computes an article’s *Trust Index* based on the

Table 1 Number of news articles by topic and political leaning

	Top Stories	Australia	World	Technology	Sport	Entertainment	Health
L	4264 (39%)	2346 (33%)	2288 (38%)	671 (46%)	1270 (41%)	1165 (35%)	396 (39%)
C	2783 (26%)	2193 (31%)	1894 (31%)	539 (37%)	609 (20%)	993 (30%)	355 (33%)
R	3768 (35%)	2504 (36%)	1876 (31%)	259 (18%)	1243 (40%)	1179 (35%)	257 (25%)
Total	10,815	7043	6058	1469	3122	3337	1008
	China	Business	Science	Finance	Human migration	Climate change	Taiwan
L	908 (41%)	720 (36%)	319 (33%)	435 (43%)	442 (41%)	318 (46%)	48 (38%)
C	846 (39%)	668 (33%)	513 (53%)	382 (38%)	436 (41%)	270 (39%)	44 (35%)
R	435 (20%)	621 (31%)	128 (13%)	200 (20%)	198 (18%)	105 (15%)	34 (27%)
Total	2189	2009	960	1017	1076	693	126

percentage of details it covers from a story – the article’s informational completeness. For example, if an article contains 7 out of 10 relevant details from the story it belongs to, its informational completeness is defined as 0.7.

The political leaning of publishers We use an external media bias dataset from *allsides*,¹ which assesses the political leaning of 473 news publishers on a five-point scale, ranging from extreme- and moderate-left to center to moderate- and extreme-right. These media bias ratings represent the average viewpoint of individuals across the political spectrum rather than the perspective of any single individual or group [24, 25]. We consolidate the extreme- and moderate-left categories into a single ‘left’ class and merge the extreme- and moderate-right categories into a ‘right’ class. This results in three political leaning classes: left, center, and right. Later, we also examine articles from publishers belonging to the ‘extreme-right’ range by *allsides*. We chose to use *allsides* due to its crowd-inclusive methodology and its focus on bias as distinct from factual accuracy. In contrast, Media Bias/Fact Check (MBFC)² relies primarily on editorial judgment and subjective assessments by a small team.

Table 1 shows the number of news articles for each topic and for each political leaning in Australian news dataset by Google News. Some studies suggest that Google news have left-leaning bias^{3,4} but in our news dataset (Table 1), while some topics show skewed proportions, there is no significant difference in the number of articles between left- and right-leaning overall.

Far-right Twitter users We use a systematically curated list containing 1496 Australian far-right coded Twitter users and collect their 3,665,809 most recent tweets using the Twitter API (at most 3200 tweets for each user). We start with 208 far-right coded users which were collected from Twitter public lists linked to far-right ideology such as ultra-nationalists and QAnon themes [26]. Then we expand the far-right coded users based on homophilic similarity [27, 28] which adds 1288 extra users (total of 1496).

The collected tweets span from February 22, 2009 to December 9, 2022. These tweets contained 1,827,162 *urls* consisting of 11,643 articles in Google News dataset (Sect. 4). We expand the articles by linking to the publishers present in either the Google News or

¹<https://www.allsides.com>.

²<https://mediabiasfactcheck.com/>.

³<https://www.allsides.com/blog/amid-google-gemini-controversy-look-google-s-history-bias>.

⁴<https://mediabiasfactcheck.com/google-news/>.

allsides dataset which results in 215,242 news articles (Sect. 5). We scraped the content of each article using `urllib.parse` Python library.⁵ We successfully downloaded the texts of 155,669 articles at the end because the rest of the links were not available.

Facebook groups At the outset of this study, we first sought to understand the online misinformation landscape in Australia, to distinguish it from other contexts like the United States of America. We began by identifying numerous Facebook groups that we suspected posted or facilitated misinformation about particular topics of concern, such as vaccination skepticism and far-right groups. This was done through a search of Facebook pages for keywords; and after identifying a relevant page, looking into what pages it followed or engaged with on the platform. We manually assemble two lists of Facebook pages for specific ideologies, namely *Australian Antivax* and *Far-right* pages. *Australian Far-right* has 14 Facebook pages and *Australian Antivax* has six Facebook pages. Using the CrowdTangle API, we collect all posts from the Facebook pages between January 23, 2021 and February 24, 2023. We collect 6017 posts from the *Far-right* pages and 2969 posts from the *Antivax* pages.

3.2 Linguistic measurements of style

We utilize three linguistic metrics that quantify linguistic attributes within text.

LIWC is one of the most widely used text analysis tools in psychology, which has recently been adopted by computational social scientists to draw insights into human behavior through computational methods [29]. This tool captures words relating to content (e.g., death, religion) and function (e.g., conjunctions, articles). *LIWC* (version 2022)⁶ has 117 categories in total. However, to capture the extreme groups' stylistic differences without contamination from content, we removed content-related features (for the list, see Table A.1), yielding a total of 89 style-related categories.

GRIEVANCE is a psycholinguistic dictionary to capture language use in the context of grievance-fueled violence threat assessment [30]. Grudge has been shown as the critical ingredient that distinguishes militant extremist mindset from social conservatism [31]; grudge (alongside confusion) is also one of the ingredients of misinformation consumption [1]. We include *GRIEVANCE* dictionary because it can extract violence and threat-specific words.

STYLOMETRIX is a grammar-related statistical representation of text. This tool allows for representing a text sample of any length with a linguistic vector of a fixed size [32] offering several preferable characteristics over well-known contextual embedding, such as BERT. First, *STYLOMETRIX* vectors encode entire documents, resolving the issue of varying text lengths. This could help when combining texts from multiple platforms, such as Facebook and Twitter. Second, *STYLOMETRIX* vectors aim to encode the entire sample's stylistic structure, not the words' meanings.

4 The content: news production and sharing

Here, we answer RQ1 and demonstrate that there is only a marginal difference in the information completeness of left- and right-leaning publishers, with the left-leaning publishers

⁵<https://docs.python.org/3/library/urllib.parse.html>.

⁶<https://www.liwc.app/>.

publishing slightly more articles. Furthermore, there are few differences in linguistic patterns between far-right and moderate publishers. However, we see a marked difference when comparing what far-right publishers *produce* compared to what far-right users actually *share* online.

4.1 Coverage and trust of publishers

Coverage by topic We extracted the 40,922 Google News articles from publishers with identified stances based on the media bias data (note that not all publishers from the Google News dataset are present in the *allsides* dataset, and we exclude any missing publishers). Table 1 shows the number of articles from the stance-identified publishers for each topic. Across all topics, except ‘Entertainment,’ we observe a higher number of left-leaning articles than right-leaning. Particularly noteworthy is the disparity in percentages within the ‘Climate change’ and ‘Technology’ topics, where the count of right-leaning articles is substantially lower than that of left-leaning articles. This indicates that the Australian news media is generally perceived to be more left-leaning.

Informational completeness by political leaning We conduct a comparative analysis of the Trust Index for articles from left-, center-, and right-leaning publishers. For each topic and political group, Table 2 provides the mean (μ), standard deviation (σ), and the number of articles (N). We assess whether there is a significant difference between Trust Index of left- and right-leaning news articles using independent samples t -tests. We report the test p -value, and we emphasize statistically significant results ($p < 0.05$). We also quantify the effect size for each test using Cohen’s d [33]. Generally, a d value of 0.2 indicates a small effect size and a value of 0.5 is considered a medium effect.

There are several observations from Table 2. First, the mean Trust Index for articles from left-leaning publishers (L in Table 2) consistently surpasses that of the right-leaning, regardless of the topic’s statistical significance. This indicates that, generally, left-leaning articles are more informationally complete. The standard deviations of the Trust Index are nearly identical across all groups and topics. Second, several topics such as ‘World,’ ‘China,’ ‘Technology,’ ‘Finance’ and ‘Human Migration’ achieve statistical significance but only moderate effect size. ‘Taiwan’ exhibits the largest effect size among all topics; however, it does not reach statistical significance due to the limited number of articles in the ‘Taiwan’ category (see Table 1). Conversely, ‘Australia’ shows a statistically significant difference, albeit with only a small effect size (Cohen’s $d = 0.11$).

Conclusion While the distinction between left- and right-leaning publishers’ information completeness is statistically significant for specific topics, the practical impact of this difference is marginal in the real world, given the small or non-existent effect size. This is unexpected, as we might assume fringe news outlets to consistently include misinformation or factual inaccuracies.

4.2 Is the content from far-right publishers different from the moderates?

Here, we demonstrate there is insufficient evidence using LIWC to show that Google News articles’ content and writing style differ based on the political leaning of the publishers. More specifically, we find this conclusion is consistent with far-right publishers (‘extreme-right’ per *allsides* classification) compared to all the others that we group here as “moderates”.

Table 2 Trust Index statistics of news articles for all topics. The *p*-values were derived from *t*-tests conducted on the L and R groups. The effect size values are calculated using Cohen's *d*. Statistically significant (*p* < 0.05) results are highlighted

	Top Stories			Australia			World			Technology			Sport			Entertainment			Health			
	μ	σ	N	μ	σ	N	μ	σ	N	μ	σ	N	μ	σ	N	μ	σ	N	μ	σ	N	
Trust Index	L	0.60	0.20	4264	0.63	0.21	2346	0.60	0.20	2288	0.58	0.19	671	0.62	0.20	1270	0.61	0.19	1165	0.61	0.20	396
	C	0.59	0.20	2783	0.61	0.21	2193	0.57	0.21	1894	0.55	0.20	539	0.55	0.20	609	0.63	0.20	993	0.58	0.21	355
	R	0.54	0.20	3768	0.60	0.21	2504	0.53	0.20	1876	0.51	0.19	259	0.55	0.20	1243	0.60	0.19	1179	0.59	0.20	257
Effect size		0.31		0.11		0.36	0.36		0.36	0.36		0.36	0.35		0.35	0.03				0.10		
p value		0.0000		0.0001		0.0000	0.0000		0.0000	0.0000		0.0000	0.0000		0.0000	0.5287				0.2218		
Trust Index	China			Business			Science			Finance			Human migration			Climate change			Taiwan			
	μ	σ	N	μ	σ	N	μ	σ	N	μ	σ	N	μ	σ	N	μ	σ	N	μ	σ	N	
	L	0.60	0.20	908	0.63	0.19	720	0.61	0.19	319	0.58	0.18	435	0.58	0.21	442	0.57	0.18	318	0.66	0.20	48
Trust Index	C	0.58	0.20	846	0.60	0.20	668	0.62	0.19	513	0.54	0.18	382	0.54	0.20	436	0.57	0.20	270	0.55	0.19	44
	R	0.52	0.20	435	0.61	0.19	621	0.60	0.19	128	0.52	0.18	200	0.51	0.21	198	0.53	0.18	105	0.57	0.20	34
Effect size		0.38		0.08			0.33		0.08		0.33		0.31		0.31	0.22				0.40		
p value		0.0000		0.1414			0.0001		0.4423		0.0001		0.0003		0.0003	0.0695				0.0793		

We do not study far-left users, because as a recent report from the Australian Institute for Strategic Dialogue found, unlike far-right groups, far-left groups do not share hyper-partisan sources online, do not weaponize conspiracy theories to spread social discord, and do not promote violence as a strategy meaning they are not a threat to public safety.⁷

Setup We qualitatively inspected randomly selected articles and identified four specific topics – ‘Top Stories,’ ‘Australia,’ ‘Finance,’ and ‘Climate Change’ – that exhibited interesting content contrasts. Table 3 shows the number of articles from far-right and moderate publishers for each category. We gather the content of each article using the beautifulsoup Python library.⁸ Next, we use LIWC to analyze the text of each article and perform *t*-tests to find whether articles from far-right publishers show statistically significant differences from moderates. To present significant results, we use the Bonferroni correction, i.e., $p < \alpha/117$ where $\alpha = 0.05$ (note: there are 117 categories in LIWC, see Sect. 3.2).

Results Table 3 presents the 21 pairs (LIWC category, article topic) for which the difference between the far-right publishers and the moderates is statistically significant – 14 LIWC categories for ‘Top stories,’ 5 for ‘Australia’ and one for each of ‘Finance’ and ‘Climate change.’ However, most of these have low ($d \sim 0.2$) or moderate ($d \sim 0.5$) effect sizes; only two such topic-category pairs show large effect sizes ($d \gtrsim 0.8$). In ‘Finance,’ far-right publishers use significantly more words from the MALE category – containing 230 words such as ‘he,’ ‘his,’ ‘him’ or ‘man’ ($d = 1.01$). The second is ‘Climate change,’ with far-right articles using statistically significantly more words in the NEGATIVE EMOTIONS category – 618 words such as ‘bad,’ ‘hate,’ ‘hurt,’ ‘worry,’ ‘fear’ ($d = 1.03$).

Conclusion The linguistic signals captured by LIWC alone are insufficient for distinguishing extreme (far-right) publishers from moderate ones. Corroborated with the conclusion from Sect. 4.1, this indicates little difference in the content and topic of what publishers of different political leanings produce overall.

4.3 Far-right articles: production vs consumption

Here, we assess whether there is a difference in linguistic patterns based on content in the articles that far-right *users share*. This differs from the analysis in the previous Sect. 4.2 in which we analyzed the difference in information *production* by far-right publishers.

Setup The far-right Twitter users shared 11,643 articles from the Google News dataset on the four topics of interest – ‘Top Stories,’ ‘Australia,’ ‘Finance,’ and ‘Climate Change.’ We process these articles’ content using LIWC. Finally, we perform *t*-tests (Bonferroni corrected) to compare articles shared by the far-right Twitter users with 1) articles produced by far-right publishers and 2) articles produced by moderate publishers. Table 4 shows select LIWC categories that show statistically significant differences. There are several noteworthy findings which we introduce below.

⁷<https://www.isdglobal.org/isd-publications/the-far-left-and-far-right-in-australia-equivalent-threats-key-findings-and-policy-implications/>.

⁸ BEAUTIFULSOUP: <https://pypi.org/project/beautifulsoup4/>.

Table 4 Analyzing far-right information consumption (Twitter-shared) and information production (publishers). We report selected LIWC features with the strongest statistical significance when compared to the far-right Twitter-shared articles. Values in gray color are non-statistically significant. We also report the effect size for each test using Cohen's d values for each topic ('Top Stories', 'Australia', 'Finance' and 'Climate change') between the Google News and the far-right Twitter-shared articles

			Words Per Sentence			Culture			politic			Lifestyle			money		
		# articles	μ	σ	d	μ	σ	d	μ	σ	d	μ	σ	d	μ	σ	d
Far-right publishers	Top Stories	148	20.64	4.09	0.93	2.27	2.18	0.43	1.37	1.97	0.47	3.60	1.70	0.77	0.44	1.10	0.69
	Australia	46	21.11	4.18	0.86	2.23	2.02	0.44	1.02	1.59	0.59	2.81	1.71	0.96	0.41	0.70	0.70
	Finance	13	21.56	3.56	0.80	1.48	0.77	0.68	0.79	0.80	0.67	9.12	2.44	~	5.30	2.44	~
	Climate change	12	21.60	4.60	~	2.03	1.80	~	1.61	1.86	~	3.64	2.24	~	0.67	1.00	~
Moderate publishers	Top Stories	9459	23.77	9.96	0.40	1.98	2.49	0.57	1.28	2.09	0.59	4.70	3.18	0.54	1.10	1.99	0.53
	Australia	6880	23.50	6.90	0.52	2.37	2.61	0.42	1.68	2.29	0.41	5.06	3.60	0.42	1.48	2.14	0.37
	Finance	1073	22.79	7.11	0.62	1.86	1.73	0.58	0.94	1.35	0.66	9.64	3.34	-0.75	6.17	3.30	-1.25
	Climate change	881	24.48	8.91	0.37	2.27	2.46	0.43	1.57	2.15	0.43	4.85	2.47	0.46	1.51	1.87	0.33
far-right Twitter shared		11643	27.13	6.98	~	3.60	3.11	~	2.78	2.86	~	6.67	4.02	~	2.46	2.92	~

Results First, we find 64 (LIWC category, article topic) pairs statistically significantly different when comparing articles from far-right producers with articles shared by far-right Twitter users. In comparison, there are 325 significantly different pairs between articles from moderate publishers and those shared by far-right Twitter users. However, only 21 pairs significantly differ between far-right and moderate publishers (see Sect. 4.2). This indicates that the far-right publishers exhibit greater similarity to moderate publishers than to the articles shared by far-right Twitter users while the far-right Twitter users share articles that exhibit less common linguistic patterns.

Investigating the mean values (μ) and standard deviation (σ) yields a similar conclusion: the differences between what the far-right Twitter users chose to share and what is produced by the far-right publishers are larger than the differences between far-right and moderate publishers.

Second, the LIWC categories for the articles that the far-right users shared showed significant differences in the categories WORDS PER SENTENCE (WPS), CULTURE, and POLITIC, with the articles Twitter shared showing significantly higher mean values. The CULTURE category includes words relating to nations, political processes (POLITIC), and ethnic identities. LIFESTYLE and MONEY also showed significant differences in the topics of 'Top stories' and 'Australia' with high effect sizes. LIFESTYLE includes words that discuss money, households, employment, and religion. This indicates that far-right users typically shared articles about common right-aligned issues like politics, societal make-up, and money.

Third, the articles shared by Twitter users use more words from the categories of CULTURE and, specifically, its subcategory, POLITIC, compared to the far-right-leaning articles from the Google news dataset. This indicates that far-right Twitter users share more political articles than the general Google News sample suggests. This seems intuitive, as most conspiracy theories touch on politics.

Conclusion The above results suggest that the information consumption patterns of online far-right users differ from the scope of article production by far-right publishers. We hypothesize that these users do not share random samples of the articles produced by far-right publishers. Rather, they selectively share the articles most useful for their arguments. We note that users may not necessarily express agreement or a positive point of view of the article that they share. In fact, far-right Twitter users regularly share articles from reputable and left-leaning publishers, potentially as evidence for a circumstance they wish to condemn. This suggests that far-right users are willing to draw on any news articles that

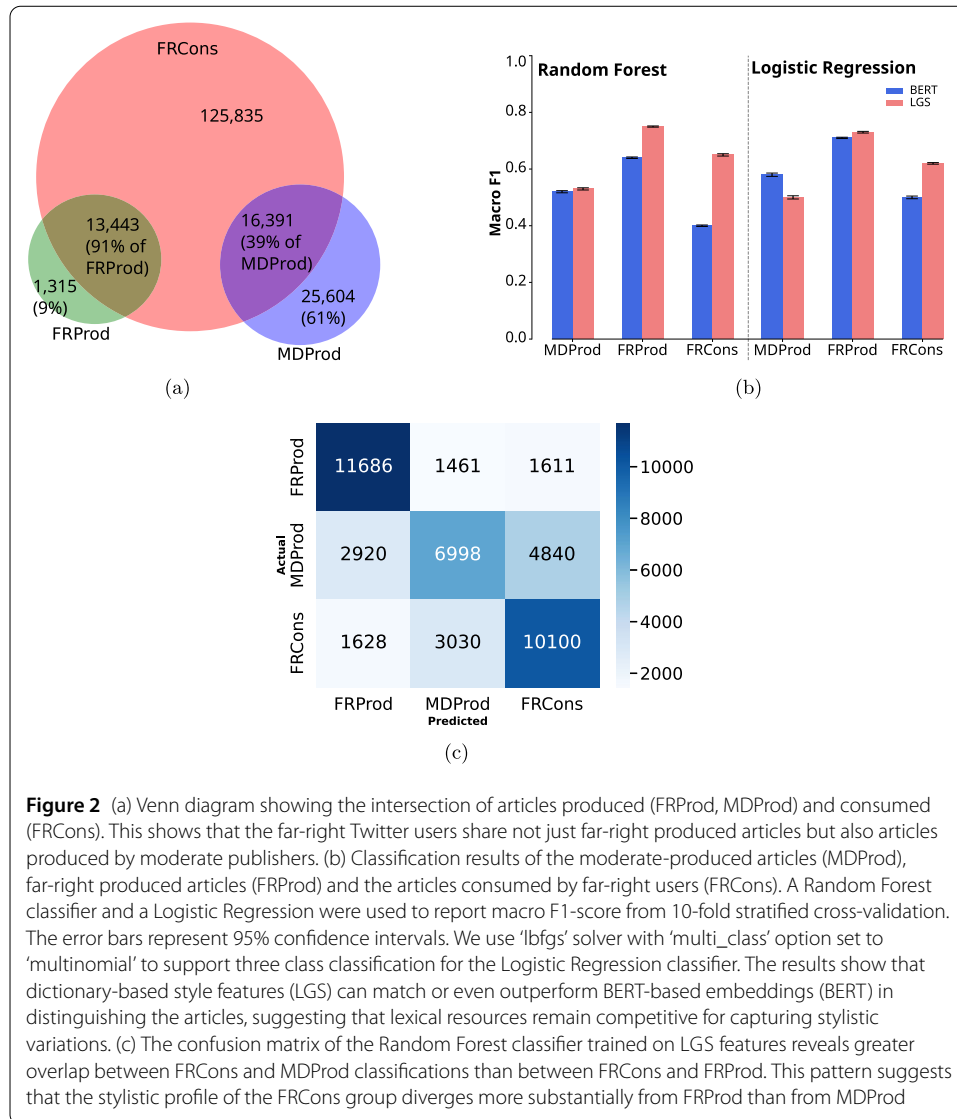


Figure 2 (a) Venn diagram showing the intersection of articles produced (FRProd, MDProd) and consumed (FRCons). This shows that the far-right Twitter users share not just far-right produced articles but also articles produced by moderate publishers. (b) Classification results of the moderate-produced articles (MDProd), far-right produced articles (FRProd) and the articles consumed by far-right users (FRCons). A Random Forest classifier and a Logistic Regression were used to report macro F1-score from 10-fold stratified cross-validation. The error bars represent 95% confidence intervals. We use 'lbfgs' solver with 'multi_class' option set to 'multinomial' to support three class classification for the Logistic Regression classifier. The results show that dictionary-based style features (LGS) can match or even outperform BERT-based embeddings (BERT) in distinguishing the articles, suggesting that lexical resources remain competitive for capturing stylistic variations. (c) The confusion matrix of the Random Forest classifier trained on LGS features reveals greater overlap between FRCons and MDProd classifications than between FRCons and FRProd. This pattern suggests that the stylistic profile of the FRCons group diverges more substantially from FRProd than from MDProd

evidence their viewpoints, regardless of the political leaning of the source (although they seem to prefer far-right over moderate produced articles, see Fig. 2a). This is quite different from what has been proposed by previous research, which is that users' engagement with a news article is heavily mediated by the source it is attributed to, and whether the user has an overall positive or negative view of the news media company itself [34].

5 The style: style over content

In the previous sections, we learned that the integrity of contents do not have strong association with the ideologies of the publishers. Additionally, contents produced by far-right publishers and consumed by far-right users are different which was unexpected. In this section, instead of comparing contents, we investigate the style of texts from different groups and distinguish between them.

Why style over content? Style words reflect how people communicate, whereas content words convey what they say in terms of meaning and topic. Style is often characterized

as a set of non-content linguistic features—including function words, syntactic structures, and punctuation—that shape the form of expression rather than its semantic content [35]. Style words are more closely linked to measures of people’s social and psychological worlds [36]. Styles encompass a range of linguistic features, including sentence structure, grammar, and punctuation patterns. Unlike the content, which continuously changes and can be influenced by subject matter, external sources, or intentional deception, stylistic features are intrinsic to one’s writing and are less prone to deliberate manipulation [37].

5.1 Identify extreme groups using styles

Here, we answer RQ2 by proposing a style classifier and demonstrating that we can distinguish ordinary online communities from different online extreme groups based on style alone.

Lee et al. [38] showed that political groups with different ideologies exhibit distinct tendencies when consuming and disseminating information on social media platforms. In order to investigate styles of ideological online groups and examine if different groups show different text styles, we use the two collections of Facebook groups that represents two datasets of extreme groups – antivax and far-right, as detailed in Sect. 3.

Method and design We begin by demonstrating the effectiveness of using style to distinguish between extreme groups and a “benign” control group. We use the scikit-learn Python library⁹ for all classifiers used in this section; Logistic regression (LR), Linear SVC (SVC), and Random Forest (RF). All three classifiers inherently support multi-class classification, and we use the default strategy for each classifier.

We design a predictive experiment aimed at evaluating and contrasting the effectiveness of three dictionary-based stylistic metrics (LIWC, GRIEVANCE and STYLOMETRIX). The task is to distinguish the posts among the three Facebook groups; two extreme Facebook groups and a NORMAL (non-extreme) group. We created the NORMAL group to encompass a wide range of discussions, from cooking to non-profit organizations. To put the performance of stylistic classifiers into context, we compare against a content-based baseline – the popular text encoding technique, BERT [39] commonly used in textual classification tasks. In contrast to the three dictionary-based encodings, BERT considers the contextual information for each instance of a given word, enhancing its capabilities. We use spacy-sentence-bert Python library¹⁰ to produce vectors of given text. The pretrained model of `en_stsb_distilbert_base` is used which transforms each sentence into a vector of 768 dimensions. spacy-sentence-bert is a wrapper of sentence-BERT package which uses mean-pooling strategy by default [40]. This approach captures contextual information from the entire input sentence and produces fixed-length vectors suitable for downstream tasks.

This exercise aims to evaluate various feature sets, not the classification algorithms. We interpret the difference in prediction performance as a difference in the representativity of the feature sets. Table 5 reports classification performance of three off-the-shelf classifiers. We randomly sampled 1000 posts from each list of groups (Far right, Antivax and NORMAL). The reported performance is the average result of 5-fold cross-validation. We

⁹<https://scikit-learn.org/stable/modules/multiclass.html>.

¹⁰<https://github.com/MartinoMensio/spacy-sentence-bert>.

Table 5 Performance of five feature sets. We distinguish between two extreme groups – *Far right* and *Antivax* – and the *NORMAL* Facebook group. We randomly sample 1000 posts for each group. We report accuracy and macro-F1 to compare among the classifiers and features used. We boldfaced the best results using BERT and LGS

		BERT	LIWC	Grievance	StyloMetrix	LGS
LR	accuracy	0.80	0.73	0.56	0.69	0.74
	macro F1	0.79	0.73	0.54	0.68	0.74
SVC	accuracy	0.76	0.66	0.57	0.71	0.68
	macro F1	0.76	0.61	0.55	0.70	0.64
RF	accuracy	0.75	0.72	0.65	0.70	0.77
	macro F1	0.75	0.71	0.63	0.69	0.76

compare the three basic stylistic feature sets (LIWC, GRIEVANCE and STYLOMETRIX) and their concatenation – denoted as LGS.

Results Since BERT leverages the content and style of a text, it is often the best-performing feature set. However, some stylistic-based classifiers outperform the content-based classifiers: the Random Forest classifier, LGS, outperforms BERT. In other words, based solely on the style and without leveraging the content of a given post, we can predict which ideological group the post came from. The true positive rate for the *NORMAL* group was 0.95, 0.79 for the *Far right* group, and 0.66 for the *Antivax* group. Some posts from *Antivax* group were misclassified as *Far right*, while only a few posts from *Far right* were misclassified as *Antivax*, likely because the anti-vax space is a niche with far-right elements, while general far-right groups discuss diverse topics.

5.2 Identifying styles

Here, we identify and classify the writing styles of posts used by people in fringe Facebook communities. This predictive exercise differs from the one in Sect. 5.1; here, we test whether human-labeled styles can be detected using a style-based classifier; we distinguished user groups in Sect. 5.1.

Method and design First, a team member with a writing background manually annotated the writing style of 100 text samples from the *Far right* and *Antivax* groups (50 for each group). She employed a deep qualitative approach in which the style labels were generated alongside the labeling. We identified styles that are extensively used in misinformation online environments by analyzing the language choices used in text-based posts across the online environment, including the grammatical features and vocabulary choices. The labeled styles in these samples are “Casual” (which used informal language as though speaking to a friend), “Empowerment” (which used angry language that sought to incite action), “Clickbait” (which mixed inflammatory words with knowledge gaps to incite curiosity and clicks), “Expert” (which used language that asserted the user’s expertise in a topic) and “Intimacy” (which used language that was highly feminine, affectionate, and vulnerable). This annotation is a result of sampling 50 posts from both *Far right* and *Antivax* then we remove the posts with no style; Table 6 shows the number of posts per style.

“Clickbait” and “Intimacy” have only two and one exemplars, respectively, and we decided to remove them from the rest of this analysis. As a result, we classify solely the styles

Table 6 Number of Facebook posts per style

	Casual	Empowerment	Clickbait	Expert	Intimacy	Total
FAR RIGHT	11	21	1	13	1	47
ANTIVAX	11	4	1	0	0	16
	22	25	2	13	1	63

Table 7 Classification results of the three styles; “Casual”, “Empowerment” and “Expert”. A Random Forest (RF) classifier and a Logistic Regression (LR) was used with a stratified split. The results show the average performance of 2-fold cross-validation due to the size of the samples

Style	Classifier	LGS		BERT	
		macro F1	Accuracy	macro F1	Accuracy
Empowerment	RF	0.60	0.58	0.51	0.54
	LR	0.49	0.52	0.67	0.66
Casual	RF	0.67	0.5	0.54	0.55
	LR	0.45	0.52	0.67	0.70
Expert	RF	0.47	0.5	0.42	0.5
	LR	0.40	0.73	0.57	0.46

“Casual”, “Empowerment” and “Expert”. We performed a binary classification for each style using One-Versus-Rest (OvR) strategy by randomly sampling negative samples to balance the sample size. For the Random Forest classifier, we used max tree depth = 3 and the number of trees = 8 due to the small sample size.

Results Table 7 reports the classification results. In general, the stylistic features (LGS) perform better with the Random Forest classifier; BERT shows better performance with the Logistic Regression. This is expected since in BERT embeddings, each dimension is not inherently interpretable or separable and thus not ideal for the Random Forest classifier which excels with features that individually carry strong predictive signals. While it is true that the Random Forest classifier may not fully leverage the rich contextual information in BERT embeddings, we aimed to assess whether it could still yield valuable insights or competitive results when compared to more complex models. This experimental choice was made to offer a comparative perspective on different feature representations and classifiers.

For all three styles, the LGS stylistic features outperform BERT with the Random Forest classifier measured by macro F1 score, which is consistent with the results presented in Table 5 for group style classification. “Expert” style was the most challenging for both classifiers since there were only 13 labeled samples.

5.3 Identify far-right articles using styles: production vs. consumption

Here, we answer RQ3 by distinguishing production from sharing. Inspired by the findings in Sect. 4.3, we hypothesize that far-right production can be differentiated from far-right consumption on the basis of style. We train the style classifier to distinguish far-right production from far-right consumption, and add moderate production for comparison purposes.

Method and design We use the 155,669 articles shared by far-right Twitter users (see Sect. 3) as far-right consumption data (FRCons) since far-right Twitter users shared these.

In order to increase the number of articles produced by far-right publishers in our sample, as well as the number of moderate articles, we expand the number of articles as follows. Each article shared by Twitter users is added to either the moderate or far-right category based on the leaning of the publisher, as long as the article's publisher is referenced in *all-sides* data. After this expansion, we have 155,669 Twitter articles shared by far-right users on Twitter, (FRCons), 41,995 articles from moderate producers (MDProd), and 14,758 far-right produced articles (FRProd).

Figure 2a shows the intersection size between these three sets of articles as a Venn diagram. 91% of all far-right produced articles are shared by the far-right consumers, whereas only 39% of the moderate articles are shared by the far-right users. Proportionally, the intersection between far-right produced (FRProd) and far-right shared (FRCons) is larger than the intersection between moderate-produced (MDProd) and far-right shared (FRCons). This shows that, while far-right users opportunistically link to articles from both far-right and moderate producers, they prefer far-right producers.

We build a textual classifier that uses the text of each article and predicts whether it is far-right produced, moderate produced or far-right shared. We downsample each class to the smallest class size (14,758 articles).

Results In Sect. 5.2, we observed that our curated stylistic features (LGS) perform better with the Random Forest classifier while BERT features outperform with Logistic Regression in most cases for identifying styles.

In Fig. 2b, we see that LGS with the Random Forest classifier consistently outperforms BERT when distinguishing production from consumption. In addition, LGS with Logistic Regression, exempting for articles from moderate publishers, outperforms BERT as well. Particularly, LGS outperforms BERT by far when distinguishing the far-right consumption group (FRCons). This indicates that it is the style, not the content, of the articles that better characterizes the far-right consumption patterns.

Figure 2c reports the confusion matrix of the three class classification from the Random Forest classifier. This result is the average of 10-fold stratified cross-validation. The most notable observation is that far-right-produced articles (FRProd) are easily distinguished from the others and far-right consumption (FRCons). The most confusions occurred between the far-right consumption (FRCons) and the moderate produced (MDProd). This result indicates that while far-right production (FRProd) and consumption (FRCons) are clearly separable, the styles utilized by moderate production (MDProd) and far-right consumption (FRCons) are similar. Also, since we did not find a significant difference between far-right and moderate articles in terms of their content (Sect. 4.2), it is implied that the styles are better signals for information consumers.

These findings reinforce the previous observations that far-right production and far-right consumption patterns are different. Specifically, far-right users do not exclusively consume far-right content, but select articles that pursue certain styles. Indeed, style itself could be a contributing factor driving political polarisation. Most notably, the findings imply that the distinctiveness of far-right texts compared to moderate sources is their style; and from this, it follows that perhaps a more distinctive approach to styling moderate views, tailored to different audiences, could help moderate perspectives compete against fringe views in today's attention economy.

6 Conclusion and discussion

We have shown that misinformation is conveyed through styled messages by detecting styles in extreme online communities and being able to distinguish the communities using styles rather than content. Specifically, Facebook pages that share misinformation can be distinguished by the style of the posts made within each group. The classification results showed that stylistic measures can outperform the content-based classifier. This is intriguing since the two groups, especially an anti-vaccination group, can be easily identified by analyzing content-related vocabulary.

We also evidenced that content produced by the far-right differs significantly from far-right online users' consumption. This is to say that while producers may cover a wide range of topics, users selected a narrow set of articles to share, which had consistent style features. Thus, we showed that the far-right production and the far-right consumption can be distinguished by using stylistic features. This indicates that misinformation consumers prefer certain styles of information as opposed to the contents of information. Though these findings are based on Australian media dataset, we believe that our insights contribute to a greater understanding of misinformation worldwide. We also note that the findings in [41] show that right-leaning users (conservatives) on Facebook are more likely to consume cross-cutting content than their counterpart (liberals) which can explain the misalignment between the far-right produced content and the far-right consumed content.

This study has limitations across multiple dimensions. Firstly, the Facebook group dataset is constrained in its post volume. Additionally, the availability of labeled style data was even more scarce, limiting the generalizability of the findings. Our plan in the future is to apply an active learning process to label more posts efficiently. Secondly, we utilized only the styles of messages in order to distinguish and observe the effects of styled text (Sect. 5). While studying online misinformation, we noticed that misinformation producers strategically adopt styles to effectively spread the information and utilize formats such as graphics, videos and reports. We can enhance the results by considering these facets of misinformation packaging. Also, some misinformation producers contribute more significantly to its dissemination [42]. Examining their patterns may help identify more discernible stylistic fingerprints.

Lastly, we explore and develop our understanding of the vulnerable demographics in online spaces. We observed that extremist online groups aimed at these personas tailor their content to these preferences. A better understanding of the patterns and styles of misinformation that vulnerable demographics are attracted to can serve as a reference for policymakers, such as for adopting a comparable strategy to effectively counter misinformation dissemination.

Appendix A: Details of stylistic measures

We summarize the linguistic measures and list the features that we used in Sect. 5. The full list of features of *STYLOMETRIX* is found here.¹¹

¹¹https://github.com/ZILiAT-NASK/StyleMetrix/blob/v0.1.0/resources/metrics_list_en.md.

Table A.1 Summary of the linguistic tools used in Sect. 5

	LIWC	GRIEVANCE	STYLOMETRIX
# features	89	22	175
features	Segment, WC, Analytic, Clout, Authentic, Tone, WPS, BigWords, Dic, Linguistic, function, pronoun, ppron, i, we, you, shehe, they, ipron, det, article, number, prep, auxverb, adverb, conj, negate, verb, adj, quantity, Drives, affiliation, achieve, power, Cognition, allnone, cogproc, insight, cause, discrep, tentat, certitude, differ, memory, Affect, tone_pos, tone_neg, emotion, emo_pos, emo_neg, emo_anx, emo_anger, emo_sad, swear, Social, socbehav, prosocial, polite, conflict, moral, comm, socrefs, substances, risk, curiosity, allure, Perception, attention, motion, space, visual, auditory, feeling, time, focuspast, focuspresent, focusfuture, Conversation, netspeak, assent, nonflu, filler, AllPunc, Period, Comma, QMark, Exclam, Apostro, OtherP	deadline, desperation, fixation, frustration, god, grievance, hate, help, honour, impostor, jealousy, loneliness, murder, paranoia, planning, relationship, soldier, suicide, surveillance, threat, violence, weaponry	

Appendix B: Ethical considerations

Data Collection and Management All data that we obtained was publicly available at the time of data collection. We discarded deleted, protected, and redacted content at the time of analysis. Therefore, the analyses reported in this work does not compromise any user privacy. This project was approved by the Human Ethics Committee of our institution (approval number redacted for review).

Broader Potential Impact of Work By exploring the landscape of Australian news media and the online misinformation space, we hope to raise awareness of how language styles used in online communities may appeal to vulnerable demographics. Understanding how misinformation is styled and packaged provides insights into how we can approach affected populations with reliable information. Awareness goes both ways, however, and the results presented in the paper could also lead to malicious users learning ways to easily manipulate a target audience. Based on our observations, however, creators of misinformation with the intention to reach a target population are already styling their messages accordingly. By showing that misinformation consumers are attracted to styled messages, we hope any entities involved in misinformation intervention campaigns can utilize these findings.

Abbreviations

LIWC, Linguistic Inquiry and Word Count, a linguistic dictionary; RQ, Research Question; GI, General Inquirer; AI, Artificial Intelligence; TDE, The Daily Edit, a data collecting platform; AU, Australia; MBFC, Media Bias and Fact Check; API, Application Programming Interface; BERT, Bidirectional Encoder Representations from Transformers; L, Left political leaning; C, Central political leaning; R, Right political leaning; WPS, Words Per Sentence; LR, Logistic regression; RF, Random Forest; SVC, Support Vector Classifier; LGS, our curated feature set that combines LIWC, Grievance and StyloMetrix; OvR, One-Versus-Rest; FRCons, Far-right consumption data, a collection of articles in our dataset which are consumed by the far-right Twitter users; FRProd, Far-right produced articles, a collection of articles in our dataset which are produced by far-right leaning publishers; MDProd, Moderate-produced articles, a collection of articles in our dataset which are produced by politically moderate publishers.

Acknowledgements

Not applicable.

Author contributions

J.L., E.B., H.F., and M.A.R. designed research; J.L. and E.B. performed research; J.L. and E.B. analysed data; and J.L. and E.B. wrote the paper; H.F. and M.A.R. reviewed the manuscript.

Funding information

This work was supported by Department of Home Affairs (PRN0006958) and Office of National Intelligence via the National Intelligence Post-Doctoral Grants (NIPG).

Data Availability

Per ethics restriction, we cannot disclose the data we used. However, the code and some news media data (headlines of news articles used) is available from the corresponding author on reasonable request.

Materials availability

Not applicable.

Code availability

The code supporting this study will be made available upon reasonable request.

Declarations**Ethics approval and consent to participate**

The study received full ethical clearance from the Human Research Ethics Committee of our university (Reference Number: ETH23-8018).

Consent for publication

All authors consent to publish.

Competing interests

The authors declare no competing interests.

Author details

¹Data Science Institute, University of Technology Sydney, 15 Broadway, Ultimo, 2007, New South Wales, Australia. ²School of Information, University of California, Berkeley, 2000 Carleton Street, Berkeley, 94720, California, USA.

Received: 20 February 2025 Accepted: 1 July 2025 Published online: 16 July 2025

References

1. Booth E, Lee J, Rizoio M-A, Farid H (2024) Conspiracy, misinformation, radicalisation: understanding the online pathway to indoctrination and opportunities for intervention. *Aust NZ J Sociol* 14407833241231756
2. Vraga EK, Bode L (2020) Defining misinformation and understanding its bounded nature: using expertise and evidence for describing misinformation. *Polit Commun* 37(1):136–144
3. Tandoc EC Jr, Lim ZW, Ling R (2018) Defining “fake news” a typology of scholarly definitions. *Dig Journal* 6(2):137–153
4. Attwell K, Smith DT, Ward PR (2021) ‘If your child’s vaccinated, why do you care about mine?’ Rhetoric, responsibility, power and vaccine rejection. *Aust NZ J Sociol* 57(2):268–285. <https://doi.org/10.1177/1440783319893468>
5. Shearer E, Mitchell A (2021) News use across social media platforms in 2020. <https://www.pewresearch.org/journalism/2021/01/12/appendix-changing-measurements-of-news-consumption-on-social-media/>
6. Kydd AH (2021) Decline, radicalization and the attack on the US Capitol. *Violence Int J* 2(1):3–23
7. Johns A, Bailo F, Booth E, Rizoio M-A (2024) Labelling, shadow bans and community resistance: did meta’s strategy to suppress rather than remove covid misinformation and conspiracy theory on Facebook slow the spread? *Media Int Aust*, 1329878–241236984
8. The Daily Edit, Inc (2023) The Daily Edit. <https://dailymedit.com/>. Accessed: 2023-01-12
9. Calvillo DP, Ross BJ, Garcia RJ, Smelter TJ, Rutchick AM (2020) Political ideology predicts perceptions of the threat of covid-19 (and susceptibility to fake news about it). *Soc Psychol Pers Sci* 11(8):1119–1128
10. Das R, Ahmed W (2022) Rethinking fake news: misinformation and ideology during the time of covid-19 global pandemic. *IIM Kozhikode Soc Manag Rev* 11(1):146–159
11. Pew Research Center (2024) Social Media and News Fact Sheet. <https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/>. Accessed: 2024-12-04
12. Park S, Fisher C, McGuinness K, Lee J, McCallum K, Yao S (2024) Digital news report: Australia 2024. News media research centre. University of Canberra. <https://doi.org/10.60836/fxcx-xq72>
13. Alkhodair SA, Ding SH, Fung BC, Liu J (2020) Detecting breaking news rumors of emerging topics in social media. *Inf Process Manag* 57(2):102018
14. Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space
15. Zubiaga A, Liakata M, Procter R (2016) Learning reporting dynamics during breaking news for rumour detection in social media. *arXiv preprint arXiv:1610.07363*
16. Horne B, Adali S (2017) This just in: fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. In: *ICWSM*, vol 11, pp 759–766
17. Sarnovský M, Maslej-Krešňáková V, Ivancová K (2022) Fake news detection related to the covid-19 in Slovak language using deep learning methods. *Acta Polytech Hung* 19(2):43–57
18. Raza S, Ding C (2022) Fake news detection based on news content and social contexts: a transformer-based approach. *Int J Data Sci Anal* 13(4):335–362
19. Whitehouse C, Weyde T, Madhyastha P, Komninos N (2022) Evaluation of fake news detection with knowledge-enhanced language models. In: *ICWSM*, vol 16, pp 1425–1429
20. Kumarage T, Garland J, Bhattacharjee A, Trapeznikov K, Ruston S, Liu H (2023) Stylometric Detection of AI-Generated Text in Twitter Timelines
21. Manning CD, Raghavan P, Schütze H (2008) Introduction to information retrieval. Cambridge University Press, Cambridge

22. Wegmann A, Schraagen M, Nguyen D (2022) Same author or just same topic? Towards content-independent style representations. In: Gella S, He H, Majumder BP, Can B, Giunchiglia E, Cahyawijaya S, Min S, Mozes M, Li XL, Augenstein I, Rogers A, Cho K, Grefenstette E, Rimell L, Dyer C (eds) Proceedings of the 7th workshop on representation learning for NLP. Association for Computational Linguistics, Dublin, pp 249–268. <https://doi.org/10.18653/v1/2022.repl4nlp-1.26>. <https://aclanthology.org/2022.repl4nlp-1.26>
23. Khalid O, Srinivasan P (2020) Style matters! Investigating linguistic style in online communities. In: Proceedings of the international AAAI conference on web and social media, vol 14, pp 360–369
24. Park S, Fisher C, McGuinness K, Lee JY, McCallum K (2021) Digital news report: Australia 2021. News and media research centre (UC). University of Canberra
25. Newman N, Fletcher R, Schulz A, Andi S, Robertson CT, Nielsen RK (2021) Reuters institute digital news report 2021. Reuters Institute for the Study of Journalism
26. Bailo F, Johns A, Rizoiu M-A (2023) Riding information crises: the performance of far-right Twitter users in Australia during the 2019–2020 bushfires and the COVID-19 pandemic. *Inf Commun Soc*, 1–19. <https://doi.org/10.1080/1369118X.2023.2205479>
27. Ackland R, Shorish J (2014) In: Cantijoch M, Gibson R, Ward S (eds) Political homophily on the web. Palgrave Macmillan, London, pp 25–46. https://doi.org/10.1057/9781137276773_2
28. Ram R, Rizoiu M-A (2022) Data-driven ideology detection: a case study of far-right extremist. Defence Human Sciences Symposium
29. Chung CK, Pennebaker JW (2018) What do we know when we liwc a person? Text analysis as an assessment tool for traits, personal concerns and life stories. In: The Sage handbook of personality and individual differences, pp 341–360
30. Vegt I, Mozes M, Kleinberg B, Gill P (2021) The grievance dictionary: understanding threatening language use. *Behav Res Methods*, 1–15
31. Stankov L (2021) From social conservatism and authoritarian populism to militant right-wing extremism. *Pers Individ Differ* 175:110733
32. Okulska I, Zawadzka A (2022) Styles with benefits. The stylometrix vectors for stylistic and semantic text classification of small-scale datasets and different sample length
33. Cohen J (2013) Statistical power analysis for the behavioral sciences. Routledge, New York
34. Yun GW, Park S-Y, Lee S, Flynn MA (2018) Hostile media or hostile source? Bias perception of shared news. *Soc Sci Comput Rev* 36(1):21–35
35. Pennebaker JW, Mehl MR, Niederhoffer KG (2003) Psychological aspects of natural language use: our words, our selves. *Annu Rev Psychol* 54(1):547–577
36. Tausczik YR, Pennebaker JW (2010) The psychological meaning of words: liwc and computerized text analysis methods. *J Lang Soc Psychol* 29(1):24–54
37. Rocha MA, Morais PSG, Silva Barros DM, Santos JPQ, Dias-Trindade S, Medeiros Valentim RA (2022) A text as unique as a fingerprint: text analysis and authorship recognition in a virtual learning environment of the unified health system in Brazil. *Expert Syst Appl* 203:117280
38. Lee J, Wu S, Ertugrul AM, Lin Y-R, Xie L (2022) Whose advantage? Measuring attention dynamics across youtube and Twitter on controversial topics. In: ICWSM, vol 16, pp 573–583
39. Devlin J, Chang M-W, Lee K, Toutanova K (2018) Bert: pre-training of deep bidirectional transformers for language understanding. arXiv preprint [arXiv:1810.04805](https://arxiv.org/abs/1810.04805)
40. Reimers N, Gurevych I (2019) Sentence-BERT: sentence embeddings using Siamese BERT-networks
41. Bakshy E, Messing S, Adamic LA (2015) Exposure to ideologically diverse news and opinion on Facebook. *Science* 348(6239):1130–1132
42. Kim S, Kim K, Xue H (2024) Fingerprints of conspiracy theories: identifying signature information sources of a misleading narrative and their roles in shaping message content and dissemination. *J Online Trust Saf* 2(2)

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)