Countering Domestic Racially and Ethnically Motivated Violent Terrorism on Social Media

Introducing the Racist and Violent Extremist Flock Tool
About This Guide

The U.S. Department of Homeland Security and other law enforcement support organizations are interested in utilizing natural language processing tools for a variety of national security purposes, specifically in countering domestic extremism and thwarting violent extremist events that originate online. These technologies are believed to offer increased accuracy, convenience, and speed to support antiterrorism objectives, allowing analysts to stay abreast of extremist activities online.

In this guide, we introduce a natural language processing tool that facilitates exploration of textual messages on extremist internet forums, identifies trending racially and ethnically motivated violent extremist terms on internet platforms, and helps flag security events likely to take place in the real world. The Racist and Violent Extremist Flock (RVE-Flock) tool analyzes textual content on social media to identify emerging terms used in racially and ethnically motivated violent extremist communities and allows users to explore connections between words and images. This guide should be of interest to officials across the United States—specifically, those responsible for the acquisition, development, and implementation of new technologies and those responsible for countering extremism and flagging bad actors.

This guide quotes and includes references to objectionable content, including hate speech and material that could be offensive or obscene. Because this content is integral to the research and our findings, it is presented with minimal editing.

RAND Homeland Security Research Division

Funding for this research was made possible by the independent research and development provisions of RAND Corporation contracts for the operation of its U.S. Department of Defense federally funded research and development centers.

This research was conducted in the Management, Technology, and Capabilities Program of the RAND Homeland Security Research Division (HSRD). HSRD operates the Homeland Security Operational Analysis Center, a federally funded research and development center sponsored by the U.S. Department of Homeland Security. HSRD also conducts research and analysis for other federal, state, local, tribal, territorial, and public- and private-sector organizations that make up the homeland security enterprise, within and outside the Homeland Security Operational Analysis Center contract. In addition, HSRD conducts research and analysis on homeland security matters for U.S. allies and private foundations.

For more information on HSRD, see www.rand.org/hsrd.

Acknowledgments

We are thankful to RAND for providing the support and resources to conduct this effort. We were fortunate to have Emma Westerman and Kelly Klima provide guidance throughout this
research effort. Heather J. Williams and Alexandra T. Evans provided us with data acquired from their extremism work and introduced us to new scholarship. We give many thanks to Kristin J. Leuschner for her assistance in editing our report. We are also grateful to our reviewers, Luke J. Matthews, Ryan Andrew Brown, and Rhianna C. Rogers, whose insightful comments and suggestions improved this guide.
Summary

In response to the increasing frequency of racially and ethnically motivated violent extremist (REMVE) attacks, greater attention has been drawn to the online spaces in which REMVEs coordinate with others and disseminate their views. Outside observers of these online spaces are hard-pressed to maintain an awareness of discussions, lingo, or phrases that can be intentionally ambiguous, evolve rapidly with time, and occur with sufficient frequency as to make manual analysis of online text infeasible. To facilitate in the analysis of these spaces, we have constructed a methodology and have developed a tool, RVE-Flock, to identify and understand emerging terms as they enter and evolve in the online lexicon of specific websites.

The analysis leverages natural language processing to characterize the extent to which different terms are emerging online based on different metrics, such as recent increases in their usage, changes in their context, or proliferation onto other forums or websites. In this analysis, we applied our metrics to posts made on 4chan’s /pol/ to illustrate the tool’s potential to identify emerging terms and facilitate exploration and definition of terms’ usage. To highlight the tool’s ability to contextualize terms and track their evolution through time, we focus in this report on a set of terms, *glowies* and *glown----r* (the second part of the term—the last six letters—is a racial slur), that are used to describe undercover agents attempting to capture social media posters admitting to or committing crimes online. We also applied the tool’s capabilities to detect how emerging terms evolve. In the case of *glowies* and *glown----*, RVE-Flock can capture how the usage and adoption of changes through time are evidenced by changes in their word associations in response to the death of Jeffrey Epstein. Specifically, the context of these words prior to the event tended to carry the connotation of authorities either imposing values or observing the population. After the event, the words became more associated with aggressive terms, suggesting that they believed that so-called glowies and glown---- were more involved participants than observing authorities. These changes are also reflected in emerging term metrics to describe contextual change and increasing usage following the event. Additionally, our use of RVE-Flock’s capabilities to identify emerging terms did succeed in identifying one term, *bantz*, during a period in which it was increasing in popularity and allowed us to track growing terms.

With the promise of these initial results, we have identified opportunities to further refine the tool, apply other methodologies, and apply it to other areas, such as targeted violence against nonracial or nonethnic types of identities (e.g., religion, sexuality, or gender) or for identifying potential events or topics of special interest to REMVEs.
# Contents

About This Guide ................................................................................................................................ iii
Summary ............................................................................................................................................... v
Figures and Tables ............................................................................................................................... ix

CHAPTER ONE
Introduction ................................................................................................................................. 1

CHAPTER TWO
The Landscape of Similar Tools and Gaps in the Market ...................................................... 5

CHAPTER THREE
Emergence Patterns from Fringe to Mainstream ................................................................. 7

CHAPTER FOUR
The RVE-Flock Tool .................................................................................................................. 15

CHAPTER FIVE
The Trending Lexicon ................................................................................................................. 23

CHAPTER SIX
Areas for Future Development ................................................................................................. 39

CHAPTER SEVEN
RVE-Flock Concluding Thoughts .......................................................................................... 47

APPENDIXES
A. Additional Market Research ................................................................................................. 49
B. Selecting REMVE Terms ...................................................................................................... 55
C. Term Emergence Exemplars ................................................................................................. 69
D. How Racially and Ethnically Motivated Violent Extremist Actors Use the Internet Before Attacks .......................................................................................................................... 77

Abbreviations ............................................................................................................................. 79
References ..................................................................................................................................... 81
Figures and Tables

Figures

3.1. *Glown***** Usage on Social Media ................................................................. 11
3.2. *Glowies* Usage on Social Media ................................................................. 12
4.1. The RVE-Flock User Interface ................................................................. 20
5.1. RVE-Flock Search Results for *Glowies*, July–October 2019 ......................... 24
5.2. Intensity of Use and Topical Similarity for *Glowies*,
    August 2017–October 2019 ................................................................. 26
5.3. Intensity of Use and Topical Similarity for *Glown*****,
    August 2017–October 2019 ................................................................. 27
5.4. Comparative Use of the Word *Bantz* on 4chan, Reddit, and Twitter ................ 32
5.5. Use of *Bantz* on 4chan ........................................................................ 33
5.6. RVE-Flock Search Results for *Bantz*, December 2016–March 2017 ................. 34
B.1. Interest over Time in *Cuckservative*, 2014–2019 ........................................... 63
B.2. Interest over Time in *Cuck*, 2014–2019 .................................................... 64
B.3. Search Volume for (((Echo))) on Google Trends ........................................... 65
C.1. Western Chauvinism Usage on Social Media ................................................. 69
C.2. *GEOTUS* Usage on Social Media ............................................................. 70
C.3. *RWDS* Usage on Social Media .................................................................. 71
C.4. *Pinochet Did Nothing Wrong* Usage on Social Media .................................. 72
C.5. *We Wuz Kangz* Usage on Social Media ...................................................... 73
C.6. *God Emperor Trump* Usage on Social Media .............................................. 74
C.7. *Venerate the Housewife* Usage on Social Media .......................................... 75

Tables

4.1. Associations Between Emerging-Term Metrics ............................................. 21
5.1. Sample of Words with High Topical Similarity with *Glowies* and *Glown***** 28
5.2. Terms with the Highest Growth Rates, December 2016–March 2017 ............... 35
5.3. Terms with the Greatest Weekly Increase in Use, March 5–20, 2017 ............... 36
5.4. Topical Dissimilarity, June–October 2019 .................................................... 37
5.5. Terms with the Largest Increases in Use, June–October 2019 ......................... 37
CHAPTER ONE

Introduction

Terrorist organizations and adversarial state actors increasingly use social media platforms to recruit, radicalize, and mobilize populations to violence, often based on differences in race or ethnicity. The harm done by such attacks goes beyond the loss of life, often inflicting trauma on targeted communities and causing societal repercussions that include decreased trust in institutions and fellow citizens. Additionally, these attacks can serve as inspiration for similar, subsequent attacks—as was the case in the May 2022 mass shooting in Buffalo, New York, by a self-avowed, internet-radicalized white supremacist (“Gunman Kills 10 at Buffalo Supermarket in Racist Attack,” 2022).

Violent extremists capitalize on the increasing availability of relatively unrestricted spaces online and develop their movements at a pace beyond the U.S. government’s capacity to respond. This domestic problem is compounded by the reality that foreign actors are keen to exploit and deepen social divides to threaten U.S. security. According to the Office of the Director of National Intelligence, escalating support from people outside the United States increases the likelihood and lethality of domestic violent extremist attacks (Office of the Director of National Intelligence, 2021). Furthermore, U.S. near-peer adversaries have used online disinformation campaigns to influence and incite violence; for example, Russia amplified disinformation on the 2020 presidential election results leading up to the January 6 Capitol insurrection and encouraged racial divisions throughout the U.S. populace in the wake of George Floyd’s murder (DiResta et al., 2019).

Racially and ethnically motivated violent extremism is the most lethal and prevalent form of domestic violent extremism in the United States and is a national security threat that can have sweeping effects on the country’s democracy (National Security Council, 2021). There is a growing need to expose and counter those voices that seek to dominate marginalized communities and perpetrate violence against them. However, the sheer volume of discussion on racially and ethnically motivated violent extremist (REMVE)—affiliated forums and the speed at which language evolves makes it difficult for even the most informed to stay apprised of the status and evolution of REMVE terminology.

Staying abreast of the dynamic communication terms and styles of REMVE groups would enable U.S. policymakers to better understand the constantly evolving REMVE landscape

---

1 For a discussion of the consequences of hate crimes, see Baali, 1988, and Iganski, 2001.
and be better postured to prevent domestic terrorist attacks. It would be useful to have a means to identify new and emerging REMVE terminology soon after it begins to be used. Identifying these REMVE terms as they become popular in extremist communities would help analysts understand the landscape in real time. In this guide, we describe a tool that enables the exploration and analysis of language used on REMVE-affiliated social media sites. The Racist and Violent Extremist—Flock (RVE-Flock) tool analyzes textual content on social media, allowing users to detect terms used in REMVE communities and identify trends on internet platforms.

To develop the tool, we first conducted market analysis to determine what extremist language detection and event prediction products were available and where gaps existed that RVE-Flock could fill. We conducted qualitative research (which included literature reviews of specific REMVE terminology and usage, a qualitative analysis of their context, and analysis of term usage and rates of increases) to understand how and when REMVE terms proliferate on social media sites and throughout the internet, from their inception to their emergence into the mainstream. We investigated several REMVE terms to identify when they were first used, when they began to emerge, and when they appeared on mainstream platforms. By exploring these terms, we identified three term emergence patterns. Herein, we focus on the term glowies and use knowledge about its introduction and emergence to demonstrate the capability of the RVE-Flock tool: to verify whether RVE-Flock could identify glowies before its usage on mainstream platforms. We also highlight RVE-Flock’s capabilities to provide additional context for words.

Although we present and draw attention to results from RVE-Flock that could be affiliated with REMVE, we want to stress that findings contained herein are limited in scope and application. We focused on one board (/pol/) on one website (4chan) that is distinct from many other, more-mainstream social media websites, such as Twitter and Reddit. Further still, we analyzed social media posts within fixed time periods, which might not be suggestive of general language because language and word associations might have drifted between the observed periods and now. Finally, the community on /pol/ is likely diverse and not composed solely by U.S.-based users, and it is hard to gauge the extent to which people earnestly hold or advance REMVE beliefs. Further still, it is not likely that all potential varieties of REMVE groups are proportionately represented on the board. As a result, these findings and associations can be tied only to the data we have analyzed, and caution should be taken in extending findings from RVE-Flock to other sites or communities.

---

2 Glow and glowy are terms used to describe members or agents of the government posing as ordinary posters or community members and attempting to encourage other posters to admit to or commit illegal actions. Note that glowy is used not only by REMVE communities and websites but also by the far right in general.
The remainder of this guide is organized into five chapters and four appendixes:

- In Chapter Two, we describe the landscape of similar tools and existing gaps in the market.
- In Chapter Three, we examine how REMVE terms have emerged in social media and on the internet.
- In Chapter Four, we describe the RVE-Flock Tool.
- In Chapter Five, we evaluate the tool’s ability to draw attention to relevant words.
- In Chapter Six, we describe areas for future work.
- In Appendix A, we describe additional market research on commercial and academic tools to detect troubling speech and to predict future threats.
- In Appendix B, we detail the keywords we used in our literature review search, the REMVE terms we considered, and two specific REMVE terms, *cuck* or *cuckervative* and *(((echo)))*.³
- In Appendix C, we describe the emergences of a few terms (*western chauvinism*, *god emperor of the United States* [GEOTUS], *right-wing death squad* [RWDS], *Pinochet did nothing wrong*, *we wuz kangz*, *god emperor Trump*, and *venerate the housewife*) and their paths into mainstream social media.⁴
- In Appendix D, we describe how REMVE actors use the internet before attacks.

---

³ *Cuck* is a term used to denigrate someone online as being submissive or similar to a shill.

⁴ *Western chauvinism* is a term used to describe the perceived superiority of Western culture and belief systems. *GEOTUS* is taken to reference Donald Trump. *Pinochet did nothing wrong* is a phrase that makes light of the atrocities of Augusto Pinochet Ugarte, indicating that they were within the bounds of acceptable behavior in the U.S. context. *We Wuz Kangz* is a phrase used to denigrate the accomplishments and heritage of precolonial African societies. *Venerate the housewife* is a phrase used to indicate value for a traditionalist role of women in society as opposed to those afforded by feminism.
CHAPTER TWO

The Landscape of Similar Tools and Gaps in the Market

In devising the RVE-Flock tool, we conducted market research to determine what extremist language detection and event prediction products were available and where gaps existed that RVE-Flock could fill. Toward that end, we complemented existing team knowledge with an open-source search to identify machine learning (ML) and natural language processing (NLP) tools and methods that analyze extremist rhetoric or predict violent extremist events. Through a combination of previous research efforts, experience with REMVE-specific lexical signatures, and a review of the literature on REMVE violence and online behavior, we were able to identify these tools and methods using aggregated lists from databases and other online sources. The literature search included searching online academic repositories, such as Google Scholar; evaluation; and selection of sources to identify themes and gaps. Our research included commercial, academic, and nonprofit research-developed offerings.

Through this research, we identified several ML and NLP tools (e.g., ExTrac, Pyrra, and IvySys Technologies) that detect concerning speech, including extremist rhetoric and deliberate disinformation. These tools predominantly detect and monitor individual threats, organized harassment and disinformation campaigns, and the parties responsible for such activity. (Some tools also claim to forecast emerging threats.) For more information on these tools, as well as those that claim exclusively to forecast emerging threats or predict future threats, see Appendix A.

In our research, we did not identify any ML or NLP tools that analyzed evolving speech or keywords in extremist subcultures, although we did find two organizations that rely on manual, human-centered means to reach these ends:

- The Sentinel Project, a Canadian nongovernmental organization that employs an early warning system to prevent mass violence in Africa, maintains Hatebase, the “world’s largest online repository of structured multilingual, usage-based hate speech. It is an attempt to create a repository of words and phrases that researchers can use to detect the

---

1 Rather than seek to identify specific entities individually, we used aggregated lists from databases and other sources, including Disinfo Cloud, undated; Morris, 2018; Pearson and Detges, 2020; and RED-Alert, undated b.
early stages of genocide” (Sentinel Project, undated b; see also Hatebase, undated, and Sentinel Project, undated a).

- The Shorenstein Center on Media, Politics and Public Policy at the Harvard Kennedy School maintains a Technology and Social Change Project, led by Joan Donovan. In an Expert Working Group on Online Extremism convened by the U.S. Department of Homeland Security in April 2021, Donovan mentioned that, in her work to analyze language shifts and identify emerging keywords, she used a manual method of detecting evolving far-right subculture keywords using “content tags.”

Accordingly, although there appear to be attempts to monitor and track evolving keywords in violent extremist subcultures, an algorithm-driven tool that can enable people to better analyze massive amounts of data more efficiently than they can analyze manually appeared to present a gap that RVE-Flock could fill.

2 Personal attendance by RAND researcher, April 2021; National Counterterrorism, Innovation, Technology, and Education Center, 2021, p. 4.
CHAPTER THREE

Emergence Patterns from Fringe to Mainstream

In this chapter, we discuss how the REMVE community’s language evolves on the internet by examining a set of REMVE terms and their emergence patterns. To do so, we used a mixed-method research design that entailed a qualitative literature review, research on terms used in REMVE and non-REMVE contexts, and an analysis of their frequency of use through time. We selected 13 terms from a larger list of more than 150 words based on their introduction timeline’s potential to coincide with the available social media data-set time profile and relevance in current REMVE communities. We were able to identify three distinct emergence patterns based on the exploration of ten terms. In this chapter, we present an exemplar term for each emergence pattern; the remainder are presented in Appendix C.

We considered a term as emerging if it met all of these three criteria:

- It was being used at an increasing rate for at a given period (signifying accepted use in online communities).¹
- It had proliferated to other websites (signifying common meaning).
- It was used in REMVE contexts without any topical change.

We considered a term to have emerged into the mainstream if it met both of these criteria:

- It was used outside of REMVE contexts or niche online communities.
- It was being discussed in a context outside of REMVE communities.

Our research on these ten terms and their corresponding emergence patterns informed the development of RVE-Flock by determining how specific types of metrics and measures could be useful, as well as verifying the functionality of the tool. This guide quotes and includes references to objectionable content, including hate speech and material that could be offensive.

¹ The increasing-rate criterion is calculated using the change of increases over three months. This means that we compared the increase in a three-month window with the increase in the previous three months. If the ratio between those two windows was consistently greater than 1 (the usage of the term was growing at an increasing rate), we considered it to satisfy this criterion.
or obscene. Because this content is integral to the research and our findings, it is presented with minimal editing.

The Rise of REMVE Terms

We began by selecting a subset of terms from the larger list of more than 150 terms or phrases whose proliferation we could study to inform the development of the tool, which was derived as a compilation of open-source searches and literature reviews. To refine this list, we compared the extent to which each term was searched on Google Trends and captured in Brandwatch’s Twitter application programming interface (API). Google Trends showcases user search data. If people are searching for a term using Google’s search engine, this could reflect users attempting to learn about a specific term or phrase and is not necessarily representative of a term’s usage rate. Brandwatch showcases term usage on social media platforms. We relied on Brandwatch’s social media data from Twitter, 4chan, and Reddit to determine whether a new term was introduced and then consistently used. With Brandwatch, we were also able to see context and how the terms were used, as well as the time periods of use. Using both Google Trends and Brandwatch helped to create timelines of a term or phrase’s use history. Informed by subject-matter expertise of research staff and the potential to find the term in our available social media data sets, we selected 13 terms to further investigate: glown-----, glowies, cuck, manosphere, Judeo-Christian values, (((echo))), GEOTUS, god emperor Trump, western chauvinism, Pinochet did nothing wrong, RWDS, we wuz kangz, and venerate the housewife.

Not all 13 terms were adequate in testing our tool. For example, according to Know Your Meme (KYM), cuck and manosphere emerged much earlier than the start of our Brandwatch data set, which goes back to only 2012 (“Cuck,” 2019; “Manosphere,” 2015). We removed Judeo-Christian values and (((echo))) because tracking usage was too difficult. Although Judeo-Christian values had high levels of usage in REMVE communities, a large majority of its mentions occurred in a political or religious debate context that showed no signs of REMVE usage, and it was not possible to sort through every post or mention used in a REMVE context in a meaningful way. Subsequent analysis would overestimate the number of instances a social media post intentionally uses Judeo-Christian values in such a way as to demean or devalue people from other cultural backgrounds rather than efforts to counter-message such statements, triggering a high number of false positives. The three- or multiple-

---

2 The literature search included searching online academic repositories, including Google Scholar, and evaluating and selecting sources to identify terms, previous analysis of them by other research efforts, and any additional information that would serve this work. The full list of terms and keywords used in the literature review is available in Appendix B.

3 Manosphere is a term used to describe the broader community centered on internet content that caters to men and advances beliefs that can be misogynist or racist. Judeo-Christian values is often used to refer to the value and belief system that centers on white, European heritage as opposed to other ethnic or racial groups’ backgrounds. The three parentheses are a typographical convention known as an echo (see Yglesias, 2016).
Parenthesis pattern, (((echo))), can be difficult to detect because most NLP techniques filter out punctuation symbols, including parentheses. The (((echo))) usage is a subtle anti-Semitic slur or effect used when someone wants to subtly identify or target another who is Jewish. The offending actor would type (((echo))) and replace echo with whatever or whomever they want to target without being noticed. It originates from a 4chan post in which an echo sound effect was used to reference a Jewish person, which then became represented in text as at least three parentheses to target Jewish people. We were not able to track this term in Brandwatch because the syntax of Brandwatch made it difficult to track three parentheses. A deeper dive into the language history of (((echo))) and cuck can be found in Appendix B.

We used Brandwatch to search for the use of specific terms and phrases across platforms for extended time periods, and we downloaded usage data from 4chan, Reddit, and Twitter for the chosen terms. 4chan and Reddit are widely used and have large REMVE communities (Tuters and Hagen, 2020; Rieger et al., 2021), while Twitter is representative of a more mainstream platform that features REMVE communities. Using these platforms, we had a manageable data set that we could use to investigate the full text of large numbers of posts to determine the context of a social media post and determine whether it was being used in a REMVE context. This involved examining whether a given term, such as cuck or (((echo))), was being used either to advance or in association with REMVE causes or denigrate specific racial and ethnic groups.

RAND’s Brandwatch account allowed exporting historical data limited to ten years prior to the download date (the limits are determined by the account type). We downloaded all data that were available from 2012 through 2022 for this research effort. The Brandwatch data included the total number of mentions for specific REMVE terms on each platform (tallied by day), as well as the full text of each post. We identified the first time each term was used, as well as large spikes and spike clusters in usage and trends of use over time in both one-week and three-month intervals. We focused on the selected available terms that arose within the past ten years and when the term was first used in the REMVE context. Several terms were used prior to 2012, which disqualified them from our study. We considered a use to be the first time if there were no mentions of the word in the data set in 2012. We then used three-month intervals between 2012 and 2022 to calculate when a term experienced increasing usage rates (i.e., when a term was being used at an increasing rate).

We then used full-text files of the posts to manually examine the content of posts during the dates of spikes or increased usage to determine whether the term was being used in a REMVE context and how it was being discussed. To provide greater context of emergence, we also examined news stories surrounding the date of the spikes, looking for events that might have motivated the poster. For example, a term might have emerged because of a news article about its use at a specific event. These articles often explain what the term means, generat-

---

4 For Twitter, Brandwatch counts each retweet as a data point. On 4chan, any reply to a thread containing the phrase is counted as a new data point. For our purposes, we counted each of these data points as an individual use. Future efforts could investigate whether these types of mentions are the same as a unique use.
ing further discussion on Twitter outside of REMVE circles. After compiling the data and analysis, we were able to identify periods of emergence and the point at which a term had fully emerged into the mainstream.

We were able to identify three distinct patterns of emergence using the selected list of REMVE terms:

- a slow buildup and a slow emergence into the mainstream (more than six months from first usage of the term)
- a slow buildup and a rapid emergence into the mainstream (more than six months from first usage but then fully merged within six months of the first spike because of an event)
- a fast buildup and a rapid emergence into the mainstream.

We used these patterns to help us categorize ways in which terms might emerge and become mainstream. In the next section, we discuss one of these patterns of emergence exemplified by *glowies* and subsequently provide a brief synopsis of terms that remain popular within REMVE communities even after they emerge into the mainstream.

**Glowies’ Slow Buildup and Slow Emergence into the Mainstream**

Some REMVE terms that we examined displayed a slow buildup of use by REMVE-affiliated communities over a time frame of greater than six months and a slow emergence into mainstream media platforms. In these instances, we saw no significant event–caused major spikes in usage that coincided with the term entering the mainstream as a widely recognized and understood REMVE-associated term. Terms displaying this slow-build, slow-emerge pattern included *glown----*, *glowies*, and *western chauvinism*. *Western chauvinism* is a term that the Proud Boys uses in its mission statement to describe its overarching ideology. Experts often refer to it as a replacement for *white supremacy*, in that the Proud Boys has sought to remain in the alt-lite and attract a more diverse membership, despite participating in REMVE actions and rhetoric (Southern Poverty Law Center, undated). *Glow---* and *glowies* and are discussed in detail below. Additional information on *western chauvinism*, as well as exemplar cases for the other emergence patterns, can be found in Appendix C.

Figure 3.1 shows a timeline of the emergence of *glown----*, which is used to indicate someone posting on 4chan whom other users say is a federal agent attempting to bait users into posting incriminating statements. It originated in response to a video posted on YouTube by Terry Baits on March 13, 2017, in which he said, “The CIA [redacted racial slur] glow in the dark, you can see them if you’re driving. You just run them over, that’s what you do.” Some sources indicate that the video appeared as early as October 2017 on 4chan, but the first usage of *glown----* in our data set was on December 4, 2017 (“Glowie,” 2022).

---

5 We acknowledge that there could be other patterns of emergence for REMVE terms outside of the ten we were able to research in depth for this study.

6 Information on the other two emergence patterns can be found in Appendix C.
From February to April 2018, there were significant spikes in rates of usage, but it was not until June 2018 that we first observed the term being used outside 4chan (on Reddit). We again saw higher rates of usage of the term from August 2018 to October 2019. The first example we found of the term being used in a non-REMVE context occurred in June 2019, when a user mentioned looking up the term on Urban Dictionary, a crowdsourced online dictionary for slang words and phrases. To determine whether a term was used in a REMVE or non-REMVE context, we read the context of the term as used and reported in the Brandwatch data set. All these data points help to build an emergence time profile.

In August 2019, there was a significant increase in use of the term glown----- in posts about Epstein and potential false-flag operations—when governments or others execute attacks or political maneuvers that they blame on another group or person. The term spiked again in April 2020 in response to COVID-19. Usage also spiked in response to civil unrest related to Floyd’s murder. This led to months of high usage of the term. Glown----- mostly stayed on 4chan, but this term and glowies were discussed in an article in The Atlantic on January 25, 2021 (Khazan, 2021). Our 4chan data end a few weeks before this article’s publication, so we...
cannot comment on the effect it had on 4chan usage, but we did see small increases on Reddit and Twitter after that publication.

As shown in Figure 3.2, glowies evolved from its more vulgar predecessor glow----. Previously used in unrelated ways as a band name, a jewelry brand, and a video game, glowies was first mentioned in a REMVE context on 4chan on February 28, 2018, when the term began exhibiting patterns of increased usage at an increasing rate. By June 2018, the term showed continuous usage on 4chan, although it did not show up in a REMVE context on Twitter until October 2019. On May 6, 2020, the first use of the term outside of the REMVE context occurred on Reddit—more than two years after its first appearance—and we deemed glowies to have emerged into the mainstream at this time. A thread on 4chan responding to riots and protests kicked off a period of consistently high levels of mentions for months. Although there were some significant spikes in usage during 2018 and 2019, it was more than six months after these spikes that we assessed the term to have moved into the mainstream because it was mentioned in the article in *The Atlantic* (Khazan, 2021). However, our 4chan data, in which this term was used primarily, end in December 2020.

Terms That Stay Popular After Going Mainstream and Long-Term Trends

We also noticed different patterns in the use of terms once they had emerged into the mainstream. Some terms so fully infiltrate the mainstream lexicon that they become a permanent

**FIGURE 3.2**

**Glowies Usage on Social Media**

![Glowies Usage on Social Media](image)

fixture in the social media landscape. Examples of such terms include *cuck, manosphere,* and *Judeo-Christian values.* Other terms can rise into the mainstream lexicon for a period of time and then eventually die out. These include such terms as *god emperor Trump,* which dropped off heavily after Trump lost the 2020 election. The other interesting category includes terms that appear to spike quickly and then disappear multiple times. Future research into the evolution of online REMVE language should consider the potential implications of these patterns for understanding (or even perhaps predicting) patterns of REMVE extremism and violence over time.

**A Summary of Emergence Patterns**

The findings about emergence patterns are intended to provide insights into the ways in which well-known REMVE terms have emerged during periods of sustained and increased usage before they have fully emerged into the mainstream public knowledge. Key findings were as follows:

- Some terms had multiple periods in which they met the qualifications we identified for being emerging, followed by periods of lower usage before they emerged into the mainstream.
- Terms can follow different patterns of emergence, and the time frames in which terms emerge into the mainstream vary significantly—from months to many years.
- Some terms eventually emerge into the mainstream somewhat organically, with their increased usage sparking the attention of non-REMVE users. Conversely, other terms emerge into the mainstream following specific events that bring the terms to public attention.
CHAPTER FOUR

The RVE-Flock Tool

In Chapter Two, we showed the various ways in which REMVE terms can emerge into usage. These results were produced without the use of RVE-Flock but helped inform the development of the tool. Specifically, the selection of time intervals and metric concepts were either informed or confirmed by the qualitative, term-specific research efforts described in Chapter Three. In this chapter, we describe the RVE-Flock Tool, which is designed to help the user track emerging REMVE terms as they peak in usage and become part of a community lexicon.

The tool’s primary purpose is to facilitate exploration of social media data using direct searches. RVE-Flock processes social media data to provide insight on queried terms by highlighting word associations that reflect the term’s meaning or context, emerging-term metrics on its usage, and topical context. As an NLP tool, it is intended to complement subject-matter expertise of users, integrate multiple data sources, and be regularly updated to process new data. This chapter provides an overview of the tool, its capabilities, and currently implemented metrics, as well as information about our methodology.

How RVE-Flock Works

The concept for RVE-Flock was inspired by the aphorism that “you shall know a word by the company it keeps,” attributed to British linguist John Rupert Firth, 1957, which suggests that the meaning of a term is tied to its context (Sadeghi, McClelland, and Hoffman, 2016). The tool’s name is a play on this idea and the better-known phrase, “birds of a feather flock together.” In short, new and emerging REMVE terms are likely to appear in the same contexts, or even the same social media posts, as well-established ones. By searching for a seed word, a user can find other words used in similar contexts on the same social media platform during a specified period. The tool is intended to be interactive (human in the loop) and used primarily for exploration of processed social media data.1 To this end, the tool provides quantitative measures of the criteria discussed in Chapter Three. Just as we evaluated a term’s

1 In this application of the tool, we focused on REMVE-affiliated language. However, to allow the tool to identify evolving language, it is not limited by a predetermined definition of what constitutes REMVE-affiliated language.
“emerging”-ness by the intensity of use and use in different contexts, the tool provides information in two ways:

- First, in its capability to facilitate exploration, it provides a list of words used in the most-similar contexts, along with quantitative metrics for each word’s topical dissimilarity and usage rates, so that the user is informed as to which words are trending or being found in different contexts. However, unlike the work done in Chapter Three, the tool lacks the capability to provide proliferation data.
- The second is the tool’s capability to provide the user with a list of words that have increased the most or have the greatest change in contexts between two time periods.

Next, we go into greater detail about the specific measures, their calculation, and utility.

**Methods and Functionality**

**Pairwise Word Topic Similarity**

RVE-Flock can identify other words that are used in similar contexts to a particular seed word. Latent semantic analysis (LSA), also known as *latent semantic indexing*, is a computational method for estimating the semantic meaning of words based on their occurrence in a body of text and other words that tend to occur with them (Dumais et al., 1988). The result of the analysis is a vector of real numbers that quantify the meaning of a word; vectors of two words can be compared to measure their similarity in meaning. Two words having similar meanings does not necessarily indicate that the two words frequently appear together in the same document. Instead, it means that the pair of words tends to appear with the same other words. As a simple example, *happy* and *glad* mean approximately the same thing, but it is unusual for someone to say, “I am happy and glad.” Instead, *happy* and *glad* tend to occur with the same other words in the context of a positive experience or sentiment. In the REMVE context, some people might use one epithet to refer to a group of people, others might use a different epithet, and still others might use a mainstream and innocuous term. The terms can all refer to the same group of people, even if they are rarely used together in the same passage. We measure the pairwise topic similarity by using cosine similarity between the two words’ vector spaces (Muflikhah and Baharudin, 2009; Tata and Patel, 2007). As similarity between two words increases, the cosine similarity will be a number that approaches 1.²

---

² In this application, we did not replace all numerical characters, so the tool can identify and process tokens that include numbers. The reason for this is that specific tokens or terms can include numbers of specific relevance to REMVE movements, such as *1488*, which is a reference to the 14 words (“We must secure the existence of our people and a future for white children”) and “Heil Hitler” (*H* is the eighth letter of the alphabet), or the QAnon slogan’s abbreviation: *WWGIWGA* (“Where we go one, we go all”). See Anti-Defamation League, 2022, and Rahn and Patterson, 2021.
Emerging-Term Metrics
Although RVE-Flock can provide some useful exploration of words and social media data, it also provides a set of three metrics that can help draw a user’s attention to language as it evolves:

- period topical dissimilarity
- intensity of use on a website
- proliferation of use across websites.

As the user runs a search, these three metrics will be presented with the search results, each providing a different facet that can inform judgments as to whether a term is emerging and warrants special attention. Although we discuss each metric in isolation, all three should be considered when using the tool to determine whether a term is new or emerging. That said, the tool provides these measures to the user for a sense of how language is shifting or evolving.

Period Topical Dissimilarity
One way to determine whether a word is emerging or has emerged is to understand the context in which the word is used and whether it is relatively stable or has changed rapidly over time. In our qualitative research of emerging terms, we found that there is a difference in the meaning or use of words once they transition to the mainstream, implying that the context in which these words appear is different from the context in which they originated. For example, as described in Chapter Three, glowies was first mentioned in the REMVE context in March 2017, became increasingly used in 2018, and showed its highest usage before breaking into the mainstream in 2019 when it was mentioned in reference to Epstein’s death. By identifying instances of topical dissimilarity, we can identify the different ways in which a word is being used either by specific groups or communities or within the mainstream.

To calculate the change in a term’s associated context, or topical dissimilarity, we rely on latent semantic vectors from RVE-Flock’s core LSA functionality from two time periods. If the contexts in which a term is found stay the same between two periods, there should be little difference in its similar words, insinuating that the meaning of the term is the same. If there is dramatic change in the contexts across the time periods, this suggests that the meaning of the term has changed. For example, if a term, such as glowies, was used in the mainstream as a popular music band but the far- and alt-right online communities began using it for something else, the topical dissimilarity of glowies would increase, highlighting it as an emerging-term candidate.

There are three ways to interpret topical dissimilarity based on context:

- First, it could mean that a word has been repurposed by a specific group or community. For instance, Turkish people on 4chan have been referred to as Turkroaches, roaches, or
cockroaches since about 2015 (O’Malley, Barlass, and Begley, 2019). The Christchurch shooter referenced Turkish people as roaches in his manifesto (SipDripTip, 2015).

- Second, it could mean that word has transitioned into the mainstream and is now associated with additional terms following greater attention on REMVE communities and their lexicon. For example, this could occur when people outside of those communities discuss and use cuck or when a major news outlet runs a story on REMVE communities and their lexicons.

- Finally, it could mean that there is some current event that is causing a topic change, as in the case of the blockage of the Suez Canal by a ship owned by the company Evergreen. This may cause the term evergreen to be more closely associated with new topics related to current events rather than how the word is typically understood. Determining which case is at play would require some awareness of current events or subject-matter expertise.

Without going into too much technical detail, the calculation of period topical dissimilarity relies on the same feature vectors that we use to identify pairwise word similarity for two words. For each word, we generated its pairwise similarity with every other word as a single vector for two time periods. We then take a measure of similarity between all pairwise vectors, comparing how that term’s pairwise similarity has changed between the two time periods. If the pairwise similarity has not changed significantly, the period topical dissimilarity should be low. However, if words’ pairwise similarities change dramatically between time periods, the period topical dissimilarity is high. Topical dissimilarity is presented in the tool as a percentile (0 to 100) to indicate how much—relative to all other words’ shifts in a period—a given word’s context has changed.

Intensity of Use on a Website
For a word to emerge and gain more traction, it must increase in use; otherwise, the term is used in isolation and cannot affect wider discourse. So, when a word is used with increasing frequency, it becomes likelier that to be considered emerging. RVE-Flock indicates whether a specific term is being used at an increasing rate. A term being used at an increasing rate could also suggest different things depending on the circumstances:

- First, it could suggest that the social media community is adopting the term and using it.
- Second, it could suggest that the term is increasingly used because there is some external event driving mentions, such as mentions of football leading up to the Super Bowl. Indeed, /pol/ responds to and discusses current political and cultural events in its most-discussed topics.

---

3 Etymology of the term is ascribed to a meme that began circulating on the internet in November 2015. See SipDripTip, 2015.
The calculation of this measure is the ratio between the relative increases in the latest and previous periods. In a sense, this is a ratio of ratios analogous to an increase of increases, normalized by being calculated relative to the previous time period. This is presented as a percentage in the tool to describe how it has increased.

Proliferation Across Websites
A measure that is not yet implemented in RVE-Flock is one that calculates the proliferation of the term across all different forums, channels, or sites. If a term becomes recognized across an entire movement, users of websites with similar user bases and topics would use it with increasing frequency. For example, 1488 is strongly associated with neo-Nazi movements, and the use of the term rose across multiple websites. Likewise, WWG1WGA (for “Where we go one, we go all”) is evident across different webpages and social media sites. This metric provides a sense of whether the term has matured and is recognized by the REMVE community. If a term proliferates widely enough, it will inevitably find itself in more public and less niche areas.

Although this is not yet implemented in the tool, the calculation of this metric would be the average increases over the previous period across all websites that are being observed and tracked.

User Interface
To process data, RVE-Flock requires only two fields per social media data set: a date and time stamp and a text field. Ideally, these fields should be aggregated in a way that is meaningful for users, depending on their use case (e.g., individual posts could be aggregated at the thread level for a tweet), and all hyperlinks, stop words, and hypertext markup language (HTML) tags should be removed. Currently, RVE-Flock is set up to work with 4chan, Stormfront, Reddit, and Discord data. The advanced version of the user interface (still in development) would provide a capability for users to perform similar searches using seed sentences in addition to seed words.

The screenshot in Figure 4.1 illustrates the results using the word nazi as a seed word for 4chan from June through October 2019. The graph shows usage of the word on the specified platform during the specified time period and how usage surged in August 2019 to double or triple its normal rate.

Beneath the graph in the left column are a few examples of how the word is used in context. The table lists the 50 words used in contexts most similar to the way nazi is used, as well as the counts for each of these terms over the same period of August to October 2019. As one might expect, there are references to Jewish people and symbols, as well as possibly anti-Semitic terminology and imagery; we caveat this as “possibly” because it is hard to say for certain without greater exploration.

To learn more about what any term might mean, the table also lists links to queries for each term on three websites: Know Your Meme (KYM), the Anti-Defamation League (ADL), and DuckDuckGo (DDG). The latter is a general-purpose web search site that does not track
users or their searches. In addition, the click-through searches can be bootstrapped, in the sense that each term in the table is hyperlinked to quickly generate a similar set of results, including a time-series chart and its own most-similar words. So, from these results, for example, a user might want to further explore the meaning, usage, and similar terms to the word jewsus (see Figure 4.1).

For each word, its relative similarity, total usage count in the latest time period, the percentage change in growth, and topical dissimilarity—referenced here as context change—is presented. Relative similarity is a quantitative measure of how close two words are to one another, with 1.0 being identical. "Usage count" is the total number of times the term was found in the previous period. "Usage change" is the percentage change in the usage counts for the selected time period. "Context change" is the topical similarity, with 0 being no change.
Synergizing Emerging-Term Metrics

Taken together, these three metrics can be used to determine whether a specific term is emerging and whether it needs further analysis. The topical dissimilarity metric will reveal whether, in the beginning of a term’s use, it has been repurposed to mean something else (e.g., *cuckold*) or its meaning is evolving. Furthermore, topical dissimilarity across time periods and increasing rate of use without proliferation to other platforms suggest that the term is emerging but is localized to a specific forum or website. If a term has topical similarity across time periods, the use of a term is no longer increasing, and it has proliferated to multiple websites without creating topical dissimilarity, then it has emerged and is a term that has a well-understood meaning among communities that span multiple webpages. Its proliferation, increase in use, and increase in topical dissimilarity can suggest multiple things but likely indicate that the term has transitioned to the mainstream. Table 4.1 describes how and whether terms could be said to be emerging in a community, given different values of our metrics.

The next section showcases RVE-Flock in action to determine whether a term is emerging based on these outlined metrics and is analyzed according to the qualitative analysis discussed in Chapter Three.

Infrastructure

RVE-Flock is written in Python and built to take advantage of PySpark and Spark NLP’s capability for large-scale parallel computing and works best when deployed on cloud computing infrastructure. Although PySpark and Spark NLP can be run on a local machine, deploying them into a cloud computing infrastructure is advantageous because operations can be scaled to take advantage of greater computing and memory resources (Kocaman and Talby, 2021; Zaharia et al., 2016).

| TABLE 4.1 Associations Between Emerging-Term Metrics |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Term                           | Topically Similar Between Time Periods |                      | Topically Dissimilar Between Time Periods |                      |
|                                | Has Not Proliferated | Has Proliferated | Has Not Proliferated | Has Proliferated |
| No increasing usage rate       | Isolated word         | Established word; not of interest | Potentially emerged; potentially responsive to an event | Potentially emerging; potentially responsive to an event |
| Increasing usage rate          | Potentially emerging; potentially responsive to an event | Potentially emerging; potentially responsive to an event | Likely has changed in meaning; potentially responsive to an event | Likely has emerged into the mainstream |
The Trending Lexicon

The aim of RVE-Flock is to assist in the discovery of REMVE terms as they emerge and evolve. Currently, the tool has the capabilities for manual exploration of language and calculates measures that characterize the extent to which a term could be considered to be emerging. To better inform the development of the tool and provide general guidelines and insights about similar methods, we applied RVE-Flock’s methodology to various time periods to those that correspond with the emergence and usage increase of the term *glowies*.\(^1\) As described in Chapter Three, the usage of *glowies* has increased over time. *Glowies* itself has entered mainstream awareness and is understood to generally mean undercover agents who post on forums—specifically, 4chan—to spy on or otherwise entrap users.

In this chapter, we discuss two general capabilities:

- The first is the tool’s exploratory capability as applied to *glowies* that provides sets of words used in similar contexts measured for the criteria discussed in Chapters Three and Four.
- The second capability leverages the metrics provided by the tool to identify emerging terms throughout different time periods. Specifically, we applied RVE-Flock to evaluate whether it could identify *glowies* when used in /pol/ during two time periods:
  - The first period is between December 2016 and March 2017 when the term was first introduced. Recognizing that new terms are likely to see low levels of usage, we varied our time frames at and below the monthly levels.
  - We also examined time windows around August 2019 for our second period of analysis. This is when the term increased in usage and appears to have experienced a large wave of popularity in response to Epstein’s death, as illustrated in Figure 3.2 in Chapter Three. Considering that the term is several years old at this point and is likely to have found some wider-spread use, we analyzed /pol/ data in three-month periods between July 2019 and October 2019.

\(^1\) In the full version of RVE-Flock, there is also a metric for the proliferation of a term across websites and forums, as discussed in Chapter Four. Because of the focus on 4chan and /pol/ exclusively, that metric is not applicable here.
We relied on /pol/ data that were previously collected by another internally funded RAND research effort (H. Williams et al., 2022). However, because we focused on the term’s evolution on a single board, we were unable to evaluate its proliferation to other boards or websites and so are able to provide data only on its increasing growth in usage and its topical dissimilarity.

In result tables in this chapter, we indicate whether and how any words we identified were related to REMVE communities. We identified specific words based on whether they were racial or ethnic slurs or whether additional research indicated some connection with broader REMVE movements or figures who were related to them. This is the type of iterative interaction we envision users to have with the tool.

RVE-Flock Exploratory Capability

RVE-Flock has an exploration function with which a user can search a term and find other associated terms, topical dissimilarity, and growth rates during a specified time frame. Figure 5.1 showcases what RVE-Flock would produce for a search for *glowies* between July and October 2019. As mentioned in Chapter Four, values closer to 1 suggest greater topical similarity for previous three-month time period (August to October 2019). The usage count

<table>
<thead>
<tr>
<th>Term</th>
<th>Learn Mass</th>
<th>Relative Similarity</th>
<th>Usage Count</th>
<th>Usage Change</th>
<th>Content Change</th>
<th>Scored 4 to 368 (Total Score Change)</th>
</tr>
</thead>
<tbody>
<tr>
<td>glowies</td>
<td>EXP AI</td>
<td>0.8</td>
<td>2520</td>
<td>13.6%</td>
<td>56.8</td>
<td></td>
</tr>
<tr>
<td>mutt</td>
<td>EXP AI</td>
<td>0.737</td>
<td>2822</td>
<td>11.2%</td>
<td>40.3</td>
<td></td>
</tr>
<tr>
<td>understand</td>
<td>EXP AI</td>
<td>0.688</td>
<td>11</td>
<td>N/A</td>
<td>38.6</td>
<td></td>
</tr>
<tr>
<td>artiest</td>
<td>EXP AI</td>
<td>0.598</td>
<td>1698</td>
<td>7.9%</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>avatar</td>
<td>EXP AI</td>
<td>0.595</td>
<td>768</td>
<td>1.5%</td>
<td>83.3</td>
<td></td>
</tr>
<tr>
<td>avatar1</td>
<td>EXP AI</td>
<td>0.564</td>
<td>194</td>
<td>3.3%</td>
<td>86.1</td>
<td></td>
</tr>
<tr>
<td>avatar2</td>
<td>EXP AI</td>
<td>0.549</td>
<td>536</td>
<td>1.6%</td>
<td>9.9</td>
<td></td>
</tr>
<tr>
<td>avatar3</td>
<td>EXP AI</td>
<td>0.541</td>
<td>496</td>
<td>2.9%</td>
<td>8.9</td>
<td></td>
</tr>
<tr>
<td>avatar4</td>
<td>EXP AI</td>
<td>0.540</td>
<td>488</td>
<td>2.9%</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td>avatar5</td>
<td>EXP AI</td>
<td>0.539</td>
<td>588</td>
<td>1.4%</td>
<td>7.2</td>
<td></td>
</tr>
<tr>
<td>avatar6</td>
<td>EXP AI</td>
<td>0.538</td>
<td>1619</td>
<td>2.1%</td>
<td>58.8</td>
<td></td>
</tr>
<tr>
<td>avatar7</td>
<td>EXP AI</td>
<td>0.535</td>
<td>207</td>
<td>0.6%</td>
<td>41.4</td>
<td></td>
</tr>
<tr>
<td>avatar8</td>
<td>EXP AI</td>
<td>0.532</td>
<td>280</td>
<td>1.9%</td>
<td>26.3</td>
<td></td>
</tr>
<tr>
<td>avatar9</td>
<td>EXP AI</td>
<td>0.532</td>
<td>29022</td>
<td>1.4%</td>
<td>24.0</td>
<td></td>
</tr>
<tr>
<td>avatar10</td>
<td>EXP AI</td>
<td>0.531</td>
<td>1658</td>
<td>10.8%</td>
<td>38.6</td>
<td></td>
</tr>
<tr>
<td>avatar11</td>
<td>EXP AI</td>
<td>0.529</td>
<td>497</td>
<td>7.0%</td>
<td>41.4</td>
<td></td>
</tr>
<tr>
<td>avatar12</td>
<td>EXP AI</td>
<td>0.527</td>
<td>280</td>
<td>1.9%</td>
<td>26.3</td>
<td></td>
</tr>
<tr>
<td>avatar13</td>
<td>EXP AI</td>
<td>0.522</td>
<td>39222</td>
<td>1.4%</td>
<td>24.0</td>
<td></td>
</tr>
<tr>
<td>avatar14</td>
<td>EXP AI</td>
<td>0.513</td>
<td>1658</td>
<td>10.8%</td>
<td>38.6</td>
<td></td>
</tr>
<tr>
<td>avatar15</td>
<td>EXP AI</td>
<td>0.511</td>
<td>497</td>
<td>7.0%</td>
<td>41.4</td>
<td></td>
</tr>
<tr>
<td>avatar16</td>
<td>EXP AI</td>
<td>0.509</td>
<td>280</td>
<td>1.9%</td>
<td>26.3</td>
<td></td>
</tr>
<tr>
<td>avatar17</td>
<td>EXP AI</td>
<td>0.508</td>
<td>280</td>
<td>1.9%</td>
<td>26.3</td>
<td></td>
</tr>
<tr>
<td>avatar18</td>
<td>EXP AI</td>
<td>0.507</td>
<td>280</td>
<td>1.9%</td>
<td>26.3</td>
<td></td>
</tr>
<tr>
<td>avatar19</td>
<td>EXP AI</td>
<td>0.506</td>
<td>280</td>
<td>1.9%</td>
<td>26.3</td>
<td></td>
</tr>
<tr>
<td>avatar20</td>
<td>EXP AI</td>
<td>0.505</td>
<td>280</td>
<td>1.9%</td>
<td>26.3</td>
<td></td>
</tr>
</tbody>
</table>

2 In the full version of RVE-Flock, there is also a metric for the proliferation of a term across websites and forums. Because of the focus on 4chan and /pol/ exclusively, that metric is not applicable here.

FIGURE 5.1

RVE-Flock Search Results for *Glowies*, July–October 2019
provides the total number of times this term was found in the latest time period. The usage change value conveys the change in increases over the time period or the percentage change from the previous time period, August to October 2019, to the previous time period, July to September 2019. Context change provides a sense of how different the contexts are in which a specific term was found, relative to all other changes in the data set.

The exploratory capability enables the user to evaluate pairwise similarity of search terms with others and provide metrics that help characterize the extent to which a term could be said to be emerging. The trend line in the graph on the left-hand side of the RVE-Flock output shows daily usage of glowies. Underneath the graph are sample texts in which glowies was found (e.g., “It would not surprise me at all if glowies were monitoring those threads”).

RVE-Flock provides the pairwise cosine similarity between the searched word and every other word in the “Relative similarity” column. RVE-Flock also produces a usage count, the total number of times that term was observed in the latest time period in the “Usage count” column. RVE-Flock calculates the intensity of word use in the “Usage change” column to signify whether the term’s usage has increased over the time period, presented as a percentage. Values close to 0 suggest relatively small increases, values below 0 suggest declining growth, and values above 0 suggest that the increase over the latest period is greater than the increase over the previous. RVE-Flock also calculates topical dissimilarity, which shows how similar the contexts are between the current and previous time periods for each term in the “Context change” column. The higher the topical dissimilarity value is, the greater the contextual differences are.

As shown in Figure 5.1, glowies lacks strong similarity to many of the terms on 4chan, but, given the terms highlighted, one could infer its association and usage within the REMVE community. For example, the closest word in similarity is mutt, which can reference two things:

- a derogative way of referencing someone of mixed racial or ethnic descent
- Mutt’s Law, which, in a similar fashion to Godwin’s Law, holds that the longer any discussion goes, the probability of an American (referred to as an Amerimutt or mutt) mentioning interracial sex approaches 1 (“Mutt’s Law,” KYM, 2022; “Mutt’s Law,” Urban Dictionary, 2020).

Wignat, meanwhile, references someone who is distinct from the modern Nazi movement and tries to portray themselves as more socially acceptable with a cleaner image by being violent and willing to engage in street fights while displaying Nazi and other white supremacist iconography liberally—thereby undermining the effort to make Nazism more tolerable to the public (“Wignat,” undated). Authorities is relatively simple—it likely reflects government authorities, such as the police or other law enforcement actors. The combination of terms on

---

3 Godwin’s Law holds that, as internet discussion continues, the probability of a comparison being made to Nazis or Adolf Hitler approaches 1 (Ohlheiser, 2017).
this list allows users to infer some meaning about *glowies* and how the term is used online. Although it is taken as a potential threat and something of which posters are wary, the fact that *glowies* is relatively obvious, over the top, and likely authorities and distinctly American could lead users to better understand its use on 4chan.

Figures 5.2 and 5.3 present the results for *glown-----* and *glowies* in /pol/’s board from August 2017 to October 2019, which includes the period in which the term was flagged as emerging. The blue line represents the intensity of use and the red line is the topical dissimilarity, where higher numbers indicate greater topical dissimilarity and lower numbers indicate greater similarity. Recall that the blue line is a ratio of growth ratios over the previous periods, so values greater than 1 mean that use of the term is growing at an increasing rate. Sharp declines following sharp increases do not necessarily mean that the term is falling in popularity or use—recall that figures in Chapter Three show these terms increasing in use over time—but it is being used at a slower rate than previous periods. Periods in which the red line is interrupted are those in which *glowies* was not used frequently enough to be processed by the tool. Both intensity of use and topical dissimilarity are metrics the tool calculates and leverages to present rules to the user. Note that, in the earliest parts of this timeline, the ratios are inflated because of the terms’ relatively low use. In the two spikes in term growth, in April 2019 and August 2019, the topical dissimilarity of the term declines, suggesting that the term finds itself in consistent “company” as the term is used in far greater numbers. However, in the period after April 2019, topical dissimilarity steadily increases until the end of this obser-

![FIGURE 5.2](image)

**Intensity of Use and Topical Similarity for Glowies, August 2017–October 2019**

NOTE: A gap in a line indicates a period in which the term was not used frequently enough for the tool to process it.
vation window when it dips—albeit nowhere near its lowest value—which was marked in our qualitative review of the term in October 2019, when the term began to emerge.

Figure 5.3 provides similar information for the more vulgar variant, *glown-----*. Here again, the blue line represents the intensity of use through time and the red line represents the topical dissimilarity. As mentioned in Chapter Three, the term was considered to be emerging starting in August 2018, which is the first opportunity to calculate its topical dissimilarity. In our qualitative review of the social media posts around this time, we noted that the term gained popularity in August 2019 because use of the term seemed to increase with mentions of Epstein’s death. In this period, the topical dissimilarity declines because the word would be keeping more-consistent company (e.g., the topic and related words around Epstein’s death) while increasing in use. After September, use of the term and its topical dissimilarity seem to increase in tandem with increasing intensity of use. Table 5.1 provides a sample of words associated with each of these two words during different time periods. The sampling of words is those that would be observed by users of the RVE-Flock tool at different time periods had they searched for either *glown-----* or *glowies*.

Several themes arise from Table 5.1. *Glown-----* and *glowies* referring to authorities is evident on several occasions but especially in the earliest periods (e.g., *auditing, designers, suppressors, supervisors, dissension*) with authoritarian overtones (e.g., *totalitarian, jail, gulag*) in the period before August 2018. During and after September 2018 until January 2019, the theme that *glowies* takes is one of some established group (e.g., *humanitarians, synagogue, councils, antiscience, institutes, humanities, researchers*) with state-aligned leanings (*propagandists*) while being used in similar contexts to slurs that could be used against Americans (e.g., *muttboy, muttoid, pug*). As time went on, however, the use of these terms to denote the insincerity of *glowies* online is suggested (e.g., *astroturfers, lobbyist, coconspirator, wrongpro- *

**FIGURE 5.3**
Intensity of Use and Topical Similarity for *Glown-----*, August 2017–October 2019
### TABLE 5.1
Sample of Words with High Topical Similarity with *Glowies* and *Glown-----*

<table>
<thead>
<tr>
<th>Year and Month</th>
<th>Glown-----</th>
<th>Glowies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2018</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May</td>
<td>countries, designers, dissension, lowskilled, norways, opportunistic, revisionism, robbed, ropers, smuggle, sucker, suppressors</td>
<td>announces, auditing, brudda, germanlanguage, gloria, propagating, slaughter, tranniesi, unrighteousness</td>
</tr>
<tr>
<td>June</td>
<td>butts, ciggies, diversifying, dulles, incessantly, islands, personnel, pipeline, rationalizing, sender, tit, tuned, weevs, westernization</td>
<td>calculator, catastrophically, conquistador, emanatism, irreligious, opportune, paperwork, sanitation, supervisors, totalitarian, transexual, vanish</td>
</tr>
<tr>
<td>July</td>
<td>antigay, combatants, doubles, feathered, handleman, insidious, insufferably, jail, jake, jewjesus, lures, mausoleum, paula, pierced</td>
<td>anticop, castrated, elections, eloquent, fratricide, gulag, shekelstein, superman, takeing</td>
</tr>
<tr>
<td>August</td>
<td>antiscience, azrbaycan, barcelona, browbeaten, councils, qlarp, guilted, humanitarians, innate, lanza, naziboos, rasmussens, selfadmitted, spectacle, synagogue</td>
<td>autos, bloodless, contra, friend, humanities, institutes, irises, leafhheet, mtdna, muttboy, propagandists, researchers, selfidentifying, shitcunt, shooting, tickles, unabashed, yenta</td>
</tr>
<tr>
<td>September</td>
<td>airline, appointed, bloke, bloodlust, catal, commence, contaminated, contractions, defeats, gracious, herwell, intervenes, kkz, mankind, mideast, mordern, motive, muttoid, others, proslam, pug, residency, seeped, shariablue, spanned, stuttering, supervisors, vatavaran, veterinary, wagenknecht</td>
<td></td>
</tr>
<tr>
<td>October</td>
<td>alibi, antisreali, astroturfers, buddhas, cinematic, cleaver, coerced, contrarion, criminaland, demoshits, derelict, francis, hospitals, kala, landfills, lgbts, predators, proud, response, starkly, strive</td>
<td></td>
</tr>
<tr>
<td>November</td>
<td>airlines, antagonize, constitution, cynic, debatable, degeneration, energize, fulfils, gds, guardok, hook, lightweight, melanin, merriment, nonnatives, pioneered, progs, prophesied, rebecame, removing, righties, situation, sundar, survived, unbox, unpersoning, vm</td>
<td></td>
</tr>
<tr>
<td>December</td>
<td>afflicting, aged, assume, bruises, championship, condo, disfunctional, earthshattering, energetic, expelling, flagged, flicks, folding, hijack, kali, kicker, offputting, plugged, receipt, regiment, richard, venezuala, warframe, whispering, yw</td>
<td></td>
</tr>
</tbody>
</table>
## Table 5.1—Continued

<table>
<thead>
<tr>
<th>Year and Month</th>
<th>Glown------</th>
<th>Glowies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2019</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>January</td>
<td>arguing, banzai, bimbos, concession, cools, detroit, entrepreneurs, exceedingly, exponentially, foot, grasping, hardy, hailed, henrik, hiking, impossible, johannesburg, lima, locusts, redistributionist, reformed, stopped, siren, tolerable, underdogs, weakness, wholeheartedly, wigger, worldshits, write</td>
<td>brampton, bury, contrarily, crossdressing, cusp, dialogue, dotted, dutertes, futile, generous, jew2, karelia, kedar, kristalnacht, libera, lobbyist, meaningless, motherf***ing, nerf, raped, sanctions, sexbot, spotify, trivialize, wraps, zinovieff</td>
</tr>
<tr>
<td>February</td>
<td>absent, bunker, butina, calibrated, certification, charges, counterculture, drilling, drive, energy, engages, erasing, immigrant, injecting, legislature, lynched, manlet, match, nearly, ozone, publicly, rightlets, rocking, scored, segregation, spiraling, supremes, triumphed, unregistered, unsuccessful, weerwolves</td>
<td>buy, conventional, destabilization, disgusted, dork, elvis, grams, heap, impassable, inhaling, linguistic, moneychangers, mouthing, neighborhoods, nuremberg, organizes, pinned, preserves, psych, reactivating, refuse, saffers, shilsl, skal, tolland, upgraded, zika</td>
</tr>
<tr>
<td>March</td>
<td>academics, babble, birther, brighton, cascading, employed, enraged, figure, gov, gyroscope, homemakers, indirect, interventions, limited, mgtow, proficiency, publically, roll, rotated, saga, sneering, tattooed, thirty, unvetted, vonhelton</td>
<td>autobiography, babe, bartender, cathedrals, commiefornia, complaining, crowley, delightful, file, forth, goli, jerusalem, labarumchirho, lgbt, malicious, masterace, normally, overthrow, php, practise, smirked, sports, stubby, tapes, velvet, wounded</td>
</tr>
<tr>
<td>April</td>
<td>agreement, astroturf, charlemagne, checkup, cohesion, colonized, contend, dial, divider, donuts, driven, ebola, editors, fermented, ftm, gavin, graaave, gypsy, harmony, honda, hypocrite, knocks, narrowing, oculus, prequel, radiates, religion, weeb, weak</td>
<td>basedjew, battlefield, berliners, buffalo, cleaned, conditional, dualcitizens, educated, fixable, gershwin, gummint, incite, insinuates, margin, monkeying, mutate, opie, ore, possibility, righters, roddy, routine, sausage, surpass, thermonuclear, viola, wataala, yeshua, zealously</td>
</tr>
<tr>
<td>May</td>
<td>albertans, campaigns, christinsanity, <strong>coconspirator</strong>, cpd, deposition, existence, garvey, ghwb, haj, hapsburgs, hilarious, ink, log, normalfags, overbear, persist, plea, roommates, scaled, slush, theoretical, toilet, torch, vasectomy</td>
<td>amerilards, architects, bootlicking, bundestag, choppa, coincidence, devices, door, emirates, esta, filipino, gaybear, goyboy, horns, moonflossing, obst, observations, policed, polter, predate, proximity, roses, sammich, sever, valuables, wrongprotop</td>
</tr>
<tr>
<td>June</td>
<td>911, alquaeda, beatings, canadatier, chump, dudes, factorio, glasnost, habeas, inventor, metabolism, opt, pompeii, prepares, pumping, quando, remittances, soyboy, toilets, worldly, yaya</td>
<td>anonymously, byproducts, chow, confucianism, desired, determines, discount, dusted, evacuation, frente, fridge, frightened, frustrations, hyperloyalty, institute, intrinsic, menwhite, minimalist, pranked, seething, shackles, silversteins, splattered, tied, yeltsin</td>
</tr>
<tr>
<td>July</td>
<td>antibrexit, authoritative, beginnings, bonnet, cato, chateau, dousing, earthy, <strong>encrypted</strong>, instantly, lowquality, muttin, norway, pax, portions, rationalization, renewables, spectacularly, stamps, surged, symptomatic, watchlist, wtfi</td>
<td>albums, angel, ebony, fairytale, fascinated, gallons, honesty, honour, internationalism, leftovers, muddied, observations, raygun, roommate, roomstay, shrinking, strapping, ticking, universe, unsurprising, vending, woops</td>
</tr>
</tbody>
</table>
Countering Domestic Racially and Ethnically Motivated Violent Terrorism on Social Media

...tip, monkeying), authoritarian nature (e.g., bootlicker, corporatist, ecofash, cybernazi, totalitarianism, reactionaries). Of special interest in the period are the associations between July 2019 to September 2019. During this period, topical dissimilarity remained relatively high for glowies while it declined for glowin----. As mentioned earlier, in August 2019, Epstein died and an increase in usage of both terms is observed. However, glowies was the relatively new one. In this time period, glowin---- appears to be more associated with clandestine activities (e.g., psyoped, secrets), which is similar to themes from earlier periods (e.g., watchlist, encrypted), while increasing in use, while glowies transitions from ideological agendas (e.g., commiefornia, confucianism, institute, internationalism) to more-aggressive ones (e.g., battle, aggressor, strife). This dramatic swing can be observed in Figure 5.2 for this time period by its relative stability, which signifies that the topic and context are evolving from period to period. Referencing Table 4.1 in Chapter Four, a term that is increasing in use but does not have topical dissimilarity could be considered potentially emerging and potentially responsive to events depending on whether it has proliferated to other websites. From the trajectory of glowin---- in Figure 5.3, it is clear that August 2019 is a time period in which the term could be considered to be emerging in either case. The period of March 2019 for glowies follows a similar pattern of either being potentially emerging or in response to an event. However, there is another spike, also in August 2019, but here the topical dissimilarity remains relatively high, suggesting that, at that point, glowies has likely changed in meaning, potentially is responsive to an event, or likely has emerged into the mainstream if it proliferated to...
other websites. Although we focused on one platform, these statistics likely reflect that it has emerged into the mainstream.

In this section, we have examined the evolution and development of *glown----* and *glowies* using two types of emerging-term metrics that we discussed in Chapter Four: increasing usage rate and topical dissimilarity. Increasing usage rate is important because a community’s familiarity with and widespread use of the term will dictate whether it becomes a meaningful part of a community’s vocabulary. The second metric, topical dissimilarity, reflects how the term’s place in topic space (alternatively, the numeric description of its connections to different topics) changes between time periods. A change in the topic space changes or a term beginning to be collocated with different sets of words suggests that there are changes in the meaning of the word. Taken jointly, RVE-Flock can describe not only how word usage is increasing or decreasing in each internet community but also whether the word context is changing in a meaningful way. In the next section, we identify the terms that have increased in usage or have changed the most in select time periods.

RVE-Flock Emerging-Term Identification

Identifying Emerging Terms as They Are Coined, December 2016–March 2017

In an effort to determine whether RVE-Flock would be able to flag *glowies*, the first time interval we considered was December 2016 to March 2017, which covers the initial introduction of the term *glowies* (more specifically, *glown----*). We consider a three-month moving window to identify how terms change in their frequency of use and their surrounding context to track larger, more-general changes. If a term’s meaning is evolving, it could be found in different contexts as it evolves. Box 5.1 highlights a few of these words for which the topical context changed most between December and March 2017.

Of the terms with the greatest topical dissimilarity, *bantz*, a word to describe offensive banter, is especially noteworthy because its identification in this time period precedes its definition on *Urban Dictionary* despite being known for some time. This indicates that it might have been specific to REMVE communities prior to its definition on *Urban Dictionary*. The earliest identification of *bantz* was in August 2006, when it was characterized as a state of mind. The British Broadcasting Corporation (BBC) recognized the term in 2014, when an article noted that it was used to describe offensive, racist, or sexist banter in a way that trivializes those who are offended by it (Sandhu, 2015; “The Backlash Against Banter,” 2014). In 2015, *bantz* was added to the *Oxford English Dictionary* as “playfully teasing or mocking remarks exchanged with another person or group, esp. among men; banter” (“bants,”

---

**BOX 5.1**

**Terms with the Greatest Topical Dissimilarity, December 2016–March 2017**

<table>
<thead>
<tr>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recollection</td>
</tr>
<tr>
<td>Bantz</td>
</tr>
<tr>
<td>Northern</td>
</tr>
<tr>
<td>Reflecting</td>
</tr>
<tr>
<td>1914</td>
</tr>
<tr>
<td>Halfway</td>
</tr>
<tr>
<td>Groping</td>
</tr>
<tr>
<td>Awesomeness</td>
</tr>
<tr>
<td>Ivanka</td>
</tr>
</tbody>
</table>

*Potentially REMVE-related term.*
undated) based on its informal United Kingdom interpretation. RVE-Flock detected *bantz* in the REMVE, U.S.-centric context prior to its definition on *Urban Dictionary*, in which it is defined as less-than-playful mocking. As shown in Box 5.1, for words with the greatest topical dissimilarities between December 2016 and February 2017, *bantz* was within the ten most-dissimilar terms. According to the *Urban Dictionary* definition entered in May 2017, *bantz* is slang term meaning “a spirited discussion between two people who share language and oppose ideals for the purpose of entertainment. Often heavy and violent, at least one source views it ‘in good fun’” (“Bantz,” undated). *Bantz* was later showcased in a quote from 4chan in a 2020 academic paper on the challenges of studying the alt-right on 4chan: “‘4chan is a hateful meme factory, promotes free speech, and needs to be shutdown.’—Academic fags. These people are idiot savants that can’t handle the bantz” (Colley and Moore, 2022).

Although this term might not be code for anything nefarious, it was around this time that the term began to be used on 4chan with greater intensity as something that posters on the site valued (Tuters and Hagen, 2020). As shown in Figure 5.4, *bantz* experienced little use on 4chan and Reddit compared with how often it was observed on Twitter.
Figure 5.5 shows the use of the *bantz* on 4chan.org.\(^4\) Note that the use of *bantz* increased until the end of 2016. This increase can be attributed to activity on the website that considered some of the divisive language that arose during the campaign season. This is evident from the linear increase in daily minimum mentions until the end of 2016 and returning to a lower level in 2017, before reaching relatively low levels after 2019. The fact that RVE-Flock flagged *bantz* as emerging during its high usage levels before it seemed to have stabilized at a lower level is useful and provides a demonstration of how the tool can provide or identify language during a period in which it was used extensively.

As Figure 5.6 shows, searching for *bantz* in the tool returns such words as *poisonous*, *taunting*, and *cheery*—results that are similar to the meaning as it is understood and used on 4chan. Had this term been flagged by people who were unaware of it, searching the term on RVE-Flock would provide words that could help triangulate its meaning.

Next, we considered term usage to evaluate whether usage of the terms was growing at an increasing rate during the same time period. To prevent the likelihood of small or unknown terms skewing the top results, we required that each term have a minimum of ten mentions per week between December 2016 and March 2017. Table 5.2 presents terms and the ratio of their usage in the two three-month time periods. The term with the highest increase in their rates of increase between these periods is *FVD*, an abbreviation for *Forum voor Democratie*

---

\(^4\) This version of the data set is limited, but it bears mentioning that the use of *bantz* can be found as early as 2014.
FIGURE 5.6
RVE-Flock Search Results for Bantz, December 2016–March 2017

(Forum for Democracy), the far-right Dutch party, which was founded in 2016. RVE-Flock also identified Geert and Wilders. Geert Wilders and the political party he founded practice anti-immigrant politics ("Dutch Far-Right Leader Geert Wilders Will Face Trial for Hate Speech," 2016). RVE-Flock also highlighted the self-identified socialist, Steve Bonnell, aka Destiny, which we attribute to his March 2017 debate with fellow YouTuber Jon Jafari (aka JonTron), a far-right personality who made highly controversial remarks on race and immigration (Gajanan, 2017).

Box 5.1 and Table 5.2 provide two sets of terms: the most topically dissimilar and the greatest growth rates. The two sets of terms illustrate different aspects or dimensions of emerging terms and topics that can be important in different ways, depending on the context of terms and the interests of users. In application of the tool, a user is likely going to examine more than 20 terms, as we highlighted here, which makes the user’s background knowledge and subject expertise crucial. Informed by our own experience and subject-matter expertise, we identified bantz as a potential emerging REMVE term and highlight FVD, Geert, Wilders, and Destiny as associated with REMVE communities and discussions.

Identifying Emerging Terms as They Are Coined, March 5–20, 2017

Next, we consider how the tool and methods perform at a finer level of granularity in our analysis of glowies because this term might have been so new that it was overshadowed by
words with greater usage and stability. We considered whether week-to-week evaluations would be able to identify *glowies*.

The second method we evaluated is RVE-Flock’s ability to identify terms that were being used at an increasing rate during the same time periods. For example, *sailor* was used almost 18 more times between March 13 and 20 than between March 5 and 12, as shown in Table 5.3. Given these results, a weekly level of analysis might be too granular for tracking substantive or significant changes in language because the tool seems to be very sensitive to current events because terms referencing public figures appear to experience the highest growth in use. *Schiff* and *Nunes* likely reference U.S. Representative Adam Schiff and then–U.S. Representative Devin Nunes in the wake of news surrounding claims of collusion and Russian involvement in the 2016 election (Raju and Schleifer, 2017). Additionally, *Destiny* here again likely also references the streamer who debated YouTuber JonTron (see *Destiny, debated*, and *Jon*).

<table>
<thead>
<tr>
<th>Term</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>FVD(^a)</td>
<td>14.81</td>
</tr>
<tr>
<td>VVD</td>
<td>10.91</td>
</tr>
<tr>
<td>Geert(^a)</td>
<td>9.18</td>
</tr>
<tr>
<td>PVV(^a)</td>
<td>8.57</td>
</tr>
<tr>
<td>Erdogan</td>
<td>8.11</td>
</tr>
<tr>
<td>Destiny(^a)</td>
<td>7.33</td>
</tr>
<tr>
<td>Wilders(^a)</td>
<td>7.25</td>
</tr>
<tr>
<td>Dit</td>
<td>7.10</td>
</tr>
<tr>
<td>Hebben</td>
<td>6.66</td>
</tr>
<tr>
<td>Thee</td>
<td>6.28</td>
</tr>
</tbody>
</table>

NOTE: PVV = Partij voor de Vrijheid, or Party for Freedom. VVD = Volkspartij voor Vrijheid en Democratie, or People’s Party for Freedom and Democracy. \(^a\) Potentially REMVE-related term.

---

5 To examine whether dissimilarity was being accurately captured and that these terms are not just the result of random noise, we also assessed the overlap in the most-similar terms for these same time periods. If the week-to-week time scale is too granular and susceptible to noise, there should also be no overlap among the most-similar terms from week to week. However, we found that there were around three times as many overlapping terms among the 500 most-similar terms as among the 500 most-dissimilar terms between these two time periods. This suggests that the results presented here are not entirely the result of random noise.
Identifying Emerging Terms as They Proliferate, June–October 2019

The previous two time periods centered on when the term glowies and its variants would have just begun to be used. It is possible that, during the very initial stages of a term’s popularization, its usage is too small to be identified by methods and tools like RVE-Flock. As seen in Figure 3.2 in Chapter Three, there is a clear usage increase of glowies around August 2019. So, to see whether RVE-Flock would be able to identify glowies as its use grew at an increasing rate on 4chan, we considered the period between June and September 2019, when we knew that the term increased in use (see Table 5.4). Between the two time periods, there appears to be overlap with gang, Khan, blizzard, and arse being present in each list. Given that two months are shared across both columns, this suggests that either August or September was responsible for changing contexts of these words. Additionally, RVE-Flock did flag Chi-coms, an offensive or disparaging reference to Chinese Communists, as shown in Table 5.4 (“Chicom,” Dictionary.com Unabridged, undated; “Chicom,” Merriam-Webster.com Dictionary, undated).

Table 5.5 provides RVE-Flock results on term growth rates. As was the case with Box 5.1, the terms that populate the top of these list seem to be those that reference current events or associated public figures. Entries for Turkish President Erdogan and the Kurds could be referencing events in October, when Turkey became more involved in Syria and Kurdish regions (McKernan, 2019). The word pewds could be referencing YouTuber PewDiePie, Felix Kjellberg, who announced and canceled a donation to ADL and had previous controversies that involved paying two men to hold signs that said “DEATH TO ALL JEWS,” for which he later

<table>
<thead>
<tr>
<th>March 5–12 Versus March 13–20</th>
<th>March 13–20 Versus March 21–28</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term</td>
<td>Ratio</td>
</tr>
<tr>
<td>Sailor</td>
<td>17.91</td>
</tr>
<tr>
<td>Schiff</td>
<td>15.55</td>
</tr>
<tr>
<td>Zap</td>
<td>15.50</td>
</tr>
<tr>
<td>Destiny(^a)</td>
<td>14.82</td>
</tr>
<tr>
<td>Debated(^a)</td>
<td>13.82</td>
</tr>
<tr>
<td>Gifs</td>
<td>13.09</td>
</tr>
<tr>
<td>Jon(^a)</td>
<td>13.27</td>
</tr>
<tr>
<td>Maddow</td>
<td>12.34</td>
</tr>
<tr>
<td>Dawg</td>
<td>11.68</td>
</tr>
<tr>
<td>Exit</td>
<td>11.62</td>
</tr>
</tbody>
</table>

NOTE: SDF = Syrian Defense Forces. To guard against the likelihood that a word mentioned once shows up at the top of this list, we limited our analysis to words that were mentioned an average of at least ten times a week.

\(^a\) Potentially REMVE-related term.
apologized (Andrew, 2019). Furthermore, *roach* and its variants are often used as derogatory slurs to reference Turkish people; the appearance of *roach* during this period could also be in response to then-current events. Meanwhile, *NDP* (perhaps for New Democratic Party), *Ber-

TABLE 5.4
Topical Dissimilarity, June–October 2019

<table>
<thead>
<tr>
<th>Term</th>
<th>June–September 2019</th>
<th>July–October 2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eventually</td>
<td>Arse</td>
<td></td>
</tr>
<tr>
<td>Sport</td>
<td>Gang</td>
<td></td>
</tr>
<tr>
<td>Gang</td>
<td>Blizzard</td>
<td></td>
</tr>
<tr>
<td>Khan</td>
<td>Retardedly</td>
<td></td>
</tr>
<tr>
<td>Weird</td>
<td>Foundations</td>
<td></td>
</tr>
<tr>
<td>Inflicted</td>
<td>Chicoms</td>
<td></td>
</tr>
<tr>
<td>Blizzard</td>
<td>Khan</td>
<td></td>
</tr>
<tr>
<td>Perpetuated</td>
<td>Gookmoot</td>
<td></td>
</tr>
<tr>
<td>Arse</td>
<td>Argued</td>
<td></td>
</tr>
<tr>
<td>Lodge</td>
<td>Dilute</td>
<td></td>
</tr>
</tbody>
</table>

*Potentially REMVE-related term.*

TABLE 5.5
Terms with the Largest Increases in Use, June–October 2019

<table>
<thead>
<tr>
<th>Term</th>
<th>June–September 2019</th>
<th>Ratio</th>
<th>July–October 2019</th>
<th>Term</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impeachment</td>
<td>17.30</td>
<td></td>
<td>Kurds</td>
<td>62.23</td>
<td></td>
</tr>
<tr>
<td>Houthis</td>
<td>16.23</td>
<td></td>
<td>Erdogan</td>
<td>50.73</td>
<td></td>
</tr>
<tr>
<td>Vaping</td>
<td>15.31</td>
<td></td>
<td>Blizzard</td>
<td>21.16</td>
<td></td>
</tr>
<tr>
<td>Pewds</td>
<td>13.80</td>
<td></td>
<td>Roach</td>
<td>19.95</td>
<td></td>
</tr>
<tr>
<td>Bidens</td>
<td>13.57</td>
<td></td>
<td>Charlie</td>
<td>16.16</td>
<td></td>
</tr>
<tr>
<td>Thurnberg [sic]</td>
<td>11.30</td>
<td></td>
<td>NDP</td>
<td>15.60</td>
<td></td>
</tr>
<tr>
<td>Vape</td>
<td>11.08</td>
<td></td>
<td>Alberta</td>
<td>15.58</td>
<td></td>
</tr>
<tr>
<td>Spying</td>
<td>10.32</td>
<td></td>
<td>Iraqi</td>
<td>12.93</td>
<td></td>
</tr>
<tr>
<td>Schiff</td>
<td>10.20</td>
<td></td>
<td>Roaches</td>
<td>12.82</td>
<td></td>
</tr>
<tr>
<td>Ukrainians</td>
<td>9.83</td>
<td></td>
<td>Bernier</td>
<td>11.78</td>
<td></td>
</tr>
</tbody>
</table>

*Potentially REMVE-related term.*
nier (perhaps referring to Maxime Bernier, a member of Parliament), and Alberta could reference Canadian political entities, given that Canada’s elections took place in October 2019.
CHAPTER SIX

Areas for Future Development

In this guide, we have considered some of the challenges of tracking online activity and applied a RAND-developed tool, RVE-Flock, to help address them by providing exploration capabilities and flagging emerging terms using specific metrics. Although the tool might be useful in a targeted search, additional metrics or methods for selecting viable time periods can be further developed to conduct more-informative social media analysis of emerging terms. In this chapter, we discuss the limitations and caveats to this guide and identify additional areas of development for the tool.

Limitations and Caveats

Although there are some specific insights to glean from our analysis, there are limits to the findings, like with any research. This subsection discusses some of the limitations and caveats of our analysis of /pol/ data, including term selection and methodology and equity-based considerations.

Term Selection and Methodology

The first set of limitations has to do with our initial pool and selection of terms and methodology. The 13 terms, and ultimately the ten for which we were able to conduct an in-depth analysis presented in Chapter Three, has a few limitations. The first is that the development of the tool was informed by our understanding of how different terms emerge and enter the lexicon. Had we selected different terms, we might have stressed other types of metrics. Second, the emergence patterns we discuss in Chapter Three are not exhaustive, and other patterns might exist, in which case we might have then chosen to represent terms’ progression differently.

In addition, our employment of the tool was limited to a specific period on 4chan, so we can speak only to language within those bounds. If we were to run the tool on other websites, we might draw different word associations from those discussed here. One reason could be that all online communities are different in some way—4chan and especially its /pol/ board are not representative of many other types of websites, let alone other boards on 4chan. Another reason could be that, as language continues to evolve, the associations we found here might not be representative of overall REMVE language but only on the board during a moment in time.
Relatedly, the fact that a term arises or is flagged by this tool does not mean that the term is inherently related to REMVE. Although we have done some additional research to determine when or how terms might be related to REMVE movements, it is hard for RVE-Flock to determine whether these terms are, in fact, REMVE related, which highlights the need for a human operator of the tool to make those determinations based on the tool’s outputs.

**Equity-Related Considerations**

Incomplete or unrepresentative training data can be a source of algorithmic bias if the data used to train an NLP or AI tool are more representative of some groups of people than others. The main concern is that unrepresentative data could result from an overreliance on online platforms that are used overwhelmingly by one category or demographic of REMVE actors (i.e., the stereotypical alt-right, cis-male, Christian, white supremacist) but not by others (i.e., those REMVE actors who are more diverse in demographics or underlying ideologies). If this is the case, the output from the model could also be systematically worse for unrepresented or underrepresented groups (Turner Lee, Resnick, and Bartin, 2019). For example, there could be a very small faction of members of a tribal nation on a fringe site using niche, derogatory language to describe Americans of Asian descent. Because this type of chatter happens too infrequently to be flagged as a trend or related to other inflammatory terms in some forums, RVE-Flock would not be able to form an appropriate connection with those terms, causing this lexicon to remain under the radar. Even if the chatter is happening somewhat frequently, if the terms being used by more-diverse REMVE actors differ significantly from those used by the mainstream, or if the chatter happens only on platforms less frequently used for REMVE conversations, it might be more difficult for RVE-Flock to make connections. Conversely, algorithms with too much data, or an overrepresentation, can skew the decision toward a particular result. Considering that we were investigating online platforms with significant REMVE activity, we anticipated there being greater influence by the communities most prevalent on those platforms.

This leads to the question of which platforms best represent the United States while also considering platforms most used by REMVE actors. Trying to determine the racial or ethnic makeup of posters on social media is a tricky thing to do even with AI tools and could lead to additional biases. Given the platforms available, we believe that the social costs of the trade-off between accuracy and fairness are justified; the data are what the data are. Additional research could be conducted to further incorporate and gauge equity within the tool and its sourced data sets (Osoba et al., 2019).

For the purposes of this study, we did incorporate a baseline understanding of diversity within REMVE-affiliated groups and how the language they use on social media might differ. We found this to be particularly compelling because some very prominent groups that support and partake in REMVE events—such as the Proud Boys—are not racially homogeneous (Center for Extremism, 2018). Following a review of literature on group ideologies and
public statements, we assessed that REMVE groups and individual actors who might not traditionally be considered “white” tend to identify with one of two ideological standpoints:

- “colorblind ideologies” that rely on language related to “reverse racism” against “whites” or that the “left” is promoting unfair advantages for certain minority groups
- multiracial white supremacist groups that welcome members who identify as “white” despite prominent non-European familial backgrounds. Many of these groups have been classified as “alt-lite” or “right-wing nationalists” rather than white supremacists but have serious REMVE implications. Some alt-lite groups have participated in acts of racially or ethnically motivated violent extremism, often based on anti–Black Lives Matter (BLM), anti-Muslim, anti-Semitic, or anti-immigrant platforms, while some alt-lite groups have participated in other forms of domestic terrorism (ADL, 2017b).

Further, racially and ethnically heterogenous REMVE actors were significant for this study because they might use different or more-coded language than traditional white supremacist groups use (Ma, 2021).

Although this might have been mitigated if we had explored some intersections of identity, we did not disaggregate the /pol/ data to explore any of these intersections solely because we had no insight as to a poster’s identity because anonymity in posting is core to that forum’s community. As a result, disaggregating and exploring language by these features is not feasible with these data and as they are leveraged. Future work, however, could take efforts to find some ways to disaggregate the data, such as identifying people who self-identify as members of a community or by specific topics and conversations that might draw different types of people. Further still, we cannot evaluate whether the posters and the content they submit to social media are representative of the demographics that might interest us (e.g., all Americans). For instance, access to the internet and devices needed to make social media posts are unlikely to be equivalent across all sections of the population, which is likely to skew demographics on these websites. Further still, not every type or persuasion of REMVE group is likely to have equivalent access to the internet, so terms that can be found in the real world might not be represented online.

Per our review of the relevant literature from such entities as ADL, the Southern Poverty Law Center, the Center for Strategic and International Studies (CSIS), academic journals, and news reports, we identified REMVE groups proclaiming a colorblind ideology to include the Proud Boys, Patriot Prayer, the Oath Keepers, and the Fraternal Order of Alt-Knights. Common terms used on social media that differ from those in traditional alt-right and white-supremacist groups included western chauvinism, corporate globalist left, Judeo-Christian values, race-baiters, MLK (presumably for Martin Luther King, Jr.), manosphere, venerate the housewife, west is the best, RWDS, and Pinochet did nothing wrong (see ADL, 2017b; Center for Extremism, 2018; Coaston, 2020; Gupta, 2018; Hermann, 2020; Ma, 2021; McQueen, 2021; Miller, 2018; Ngangura, 2021; Southern Poverty Law Center, undated; Trouillard, 2021; and “Understanding Multiracial Whiteness and Trump Supporters,” 2021). To ensure equity of
analysis and avoid bias, we included several of these terms (*manosphere, western chauvinism, venerate the housewife, RWDS, and Pinochet was right*) in our analysis on term emergence discussed in Chapter Three and intend to include terms from more-diverse REMVE actors in our validation of the RVE-Flock tool going forward.¹

**Additional Applications of RVE-Flock**

Although the tool is developed with an eye toward the tracking and identification of emerging terms, there are other instances or potential applications. For instance, we recognize that /pol/ board seems relatively international—and, as such, is a forum to which international users could contribute. Given the realization that foreign adversaries could exploit this fact to deepen divisions, the tool’s capabilities it could be leveraged to compare discourse on these forums with discourse on other forums that might be more domestic. Such comparisons could reveal deliberate efforts to advance REMVE terms or that more-domestic forums serve as a place wherein foreign REMVE actors connect with domestic REMVE groups (DiResta et al., 2019; Graphika, 2021). Evaluating how language, terminology, or concepts can proliferate from foreign to domestic groups could be accomplished by way of developments discussed in the next section.

Additionally, tracking the evolution of extremist language could be useful in efforts that include an attempt to counter radicalization toward people on the basis of race or ethnicity, religion, sexuality, or gender or gender identity. Given the speed at which language evolves on 4chan and other relatively unmoderated websites, staying abreast of terminology without prior familiarity is difficult. The tool could support counterradicalization efforts by being able to flag new terms of interest for nonprofit organizations to inform intervention efforts. This could take the form of addressing radicalization in youth by advising school counselors about specific terms or phrases that exist within the wider REMVE lexicon but are not yet well-known outside that community.

**Additional Developments to RVE-Flock**

From our analysis of /pol/ data, we were able to obtain intriguing information from RVE-Flock’s exploration functions. The metrics that were intended to capture developing language flagged chatter about current events in politics, international affairs, and internet culture. This is useful information to analysts because it highlights hot REMVE topics, providing an opportunity to identify terms that might be used in that specific community about current events that would not be found elsewhere (e.g., the use of the word *roach* to reference Turkish people is unlikely to remain in more-moderated communities). On the other hand, these current events seem to overshadow the gradual changes in the REMVE lexicon. Our qualita-

¹ We found that groups we had identified as multiracial white supremacist (including League of the South and the Nazi Low Riders) used language similar to that used by traditional alt-right and white supremacist groups.
tive research showed that *glowies* increased in popularity during the periods we examined but was nowhere near the terms with the greatest amount of change on either dissimilarity or usage metrics. There are various approaches we could use to improve the tool’s ability to provide meaningful results:

- Allow the user to refine results by entering in a reference set of words that are related to a topic of interest. For example, a user could enter such words as *Nazi*, *Hitler*, and *Holocaust* to find words that occupy similar topic spaces to those. This could have the effect of prioritizing relevant sets of words in which the user is interested. RVE-Flock does not leverage social media post metadata to attach results to any location. However, with the use of this capability, it might be possible to consider word associations between *revolution* and a given location, such as Washington, D.C. Additionally, being able to visualize the drift between a search term and the reference group and other words could help communicate the change in usage through time.

- Weight results by an information entropy measure based on a reference corpus (e.g., Wikipedia) to highlight “surprising” or unlikely associations in social media data that posters are suddenly using en masse. Information entropy applied in NLP can reflect the “surprise” that arises when one word follows another. For example, *heil* is likely followed by *Hitler*—so the surprise and entropy for that association is low. However, another word pairing, such as *Epstein* and *glowies* might be more surprising because that pair of words is seldomly—if ever—found on forums outside 4chan. Using information entropy to weight word pairings could help provide the user with useful associations that they might not encounter otherwise by prioritizing those with higher entropy (more surprising) over those that have lower entropy (less surprising, or expected).

- Filter terms based on growth rate metrics to find words that have specified increases in usage across time periods to counteract the overrepresentation of terms related to current events. As suggested by the relatively low growth ratio of *glowies* and relatively high growth ratio of terms associated with current events, a user might be interested in words that grow at slower rates. This suggests that allowing the user to filter based on these ratios would provide them with results that are less sensitive to offline events.

- In combination with one or more solutions above, construct a composite measure of topical dissimilarity, usage rates, and similarity to reference sets to order and prioritize the presentation of the most-relevant or -insightful associations. In the tables in Chapter Five, we provided measures of topical dissimilarity and growth separately. However, it would be possible to return a list of words in which both topical dissimilarity and growth were considered so that words that grew rapidly and were found in dissimilar topics were prioritized over words that either only grew rapidly or were found only in changing contexts.
Additional Capabilities for RVE-Flock

In addition to modifying the metrics that are used to flag emerging terms to users, we are also considering additional capabilities to improve users’ ability to track and understand language on social media platforms.

Semantic Search

RVE-Flock has the initial capability to determine a term’s meaning based on sentence context, or context clues, using word embeddings. Like LSA, a word embedding is a learned representation for text in which words that have the same meaning have a similar algorithmic representation. The difference is that word embeddings can extract meaning from surrounding context and, if trained well, can use an algorithmic expression to solve for another term’s meaning. For example, \( V( ) \) represents a function that returns the vector (or set of specified numbers) for any given word and mathematical operations as follows:

\[
V(\text{Washington, D.C.}) - V(\text{United States}) + V(\text{France}) = V(\text{Paris}).
\]

This can make semantic sense. Stated plainly, this means that, if one takes the vector for Washington, D.C., and does elementwise subtraction by the vector for the United States, the result is a vector or set of numbers that represents the abstract idea of a capital. If one then performs elementwise addition with the vector for France, the resulting set of numbers will be a combination of \textit{France} and \textit{capital}, which, in turn, should be approximately equivalent to \textit{Paris}.

Further, several methods also exist to extend word embeddings to sentences to aid in the search for similar sentences. Common ways of doing so involve either training new embedding models for sentences or using the vector values of a sentence’s words and calculating vector elementwise sums, or averages. In our case, for each sentence in a social media post, we generate a sentence-embedding vector that can then be searched by the user. The user would enter a sentence, and RVE-Flock would return a set of similar sentences based on these sentence embeddings and highlight the associated websites and dates.

Image-Hashing Search

Usually social media data leverages text data for drawing associations and meanings. However, many social media posts contain images or memes with text. If optical character recognition can be used to “hash,” or represent, an image as a series of 0s and 1s, differences between any two images could be determined, resulting in the ability to search for related images and, ultimately, flag emerging memes.

Event Tracking

In anticipation of future use of the RVE-Flock tool to predict REMVE attacks, we performed an initial assessment of existing databases on REMVE attacks in the United States, many of which were comprehensive and well vetted, such as the CSIS Transnational Threats (TNT) Project or the National Consortium for the Study of Terrorism and Responses to Terror-
ism’s Global Terrorism Database. In future work, we hope to use these databases of historical events to track the evolution of language leading up to them. ADL’s Hate, Extremism, Antisemitism, Terrorism (H.E.A.T.) Map would be a great resource for this task because it is widely cited and respected in the literature. Its sources are well vetted and include news reports, police reports, victim reports, and extremism-related sources. It also includes terrorist plots in addition to attacks, which will facilitate our tracking of REMVE language leading up to events that were thwarted, as well as events that were carried out.

2 The CSIS TNT Project data set and the Global Terrorism Database are also excellent sources, although they do not include plots that were thwarted or those for which motivations could not be verified.
RVE-Flock Concluding Thoughts

With its emergence metrics, RVE-Flock can identify when the context of a word changes, which can be leveraged to identify emerging terms or terms with new meanings. RVE-Flock characterizes a term’s emergence using two metrics: intensity of use and topical dissimilarity. We highlight the relationship between these emergence metrics, which reflect the insights and results of our qualitative examination of glown-----and glowies. Exploration of the RVE-Flock for these terms showcased associated terms that would help the querying user understand the meaning of these emerging terms. During our investigation of glowies, RVE-Flock also identified another emerging REMVE term, bantz, that was, at the time, a relatively recently coined term. In conclusion, RVE-Flock can define terms as they emerge and flag them for REMVE analysts. The RVE-Flock tool offers unique capabilities for analysts who need to stay abreast of the ever-changing REMVE landscape.
Additional Market Research

As discussed in Chapter Two, before devising the RVE-Flock tool, we conducted market research to determine what extremist language detection and event prediction products were available and where gaps existed that RVE-Flock could fill.

This appendix outlines the results of this market research. The first section provides an overview of the several tools we identified that detect troubling speech, which includes both extremist rhetoric and deliberate disinformation. The second section summarizes the tools that claim to forecast emerging threats or predict future threats. Each of these sections presents commercial tools separately from those developed by nonprofit or academic research. The final section outlines the current state of the ML and NLP market—specifically, its current focus on disinformation detection tools.

Tools to Detect Troubling Speech

We identified several ML and NLP tools that detect concerning speech, including extremist rhetoric and deliberate disinformation. This detection is used predominantly to identify and monitor individual threats, organized harassment and disinformation campaigns, and the parties responsible for such activity.¹ Some of these tools also claim to forecast emerging threats.

Commercial Offerings

ExTrac “combines real-time attack and communications data with artificial intelligence (AI) to provide actionable insights for [counterterrorism] and [countering violent extremism] policymakers and practitioners.” Specifically, one offering is Comms Tracker, which “tracks tactical and strategic shifts in messaging around conflict, hate speech and extremism” (ExTrac, undated).

¹ Not included below is the IntelCenter Database (ICD), which calls itself the “most comprehensive open source counterterrorism database available today with more than 94 million data points on terrorist activity . . . covering raw terrorist materials such as video and audio releases, photos, graphics, magazines, books, training manuals and cyber tools.” Given that IntelCenter boasts that “all core data points in the ICD are manually collected and processed” and that “IntelCenter does not utilize scrapping [sic] tools to populate the ICD,” we have not included it among the algorithm-derived tools described in this appendix (IntelCenter, undated).
Pyrra’s platform, formerly the Extremist Explorer tool developed by Human Rights First, “collects data from a variety of social media sites, forums and chatrooms, and algorithmically identifies . . . extremist language, violent threats, and harmful disinformation” (Human Rights First, 2021; Pyrra, undated).

IvySys Technologies offers social network analysis (SNA) tools that accurately model and predict the way information and behaviors propagate over [online social networks] . . . The goal is to accurately detect—early in the diffusion process—alarming social media signals representing orchestrated spreading of misinformation that will propagate widely in the future. IvySys SNA tools use model-based predictive analytic techniques based on digital signal processing and machine learning to understand and detect unfolding disinformation campaigns. (IvySys Technologies, undated)

Research and Academic Entities
The Institute for Strategic Dialogue’s BEAM platform detects and tracks disinformation, conspiracy theories, hostile state information operations and targeted harassment campaigns transparently, robustly, and using the expertise of communities affected by these threats. It mixes machine learning with social science and combines broad, continuous monitoring with focused investigation.

BEAM’s NLP algorithms are “trained . . . to detect disinformation narratives, their salience, the actors involved and any signals of coordination or inauthenticity” (Disinfo Cloud, 2021).

The European Union (EU)–sponsored Real-Time Early Detection and Alert System (RED-Alert) project, convened by a consortium of academic institutions, government agencies, and private-sector entities, promised to develop “novel natural language processing (NLP), semantic media analysis (SMA), social network analysis (SNA), Complex Event Processing (CEP) and artificial intelligence (AI) technologies” to “collect, process, visualize and store online data related to terrorist groups” and allow law enforcement clients to “take coordinated action in real-time while preserving the privacy of citizens” (EU, undated; RED-Alert, undated a).

The German government sponsored a consortium of universities and research institutes to develop the Monitoring System and Transfer Platform Radicalization (Monitoringsystem und Transferplattform Radikalisierung, or MOTRA), a comprehensive monitoring tool to analyze aggregated data to monitor high-profile societal developments. The objective is to detect changes in attitudes, which can potentially serve as an early indicator of criminal activity. Systematic monitoring enables faster identification and classification of new trends and serves as the basis for prognostic statements enabling the development of a security policy that is evidence-based, repressive and preventive.
It is scheduled for development until 2023 (Monitoringsystem und Transferplattform Radikalisierung, undated).

Tools That Predict Future Threats or Forecast Emerging Threats
We identified a few ML and NLP platforms that detect emerging threats or predict violent events. We also found numerous academic papers on event prediction methodologies.

Commercial Offerings
The Network Contagion Research Institute features a “proprietary platform that has played a critical role in the identification and forecasting of emerging threats that threaten the economic, physical and social health of civil society.” One such example is Mapping Mistrust, an “interactive tool that maps the geospatial analysis of civil-unrest activity, relating to the COVID-19 Pandemic, in the United States” (Network Contagion Research Institute, undated).

Dataminr offers “real-time AI for event and risk detection,” asserting that its platform “detects the earliest signals of high-impact events and emerging risks from within publicly available data” (Dataminr, undated).

Research and Academic Entities
The EU describes the Interconnected Next-Generation Immersive Internet-of-Things Platform of Crime and Terrorism Detection, Prediction, Investigation, and Prevention Services (CONNEXIONs) project, convened by a consortium of academic institutions, government agencies, and private-sector entities, as aiming to develop and demonstrate next-generation detection, prediction, prevention, and investigation services. These services will be based on multidimensional integration and correlation of heterogeneous multimodal data, and delivery of pertinent information to various stakeholders in an interactive manner tailored to their needs, through augmented and virtual reality environments. The CONNEXIONs solution encompasses the entire lifecycle of law enforcement operations including: pre-occurrence crime prediction and prevention; during-occurrence LEA [law enforcement agency] operations; post-occurrence investigation, and crime-scene simulation and 3D reconstruction. (CONNEXIONs, undated)
Countering Domestic Racially and Ethnically Motivated Violent Terrorism on Social Media

We identified more than a dozen academic papers outlining methodologies for terrorist event and discourse prediction. To demonstrate the range of these methodologies, we provide three examples. 

- A doctoral candidate at the Universita Degli Studi di Milano proposed the TENSOR (clusTErriNg terroriSm actiOn pRediction) framework, a “near real-time reasoning framework for early identification and prediction of potential threat situations (e.g. terrorist actions)” (Sormani, 2016, abstract). The goal of the framework is to “show how patterns of strategic terrorist behaviors, identified analyzing large longitudinal data sets, can be linked to short term activity patterns identified analyzing feeds by ‘usual’ surveillance technologies and that this fusion allows a better detection of terrorist threats” (Sormani, 2016, abstract).

- The makers of the Violence Early-Warning System (ViEWS) describe it as a “publicly available, data-driven forecasting system” based at Uppsala University and Peace Research Institute Oslo. It generates “probabilistic assessments of the likelihood that fatal political violence will occur in each country and 3,000km2 throughout Africa” (Violence Early-Warning System, undated).

- The developers of Early Model Based Event Recognition Using Surrogates (EMBERS) describe it as an “automated, 24x7 continuous system for forecasting civil unrest across 10 countries of Latin America using open source indicators such as tweets, news sources, blogs, economic indicators, and other data sources” (Ramakrishnan et al., 2014, abstract).

Today’s Market Focus on Disinformation

With the global war on terrorism and especially the rise of the Islamic State in 2014, there was a considerable focus on identifying and countering terrorist propaganda online, especially on social media. However, since the 2016 U.S. presidential election, with documented evidence of Russian disinformation operations, there has been an increasing government, commercial, and research focus on tracking disinformation rather than extremist content (Disinfo Cloud, undated).

In addition to some of the entities already mentioned, other new, prominent firms boast using ML and NLP tools to identify and counter disinformation online:

- Omelas says that it aims to expose “how authoritarians manipulate the internet,” combining “comprehensive data collection, advanced AI, and deep subject matter expertise.” Its flagship offering is Wolf Totem, “the world’s best IO [information operations?]

---

2 The others were Biswas, 2021; Colbaugh and Glass, 2012; Compton et al., 2013; Dickerson, Simari, and Subrahmanian, 2013; Ding et al., 2017; Huamani, Alva Mantari, and Roman-Gonzalez, 2020; Olabanjo et al., 2021; Pan, 2021; Python et al., 2021; Scanlon, 2014; Toure and Gangopadhyay, 2016; Uddin et al., 2020; and Xu et al., 2014.
analytics, the most advanced natural language processing, atmospherics on dozens of countries, all in a sleek, intuitive, accessible interface” (Omelas, undated).

- Yonder by Primer, formerly New Knowledge, describes itself as an “AI software company that discovers the hidden groups who control and amplify online narratives, so companies can navigate an unpredictable, ever-evolving internet with confidence.” Its Incident Management solution is a “social intelligence tool that helps brands detect and understand potentially harmful online narratives before they go viral, impacting brand trust and market valuation” (Yonder, undated).

- Graphika says that it “leverages the power of artificial intelligence to create the world’s most detailed maps of social media landscapes.” It helps clients “discover and investigate disinformation campaigns,” publishes “groundbreaking studies on disinformation dissemination,” and “monitor[s] digital threats targeting vulnerable users” (Graphika, undated).
APPENDIX B

Selecting REMVE Terms

This appendix details the literature review search terms used to generate the list of terms we considered and the full list of REMVE terms considered and provides a deep dive into REMVE terms the research team explored while assessing the emergence of these terms on social media and the internet. We identified these REMVE terms by completing a literature view on various sources that contained lists or definitions of words. To find these terms, we searched Google, Google Scholar, and SAGE Publishing and ultimately compiled this list of REMVE terms from 20 sources:

- 4chan
- ADL
- arXiv
- BuzzFeed
- CBS News
- Centre for the Analysis of the Radical Right
- Hatebase
- Incels Wiki
- JSTOR
- KYM
- Mother Jones
- Quartz
- RationalWiki
- Reddit
- SAGE Publishing
- The Atlantic
- UHURU Magazine
- Urban Dictionary
- VICE
- WordPress-hosted blogs.

We found these source types, which included academic articles, shorter news articles in such outlets as The Atlantic, and online databases, such as Hatebase. These sources included articles focused on a single word, longer glossaries, and sources with hundreds of entries.
To better understand the context of these terms, we also used KYM and Urban Dictionary, which contain user-submitted (crowdsourced) definitions of terms. It was important to include more-casual sources alongside academic journal articles and news stories because online language evolves very fast, and academic articles are significantly behind in identifying and defining terms. Some of the terms were compiled from a list that RAND researchers previously curated for a different project.

Keywords Used in the Literature Review Search
Different combinations of the following words were used to identify different lists, glossaries, articles, and dictionaries that generated the list of terms above:

- 4chan
- cuck
- dictionary
- far right
- glossary
- glowies
- hate speech
- incel
- lexicon
- memes
- misogynist
- online
- racist
- racist online language
- red pill
- REMVE language
- sexist
- terms.

REMVE Terms Considered
This list contains the 166 REMVE-related terms that we considered for this study:

- (((echo)))
- 14
- 14 words
- 1488, 14/88, 14-88, 14.88
- 6 gorillion
- 88
- alphabet
• Amerilard
• Amerimutt
• Ands Breivik
• animeright
• antiwhite movement
• ape supremacy
• bane, baneposting
• based
• Becky
• Bernard the polar bear
• billionairebros
• black don’t crack
• black pill
• black sun
• blood and soil
• blown the f--- out, BTFO
• blue pill
• Bogdanoff
• Bogpilled
• boogaloo
• Brad
• brother wars, no more brother wars
• centipede, centipedes
• Chad
• Chad-lite
• chinkoid
• Christcucks
• Congresscritter
• cuck, cucks, cuckservative, cuckservatives, libcuck, cuckbook, starcucks, cuck Schumer
• cultural enrichment
• cultural Marxist, cultural Marxists
• decile scale
• demographic replacement
• deus vult
• dindu nuffin, dindu, dindus
• Ebba Akerlund
• faggotism
• fedjacket
• feefees
• feels before reals
• feelsman, feels guys
• femcelish
• femoids
• fourteen words
• framecel
• frog Twitter
• gas the kikes – race war now, GTK-RWN
• gassed
• GEOTUS
• gigachad
• gigastacy
• glow
• glow posting
• glowies
• glowing
• glown-----
• god emperor Trump, god emperor
• good little goy
• goy
• gymcel
• hail victory!
• he will not divide us, HWNDU
• helicopter ride, helicopter rides
• high-tier Becky
• high-Tier Normie
• hikikomori
• Holohoax
• identitarians
• in minecraft
• incelish
• incels
• It’s [current year]!
• kebab removal
• kek
• kek, Kekistan
• la creatura
• le 56% face
• le happy merchant, happy merchant
• leftoids
• legacy Americans
• let’s go Brandon
• Lolbert
• looksmaxxing
• low-tier Becky
• Luca Traini
• Lügenpresse
• macacinho
• Melvin
• meme magic
• MIGA [for Make Israel great again]
• mogged
• Moishe
• monkey man
• moon man
• MSM [for mainstream media]
• muh
• muh oppression
• nabob
• neoreaction, neoreactionaries, neoreactionism, NRx, NrX, dark enlightenment
• nimble navigator
• Nirvana fantasy
• nog
• normie, normies
• not in education, employment, or training, NEET
• onions boy
• Operation Google
• Overton window
• Patrick Crusius
• Pepe the frog (or Pepe)
• political correctness
• pro white
• progressive stack
• racial holy war, RAHOWA, rahowa
• red pill
• REEEEEEE! (the number of E’s can vary)
• remove kebab
• retrowave, synthwave
• rope yourself
• Sam Hyde
• Sargon of Akkad
• Sebastiano Venier
• Shabbo Goyim
• Shadilay
• shitlord
• shitpost, shitposting
• shitskin
• smuggies
• Sneed
• snowflake (or special snowflake)
• social justice warrior, social justice warriors, SJW, SJWs
• sonnenrad (or schwarze sonne, black sun)
• spicoid
• Stacy
• Stacylite
• succubus
• suck my own cock, SMOC
• supreme white power, SWP
• Tanner
• tendies
• trash dove
• transtrender
• trucel
• ugylecel
• vibrant diversity
• virtue signaling
• we wuz kingz, we wuz kangz, we wuz kaingz
• white genocide
• white growth rates
• white liberation
• white power, white pride, WP
• white pride world wide, WPWW
• wizard
• wizard apprentice
• wristcel
• Zionist occupied government, ZOG.

This list was compiled from various online sources.

A Deep Dive into Some REMVE Terms
Cuck, Cuckservative
The term cuckold, defined as an “unknowing husband of an adulterous wife,” has a long history in the English language. Derived from Old French (cucualt), it dates back to Middle English, first appearing in the written word in an epic poem in 1250 CE (as kukeweld) and later
in the works of Shakespeare (e.g., *As You Like It*) and Chaucer (*Miller’s Tale*) (Bremmer, 2017; Online Etymology Dictionary, 2018; Sturtevant, 2016). The etymology of the terms stems from the cuckoo bird, which engages in an act of “brood parasitism,” a phenomenon whereby another bird unknowingly raise a cuckoo’s chicks after the cuckoo replaces the other bird’s unhatched eggs with its own (Clair, 2016).

After falling into obscurity, the term found new resonance in the modern era through pornography. Cuckhold porn emerged as a type of fetish pornography, a form of masochism whereby the cuckhold enjoys being humiliated (Rufus, 2019). Typically, it is a “white husband losing his white wife to a black man (and often watching or participating),” adding a racial element to what is already considered humiliating (Clair, 2016).

*Cuck*, an abbreviation of cuckhold, was added to *Urban Dictionary* in 2007 (“Cuck,” 2019). But it did not emerge as an insult until the GamerGate controversy in August 2014. The main protagonist of GamerGate accused his girlfriend, a video game developer, of sleeping with critics for positive reviews. After the accusations, he was tarred as a “beta cuck” on 4chan and Reddit (“Cuck,” 2019). The term soon gained traction in the misogynistic *manosphere* (e.g., Reddit’s /r/RedPill, 4chan’s /pol/, 8chan)—where men lament mistreatment by women and society in general as being too empowering of women at the expense of men—to “shame the men they despised” for being too weak (Sturtevant, 2016). Given the racial element described above in its original pornography context, the term quickly migrated to racist forums as well (e.g., Reddit’s /CoonTown/) (Squirrell, 2021).

According to Sturtevant, *cuck* is the “perfect” slur: “Its meaning is tantalizingly filthy, its sound abrasive, and its relative obscurity fits well in a world where language is coded and cryptic (both to establish the in-group from the out and as a method of obfuscation)” (Sturtevant, 2016). It pushes “psycho-sexual hot buttons,” particularly racial ones. According to an article in *The New Republic*, “Racism and sexism have always been connected, with one of the prime justifications for racial hierarchy being the supposed need to protect white women from black men” (Heer, 2015). It is no surprise, then, that the term was co-opted by the alt-right during its rise to mainstream prominence.

The portmanteau *cuckservative*, combining *cuck* and *conservative*, arose as an alt-right insult to “lambast mainstream Republicans,” mostly politicians, whom the term’s user “perceived to have sold out their conservative values to be humiliated by a more liberal agenda” (Clair, 2016). Given the white nationalist nature of the alt-right, it was conceived mostly in racial terms, describing “a white Christian conservative who promotes the interests of Jews and non-whites over those of whites” (Mayo, 2015).

---

1 According to medievalist researcher Paul Sturtevant, “By the 20th century, the word had faded enough into obscurity that the word was often defined in the margins of editions of Shakespeare or Chaucer” (Sturtevant, 2016).

2 Technically, this is an inaccurate use of the word. The more accurate term is *wittol*, a “compliant cuckold” (Online Etymology Dictionary, 2018).

3 Beta, or submissive, in contradistinction to an “alpha” male.
The alt-right lamented the fact that, although “white people are the predominant constituents of conservative politics . . . conservatives in power rarely promote white interests” (Hawley, 2017, p. 94). Or, in their own words, to quote prominent alt-right activist Richard Spencer,

It is the cuckold who, whether knowingly or unknowingly loses control of his future. This is an apt psychological portrait of white “conservatives,” whose only identity is comprised of vague, abstract “values,” and who are participating in the displacement of European Americans—their own children. (Weigel, 2015)

Alternatively, it is not the conservative politicians who are “cucked” but rather their supporters. Again, in their own words, quoting Gregory Hood of the alt-right online journal American Renaissance,

Grassroots Republicans who thought they had retaken the government are confused, angry, and powerless. They worked for Republicans in good faith, but get only scorn and contempt. They’ve been deceived, cheated, and exploited. In short, they’ve been “cucked,” or cuckolded. (Hood, 2015)

The first emergence of the term cuckervative as used by the alt-right is unknown. According to Twitter user @cuckservative, the “precise origins is something nobody in the alt-right has been able to pin down . . . it could have easily been MPC [My Posting Career] or TRS [The Right Stuff] where I first saw it used, but I can’t say with certainty” (Bernstein, 2015). Although an April 2015 post on MPC contains the term cuckervative to demean Republican antitax activist Grover Norquist, the term appeared a month earlier on Twitter from user stompthewaffle: “Ok mister Cuckservative, tell me more how threats to jews and israel are more important than threats to myself and my own nation” (Bernstein, 2015). In mid-July 2015, the term began to migrate from the “fringes of the ‘alt-Right’ internet to the cusp of the mainstream political conversation” when conservative writers Matt Lewis and Erick Erickson discovered and denounced the term, which brought awareness to their followers and brought it to the conservative mainstream (Bernstein, 2015; Hawley, 2017, p. 96; Heer, 2015).

---

4 The first time the term cuckervative was posted on Twitter was in April 2010: “A liberal is just a conservative that hasn’t been cucked yet . . . #cuckservative” (glopdemon, 2010). A play on the aphorism “a liberal is just a conservative that hasn’t been mugged yet,” this usage is intended as an “ironic slur against the right,” posted by a left-leaning Twitter user who frequently mocked conservatives. The tweet itself was mostly ignored ( garnering only four retweets and 16 favorites), and even glopdemon acknowledged that the alt-right probably came up with the term independently (Frost, 2015).

5 Bernstein described MPC as an “8chan-meets-far-far-right politics forum.” He described TRS as a white supremacist blog (Bernstein, 2015).

6 Perhaps not coincidentally, stompthewaffle was a proto-GamerGater, having feuded with GamerGate “villain” Anita Sarkeesian as early as 2011 (Bernstein, 2015).
But the term did not truly become ubiquitous until July 2015 when fans of Donald Trump “warred with the movement conservatives who opposed” them (Weigel, 2015). 4chan users promoted its use on Twitter to attack Republican presidential candidates who strayed from conservatism (Rappeport, 2015). Having exploded on Twitter, the term was picked up by multiple mainstream news outlets that published articles on it, including The Washington Post, BuzzFeed, and The New Republic (Bernstein, 2015; Heer, 2015; Weigel, 2015). Cuckservative peaked at the end of July 2015, when it was used more than 5,000 times in one day (Rappeport, 2015) (see Figure B.1).

Perhaps not surprisingly, many of the 4chan users who promoted use of the term on Twitter against most Republican candidates simultaneously “rained praise” upon candidate Donald Trump (Rappeport, 2015). By February 2016, when Trump was on his way to securing the Republican nomination, the abbreviated term cuck reemerged on Reddit’s /The_Donald/, now as a “catch-all insult for any man they despised” (Sturtevant, 2016). According to Squirrell’s article in The New Statesman,

Before The_Donald, misogynists, hipster racists, and other alt-right commentators existed on the peripheries. The_Donald has provided a means for all these groups to come together, and they found in “cuck” a term that encapsulated everything they hate. It’s all things to all bigots: a racially tinged term of abuse, a slur against men who trust women,

FIGURE B.1
Interest over Time in Cuckservative, 2014–2019

SOURCE: Google Trends, undated b.
a label for conservatives who aren’t conservative in the right ways, and an Islamophobic dog-whistle that propagates the narrative that Europe and America as a whole are being screwed by Muslims. With each Trump victory, cuck gained linguistic currency in The Donald. From there, it migrated to other communities on Reddit. (Squirrell, 2021)

It was soon being used to describe progressives, liberals, and others who “would not brand themselves as conservatives in the first place” (Clair, 2016).

In essence, cuck was used as a “shorthand for any perceived weakness, or rather, perceived reluctance to exploit strength” (Schwartz, 2016). (Not coincidentally, the term cuck is shorter than cuckservative, which is useful on Twitter, where economical use of characters is a priority [Clair, 2016].) Over time, other forms of the term emerged. Libcuck—added to Urban Dictionary in June 2016—describes any social progressive (“Libcuck,” 2016; “What Is a ‘Libcuck?’” 2016). Eventually, like with many insults hurled at marginalized groups, the term was coopted and used ironically by the left and libertarians as a self-description to disarm the term’s power (Bremmer, 2017). Having run its course in the ever-fleeting life span of internet memes, the term has once again fallen to the periphery (see Figure B.2).

(((echo)))

The history of (((echo))) as a symbol in the far-right discourse traces back to 2014. Its first appearance took place in an aural form. A segment titled “Merchant Minute” from an anti-
Semitic podcast, “The Daily Shoah,”7 applied an echo sound effect to the Jewish names mentioned on air (Smith and Fleishman, 2016). To represent this effect in a textual form and visually emphasize a Jewish-sounding name, the multiple (most frequently, triple) parentheses were placed around it. In this format, the (((echo))) made its first appearance in the /pol/ discussion of Emma Lazarus, a 19th-century Jewish author of a poem placed on the pedestal of the Statue of Liberty (Tuters and Hagen, 2020). The post from September 1, 2015, was based on a popular anti-Semitic “happy merchant” meme, depicting a man drawn using Jewish stereotypes, rubbing his hands with a greedy grin (ADL, undated). In this case, the face of the man was covered with a drawing of a woman’s face, and the drawing was captioned, “She was a jew by the way. Emma (((Lazarus))))” (Tuters and Hagen, 2020).

However, the widespread use of (((echo))) likely did not begin until 2016. We could not find any explanations for it gaining popularity at that particular moment in the academic literature, including through the text analysis of social media posts and in media publications. The search volume depicted by Google Trends (Figure B.3) shows that the term had been searched for even prior to that date. Although it is impossible to confirm fully, the most widespread

---

**FIGURE B.3**

Search Volume for (((Echo))) on Google Trends

![Search Volume for (((Echo))) on Google Trends](image)

**SOURCE:** Google Trends data.

---

7 The word Shoah in the name of the podcast is a Jewish word for the Holocaust, so the podcast’s title plays off “The Daily Show.” See Smolla, 2020.
assumption about the earlier use is that it signified a virtual hug or an attempt to emphasize certain parts of the message (Arviv, Hanouna, and Tsur, 2020).

In 2016, several events coincided that could potentially serve as an impetus for broader use of (((echo))), but there is no evidence that would support a finding of any causal relationships. First, to facilitate the use of the meme, a browser extension for Google Chrome was developed. The plug-in, titled the Coincidence Detector, is based on a list of Jewish and Jewish-sounding names generated and frequently updated by its users. When enacted, the plug-in automatically detected the names from this list on any internet resources the users were on and placed triple parentheses around them. Several sources have reported that many browsers no longer include the plug-in in the list of available extensions. However, the website offering the plug-in and instructions for its download is still functional (Coincidence Detector, undated).

The other event that made (((echo))) known to an audience that might not have noticed it otherwise is a New York Times article published on May 26, 2016, titled “The Nazi Tweets of ‘Trump God Emperor’” (Weisman, 2016). The author, Jonathan Weisman, described his own experience of being targeted on Twitter with the use of this meme. One response to his post on Robert Kagan’s essay on the emergence of fascism in the United States contained the text, “Hello ((Weisman)).” He asked the person who posted it to explain the use of the parentheses. The response was that it was a dog whistle to draw other users’ attention to his profile. Weisman characterized this call to action as successful. After the response to his original tweet, he was harassed by multiple Twitter users, attracting the use of (((echo))) for a long time (Weisman, 2016).

The magnitude of the attack that followed Weisman’s post speaks for profound use of (((echo))) well before it was publicly exposed. Many Twitter users who confronted him were apparently aware of the meme’s meaning and acted rather synchronously. It is likely, though, that the use of (((echo))) was limited to the communities for which it had been intended in the first place. A search on Google Trends (Google Trends, undated c) shows that the interest in the triple parenthesis as a topic (encompassing all possible variations of this expression as a search term) skyrocketed in the days after the New York Times article (Figure B.3). It might be safe to presume that someone searches for a term only once, after their first encounter. A significant increase in the number of people who wanted to find some information might signify the beginning of more-widespread use.

TRS has explained the meaning of (((echo))) (Z. Williams, 2016). Its lexicon page refers to it as a “parenthesis meme.” The meme symbolizes that “all Jewish surnames echo throughout the history” (Yglesias, 2016). So, the “damage” instigated by the Jews is still observed and felt

---

8 As of March 24, 2022, the search on the website did not yield any results. A search through the Internet Archive returns the message, “Sorry. This URL [uniform resource locator] has been excluded from the Wayback Machine.” However, the sources refer to TRS website lexicon page defining this term. For example, see Yglesias, 2016.
as time goes by, decade after decade. *Mic* provides an even more detailed explanation, citing an email by the editors of TRS:

> The inner parenthesis represent [sic] the Jews’ subversion of the home [and] destruction of the family through mass-media degeneracy. The next [parenthesis] represents the destruction of the nation through mass immigration, and the outer [parenthesis] represents international Jewry and world Zionism. (Smith and Fleishman, 2016)

The use of parentheses to publicly disclose the people of possible Jewish origin is rooted in a common far-right practice of “dog-whistling” (Saul, 2019). This technique helps disguise the true meaning of a seemingly innocuous expression to fellow community members. Some trace this to the “multivocal communication” (Åkerlund, 2021) that politicians use to simultaneously appeal to their regular voter bases while also attracting those with more-extremist views. Different audiences interpret the dog whistles differently, which helps mobilize the supporters for a particular action while being primarily unnoticed by the general public.

The triple-parenthesis meme is significant in many regards. The search engines usually disregard the nonalphanumeric characters. Therefore, it is difficult to track the use of (((echo))) at a scale for an ordinary user (Gunaratna, 2016). A researcher or journalist intentionally studying the issue can apply some advanced computational data analytics tools, but such tools might not be accessible to a typical social media user. Moreover, the triple parentheses can also be used for other, completely harmless purposes. For instance, as described above, when placed around a name, they can signify a “virtual hug” or emphasize a particular part of the message (Arviv, Hanouna, and Tsur, 2020). It once again shows that (((echo))) suited its purpose of going under the radar for most but sending the signal to those who could decode the meaning. Research shows that (((echo))) users tended to cite or mention one another, which indicates that they belong to close-knit communities (Ward, 2021).

Also, the (((echo))) is significant in that it is directly based on the practice of “othering” in which the extremist ideology is based (Tuters and Hagen, 2020). Extremist communities pay close attention to identifying those who do not belong. The researchers call this distinction of “us” versus “them” a “memetic antagonism” (Tuters and Hagen, 2020). The (((echo))) is sometimes compared with the Nazis’ use of the yellow star and other symbols to identify populations they sought to eliminate. This tactic of visibly separating the “us” from the “others” is assumed to pave the way for further segregation and annihilation (Gunaratna, 2016). For that reason, although (((echo))) is used primarily to denote Jewish-sounding names, there are examples of it being applied to single out other groups, such as “illegal Mexicans” (Tuters and Hagen, 2020). Also, the parentheses were placed not only around the names of the people with presumably Jewish origin but also on more-general concepts, such as *they*, *them*, and *their* (Tuters and Hagen, 2020). They could be used to mark abstract entities, such as *special interests* (Main, 2018), *bankers*, and *globalists*, that reference Jews in a covert manner or organizations that are allegedly controlled by Jews, such as the International Monetary Fund (Arviv, Hanouna, and Tsur, 2020).
Widespread use of the (((echo))) incentivized a countermovement. Many people placed triple parentheses around their names on social media to express solidarity. It was believed that the (((echo))) becoming ubiquitous would lose its original purpose of marking only a specific group. Such mobilization also hoped to increase awareness of extremist communities’ use of the (((echo))) for online harassment of those who were considered enemies (Gunaratna, 2016). Some suggested that this effort resulted in “effectively removing the anti-Semitic purpose of the symbol” (Miller-Idriss, 2022). Jeffrey Goldberg, editor in chief of The Atlantic, who also used the (((echo))) this way, called this an “act of (((cultural appropriation)))” (Coatney, 2021). It is sometimes compared with other terms initially used as offensive but successfully reclaimed by the very communities they intended to insult, such as queer (Gunaratna, 2016).

However, there is an alternative opinion on the adoption of the (((echo))) and its widespread use by people of both Jewish and non-Jewish origin: The parentheses represented not only a way to visibly mark a person of Jewish origin but served as a symbol of a camp. It visually “contained” the Jews in a symbolic virtual prison (Moshin, 2018). Furthermore, by exposing the Jews to the world, the extremists intended to emphasize the omnipresence of Jews, first in media but also in other aspects of social life (Hawley, 2017). From this perspective, placing the triple parentheses around one’s name, which was still primarily embraced by the people with Jewish origin, only made the extremists’ job easier. One argument that might speak in favor of the first theory that multiple people’s use of the (((echo))) effectively removed its original purpose and against the argument that the extremists benefited from it is that the extremist communities started to place inverted parentheses ))) ((( around their names to show their non-Jewish affiliations (Arviv, Hanouna, and Tsur, 2020).

We were unable to track the usage of the (((echo))) using Brandwatch because we were unable to query the parentheses that are used in its syntax. We were, however, able to query adjacent terms.
APPENDIX C

Term Emergence Exemplars

Terms That Had a Slow Buildup and Slow Emergence into the Mainstream

Western chauvinism is a term used by the Proud Boys in its mission statement to describe its overarching ideology. Experts often refer it as a replacement for white supremacy in that the Proud Boys has sought to remain in the alt-lite (see Chapter Six) and attract a more diverse membership, despite the fact that the group participated in REMVE actions and rhetoric. As shown in Figure C.1, the term western chauvinism was first put forward on social media on November 4, 2014, through a retweet of a conversation with future Proud Boys founder

FIGURE C.1
Western Chauvinism Usage on Social Media

NOTE: Throughout this report, by mentions, we mean tweets, retweets, and quoting tweets for Twitter; for Reddit, we mean comments and posts.
Gavin McInnes and former governor Jesse Ventura (Ventura, 2014; McInnes, 2014). The term was used at low levels in REMVE communities over an extended period, with relatively low but increasing rates of usage from July to December 2014. Its first mention outside of Twitter was on 4chan just over a year later, on November 12, 2015 (Jeff Mangum Died for Your Sins, 2015). We consider the term to have gone through two periods of emergence, at which point it was being used by REMVE actors on multiple platforms with increasing rates of usage, from both April 2016 to June 2016, and again from April 2017 to June 2017. At the end of the second period of emergence, the term fully emerged into the mainstream in July 2017, at which point non-REMVE actors began discussing the term on social media. The largest spikes in its usage did not come until 2020 after Trump mentioned the Proud Boys in a presidential debate.

Terms That Had a Slow Buildup and Rapid Emergence into the Mainstream

GEOTUS, as seen in Figure C.2, first showed up on social media on January 9, 2017, in a tweet (Dorfman, 2017). We saw patterns of increasing rates of usage beginning in February 2017, at which point the term was also being used in the same context on all three platforms by REMVE-associated actors. This period of increased usage at an increasing rate contin-

**FIGURE C.2**

**GEOTUS Usage on Social Media**


NOTE: Throughout this report, mentions in the context of 4chan means threads and responses to them.
ued until February 2018, but we consider the term to have emerged into the mainstream by October 2017, when the term saw widespread usage by non-REMVE-related actors. This coincided with a spike in usage around President Trump’s comments about National Football League (NFL) players kneeling in support of BLM (Crokin, 2017). There was another major spike in usage on November 19, 2017, due to a retweet of user @LizCrokin calling former president Trump a superhero (Aguanno, 2017). Later, large spikes in usage on social media occurred in 2018, 2019, and 2020 around QAnon information drops, the appointment of federal judges, and the presidential election.

Shown in Figure C.3, RWDS is commonly featured in messages calling for enemies of the alt-right to be killed. There was some white noise surrounding this term because RWDS was also used in a non-REMVE context to discuss music. The first REMVE-connected mention was on April 1, 2012 (CAN COOKZ? 2012), although there were likely previous mentions that fell outside of our data set. We saw increased rates of usage at an increasing rate occurring from July to September 2014, and then again from July to December 2016. We consider the term to be emerging during these time frames because REMVE actors were using it on multiple platforms during these periods. There were some initial larger spikes in usage on 4chan in September 2016 and May 2017 related to Trump’s presidential campaign and a story about African migrants crossing into Europe that catalyzed direct calls for violence. We do not consider the term to have emerged into the mainstream until June 28, 2017. However, there were

**FIGURE C.3**

**RWDS Usage on Social Media**

![Graph showing RWDS usage on social media](image)
concerns about actual RWDS in Venezuela on Twitter, including from users both inside and outside of the alt-right community.

Other REMVE terms were characterized by a slow buildup of use by REMVE-affiliated communities over a time frame of greater than six months, followed by a rapid emergence into the mainstream. This rapid emergence into the mainstream was characterized by a significant spike in usage—usually related to a specific event—which preceded full emergence by less than six months. *Pinochet did nothing wrong* is a phrase that followed this emergence pattern.

As seen in Figure C.4, *Pinochet did nothing wrong* first appeared on social media on September 11, 2012. This term is a reference to Augusto Pinochet, a violent Chilean dictator known for arresting and killing hundreds of thousands of leftists who opposed his rule (KOIN 6 news staff, 2018). The implication is that this would be an acceptable strategy for a leader to take now. From January 2016 to March 2016 and again from October 2016 until June 2017, the term showed rapid increase in usage, as well as proliferation to multiple social media platforms by REMVE actors using it in the same context—we consider the term to be emerging at both points. The first major spike in usage did not occur until August 4, 2018, in response to leaders at a far-right rally in Portland wearing shirts with the phrase on them. This spike in usage involved messages on Twitter from people both inside and outside of the

**FIGURE C.4**

*Pinochet Did Nothing Wrong* Usage on Social Media

![Chart showing the usage of the term “Pinochet did nothing wrong” on social media](chart.png)

- **Date Range:** 2013–2022
- **Platforms:** 4chan, Reddit, Twitter
- **Key Events:**
  - The term has emerged into the mainstream
  - The term is emerging
  - Twitter users discuss violent behavior and shirts bearing the phrase (1,020 mentions on Twitter)
  - Twitter users discuss and condemn use of the term (1,793 mentions on Twitter)

**SOURCE:** Brandwatch data, accessed April 13, 2022.
REMVE community, with heated debate for and against the shirts, and we consider the term to have fully emerged into the mainstream at that point.

Terms That Had Rapid Buildup and Rapid Emergence into the Mainstream

A small number of REMVE terms analyzed displayed a rapid buildup of use by REMVE-affiliated communities (over a time frame of less than six months) followed by a rapid emergence into the mainstream. This rapid emergence into the mainstream was characterized by a significant spike in usage, which preceded full emergence by less than six months. We wuz kangz exemplifies this emergence pattern.

As seen in Figure C.5, we wuz kangz is used to mock people who present evidence that Black societies that predate such institutions as slavery had significant achievements and contributions to the world. The term took off quickly and is an important example of rapid emergence into the mainstream. According to Brandwatch, the first usage in our data set was on 4chan on October 18, 2015, and it experienced a rapid increase in usage from January through June 2016 and almost immediately proliferated to multiple social media platforms by REMVE-associated actors using it in the same context. There was a large spike in usage on February 26, 2016, in a thread about BLM that included racist and sexist language and

FIGURE C.5
We Wuz Kangz Usage on Social Media

another on April 4, 2016, in a racist 4chan thread with 720 mentions, including threats of violence. Although we found a one-off comment on 4chan in December 2015, we did not see other widespread evidence of the term being used outside of REMVE contexts until June 2016, at which time conversations began emerging about how the term was racist. We consider the term to have emerged into the mainstream at this point. In a Graphika report, suspected Russian state actors targeted right-wing extremists in 2020 using *we wuz kangz* to play on long-standing racist tropes to further divide Americans. According to the report, these influence campaigns are “still active, and build on previous foreign influence efforts likely conducted by the same actors that Graphika exposed ahead of the 2020 U.S. presidential election” (Graphika, 2021), further highlighting the importance of keeping up with a rapidly changing extremist lexicon.

*God emperor Trump*, as seen in Figure C.6, was first mentioned in our data set on August 7, 2015, in a tweet from a Trump supporter about the “ultimate battle for the earth” between Trump and Russian president Vladimir Putin (irrebæ, 2015). This term is used to describe Trump as an all-powerful dictator. There were two modest spikes in usage in December 2015 by REMVE-associated social media users. From November 2015 to January 2016 were patterns of increasing rates of usage, along with the term being used on multiple platforms within REMVE communities, that indicate that the term was emerging. We assess the term as having fully emerged by January 2016, when it was first used outside of REMVE com-

**FIGURE C.6**

*God Emperor Trump* Usage on Social Media

![Graph showing the usage of *God Emperor Trump* on social media platforms from 2015 to 2022, with peaks in 2016 and 2021.](SOURCE: Brandwatch data, accessed May 9, 2022.)
munities. By September 3, 2016, Twitter usage of the term surpassed 4Chan usage as users retweeted a comment about Facebook banning the “God Emperor Trump” page. There were other large spikes in the following years, both by those making fun of those using the term and by those in support of former president Trump.

*Venerate the housewife*, as seen in Figure C.7, used primarily by alt-lite actors, such as the Proud Boys, to discuss the “natural” place of the woman as being in the house and traditional gender roles. The term never crossed 100 mentions, but there were rapid spikes within the first six months that exemplify how rapid emergence occurs. The first mention in our data set was on June 30, 2016, on Twitter (NinjaDragonFart, 2016). Within a month, there was a spike to 96 retweets of a comment that McInnes made on Twitter (e.g., McInnes, 2016, and Grun, 2016). By April 2017, the term was being used outside of the REMVE community to mock the Proud Boys. The term faded in and out of popularity as time progressed, with a period of increased usage at an increasing rate between October and December 2020 after President Trump’s mention of the Proud Boys during a presidential debate.

**FIGURE C.7**

*Venerate the Housewife Usage on Social Media*

APPENDIX D

How Racially and Ethnically Motivated Violent Extremist Actors Use the Internet Before Attacks

Ample research has shown how the internet and social media have been used as tools of recruitment and radicalization for violent extremist organizations, including REMVE actors. In this project, we did not seek to recreate or analyze that research. However, to verify the usefulness of the RVE-Flock tool for future use in event prediction, the project team did conduct a literature review on the nature of online indicators and the evolution of language used by perpetrators of imminent violent events. Although the discussion in this appendix is not an exhaustive review of this literature, it does provide a brief overview of recent findings that lend strong credibility to the importance of monitoring the use of language on social media in the lead-up to REMVE events.

Academic and policy-oriented research provides strong evidence that violent extremists, including REMVE actors in the United States, utilize online social media platforms in ways that either directly or indirectly indicate their desire or intention to carry out violent attacks.

According to the Federal Bureau of Investigation (FBI), online indicators of homegrown violent extremism for imminent threats include preparing and disseminating martyrdom videos or statements (while near-term threats might be indicated by attempting to mobilize others to violence); communicating intent to engage in a violent activity with justification; and producing extremist content (Federal Bureau of Investigation, National Counterterrorism Center, and U.S. Department of Homeland Security, 2019). Although this list was specific to Islamist-based terrorism, further research indicates that right-wing extremists are likelier to disclose their violent intentions online than extremists of other ideologies (Bouhana et al., 2018). Other recent peer-reviewed research has found top online indicators of right-wing extremist violence to include expressing racially or ethnically motivated grievances, engaging other in ideological discussions, demonstrating ideological outbursts, and advocating and encouraging violence (Scrivens et al., 2021). Despite the visibility of recent large-scale REMVE events by small groups, studies have found that around 50 percent of REMVE attacks are carried out by lone actors (ADL, 2017a). The prevalence of lone REMVE actors has prompted a diversity of research on preattack behaviors, in which it has been revealed that these lone actors largely disregard operational security measures; they often share, or at least
hint at, their violent plans and “leak” their convictions about or involvement in suspicious, violent activity months or even years before attacks (such as stating they want to “kill someone”) (Schuurman et al., 2018). Additionally, threat actors’ linguistic patterns can change, becoming less complex in the month leading up to the attack.
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADL</td>
<td>Anti-Defamation League</td>
</tr>
<tr>
<td>AI</td>
<td>artificial intelligence</td>
</tr>
<tr>
<td>BLM</td>
<td>Black Lives Matter</td>
</tr>
<tr>
<td>CONNEXIONs</td>
<td>Interconnected Next-Generation Immersive Internet-of-Things Platform of Crime and Terrorism Detection, Prediction, Investigation, and Prevention Services</td>
</tr>
<tr>
<td>COVID-19</td>
<td>coronavirus disease 2019</td>
</tr>
<tr>
<td>CSIS</td>
<td>Center for Strategic and International Studies</td>
</tr>
<tr>
<td>DDG</td>
<td>DuckDuckGo</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FVD</td>
<td>Forum voor Democratie, or Forum for Democracy</td>
</tr>
<tr>
<td>GEOTUS</td>
<td>god emperor of the United States</td>
</tr>
<tr>
<td>ICD</td>
<td>IntelCenter Database</td>
</tr>
<tr>
<td>KYM</td>
<td>Know Your Meme</td>
</tr>
<tr>
<td>LSA</td>
<td>latent semantic analysis</td>
</tr>
<tr>
<td>ML</td>
<td>machine learning</td>
</tr>
<tr>
<td>MPC</td>
<td>My Posting Career</td>
</tr>
<tr>
<td>NA</td>
<td>not applicable because there were too few observations to calculate the metric</td>
</tr>
<tr>
<td>NFL</td>
<td>National Football League</td>
</tr>
<tr>
<td>NLP</td>
<td>natural language processing</td>
</tr>
<tr>
<td>RED-Alert</td>
<td>Real-Time Early Detection and Alert System</td>
</tr>
<tr>
<td>REMVE</td>
<td>racially or ethnically motivated violent extremist</td>
</tr>
<tr>
<td>RVE-Flock</td>
<td>Racist and Violent Extremist Flock</td>
</tr>
<tr>
<td>RWDS</td>
<td>right-wing death squad</td>
</tr>
<tr>
<td>SNA</td>
<td>social network analysis</td>
</tr>
<tr>
<td>TRS</td>
<td>The Right Stuff</td>
</tr>
</tbody>
</table>
References

ADL—See Anti-Defamation League.

Aguanno, Seb (@Seb_Aguanno), “Some call him POTUS, others call him GEOTUS, I call him BATMAN! Trump is a real life super hero! 🦇 No pedo is safe in Gotham anymore! 🕷️ https://pic.twitter.com/7127PA27iC,” Twitter, November 18, 2017, 7:43 p.m. As of September 23, 2022: https://twitter.com/Seb_Aguanno/status/932046587647418368


———, “From Alt Right to Alt Lite: Naming the Hate,” backgrounder, July 12, 2017b. As of February 18, 2022: https://www.adl.org/resources/backgrounders/from-alt-right-to-alt-lite-naming-the-hate


Countering Domestic Racially and Ethnically Motivated Violent Terrorism on Social Media


CAN COOKZ? (@_Coox), “Oga please, I’ve got my gun on meiii, pump it up real niggas go pump it up! RWDS Baby!!” Twitter, April 1, 2012, 8:12 a.m. As of September 23, 2022: https://twitter.com/_Coox/status/186425683881377793


Coincidence Detector, homepage, undated. As of March 24, 2022: https://coincidencedetector.com/


Crokin, Liz (@LizCrokin), “Some call him POTUS, others call him GEOTUS, I call him BATMAN! Trump is a real life super hero! 🦇 No pedo is safe in Gotham anymore! 🌋” Twitter, November 19, 2017, 7:43 p.m.


Dataminr, homepage, undated. As of September 9, 2022: https://www.dataminr.com


DiResta, Renee, Kris Shaffer, Becky Ruppel, David Sullivan, Robert Matney, Ryan Fox, Jonathan Albright, and Ben Johnson, The Tactics and Tropes of the Internet Research Agency, prepared for the U.S. Senate Select Committee on Intelligence, October 2019. As of September 12, 2022: https://digitalcommons.unl.edu/senatedocs/2/


———, “Stage 1 Technology Assessment Summary: BEAM by ISD and CASM,” webpage, 2021.

Dorotman, Anna (@doorsixteen), “Also appreciated: @realDonaldCtxxt. Contextual fact-checking of GEOTUS's tweets,” Twitter, January 9, 2017, 11:21 a.m. As of September 23, 2022: https://twitter.com/doorsixteen/status/818492819870363650


EU—See European Union.


ExTrac, homepage, undated. As of September 9, 2022: https://extrac.io


glopdemon (@glopdemon), “A liberal is just a conservative that [sic] hasn’t been cucked yet. Looking for paleo-con man missile to do this. #cuckservative #tcot #tcuck,” Twitter, April 1, 2010, 5:48 p.m. As of September 24, 2022: https://twitter.com/glopdemon/status/11446292205


Graphika, homepage, undated. As of September 9, 2022: https://graphika.com

———, *Posing as Patriots*, June 2021. As of September 12, 2022: https://graphika.com/reports/posing-as-patriots

Grun, Josh (@JoshGrun10), “RT @Gavin_McInnes #IamaFeminist because I venerate the housewife https://t.co/84njrQdJ5X,” Twitter, July 28, 2016, 1:15 p.m.


Hatebase, homepage, undated. As of September 9, 2022: https://hatebase.org


IntelCenter, “IntelCenter Database (ICD),” webpage, undated. As of September 12, 2022: https://www.intelcenter.com/icd


irrebae (@irrebae), “@KyleKulinski I cant wait to see the ultimate battle between God Emperor Trump and Putin for the fate of Earth” (animated image omitted), Twitter, August 6, 2015, 8:28 p.m. As of September 23, 2022: https://twitter.com/irrebae/status/629448888059981824


Jeff Mangum Died for Your Sins, “Even putting my Western chauvinism aside, the King James Bible is a much more impressive piece of literature than any English translation of the Quran. The Quran is too disjointed, and too obviously plagiarized from the Judeo-Christian tradition,” 4chan, November 12, 2015.


“Manosphere,” Know Your Meme, updated c. 2015. As of September 12, 2022: https://knowyourmeme.com/memes/subcultures/manosphere


McInnes, Gavin (@Gavin_McInnes), “@GovJVentura, “I’m talking to @Gavin_McInnes about Dan Snyder, his push for ‘western chauvinism’ towards Islam & midterm elections. Any questions for him?” Twitter, November 4, 2014, 7:45 p.m.

———, “#IamaFeminist because I venerate the housewife https://t.co/84njrQdJ5X,” Twitter, July 28, 2016.


Monitoringsystem und Transferplattform Radikalisierung, homepage, undated. As of September 10, 2022: https://www.motra.info


MOTRA—See Monitoringsystem und Transferplattform Radikalisierung.


References


NinjaDragonFart (@NinjaDragonFart), “@badvibratjons @itsbetterifudo @FashNational i [sic] do. I venerate the housewife. I want better for you than loneliness and crippling depression,” Twitter, June 30, 2016, 7:12 p.m.


Omelas, homepage, undated. As of September 10, 2022: https://www.omelas.io/


———, “Similar Projects,” webpage, undated. As of September 9, 2022: https://redalertproject.eu/about_us/similar-projects

RED-Alert—See Real-Time Early Detection and Alert System.


———, “Hatebase Projects,” webpage, undated b. As of September 9, 2022: https://thesentinelproject.org/project/hatebase/


Ventura, Jesse (@GovJVentura), “I’m talking to @Gavin_McInnes about Dan Snyder, his push for ‘western chauvinism’ towards Islam & midterm elections. Any questions for him?” Twitter, November 4, 2014, 7:45 p.m. As of September 23, 2022: https://twitter.com/GovVentura/status/529434232511135744

Violence Early-Warning System, Department of Peace and Conflict Research, Uppsala University, homepage, undated. As of September 10, 2022: https://www.pcr.uu.se/research/views


Racially and ethnically motivated violent extremism is becoming an increasingly common occurrence in the United States. Racially and ethnically motivated violent extremist (REMVE)–related terrorism has consequences beyond loss of life: It undermines the sense of safety that targeted groups feel in their country and unravels the social fabrics of trust that are necessary for society to function. Further still, REMVE attacks can motivate other like-minded attackers to follow up with their own attacks, as was apparently the case with the May 14, 2022, mass shooting in Buffalo, New York, by a self-avowed, internet-radicalized white supremacist whose manifesto drew heavily from the March 15, 2019, Christchurch mosque mass shootings, which also inspired a mass shooting in El Paso, Texas.

Two key challenges for those who observe online spaces in which radicalization occurs are the sheer volume of data and the idiosyncrasies of online communities. Website-specific language and memes are difficult to track and parse; even if emerging terms are detected, defining them can be difficult.

RAND Corporation researchers developed the Racist and Violent Extremist Flock (RVE-Flock) tool to explore and analyze textual content on REMVE-affiliated social media. The user can identify emerging terms used in REMVE communities and trends on internet platforms. In this guide, the authors characterize term proliferation in online communities by exploring various REMVE terms and demonstrate the tool’s functionality. To conclude, the authors identify additional applications of this work and potential refinements of the tool.