How Discourse Strengthens Online White Supremacy: Strategic Framing on Stormfront

By

Maria-Henriete Pozsar

Submitted to Central European University Department of Political Science

In partial fulfillment of the requirements for the degree of Master of Arts

Supervisor: Levente Littvay

Vienna, Austria

2021

"Becoming a member of the alt-right, it turned out, wasn't exactly a process of logical persuasion. It was more like a gradual shift in your mental vocabulary. (...) She'd spent months convincing her friends, and herself, that the movement was more innocuous than it seemed, that their fears were misplaced. But what if the ironic racism really was just racism? What if white separatism and white nationalism and racial realism were all epithets for the same old-fashioned violent impulse?"

(Marantz, 2019)

Abstract

Far right political violence has undoubtably been getting bolder in the United States over the past decade. One expert hypothesis is that online white supremacist spaces might contribute to this phenomenon because they may be conducive to processes of radicalization. This project aims to provide a building block for the fields of study dealing with white supremacist online radicalization by examining ideological framing processes present in an original dataset containing seventeen years of speech from the oldest online platform dedicated to white supremacy, Stormfront. In doing so, this project also aims to present an example of a mixed-methods study approach to the topic, including Natural Language Processing and Social Network Analysis, applied on the basis of social science theories from social movements and communication studies, such as framing theory, the concept of political dog-whistles and rhetorical distance. The results of the analysis come in opposition to previous findings obtained mainly from qualitative study designs which support the idea that strategic re-framing of racist and white supremacist ideas is prevalent, or at least, on the rise, in online white supremacist communities. The current findings discover that the opposite frame, dealing with racist speech and ideas in an overt, direct way, may have been underestimated, and continues to occupy an important position in the Stormfront network.

Acknowledgements

As a political scientist without a strong computational methods background, acquiring the skills needed for the completion of this study design was more than a little bumpy. Somewhere in the process however, faced with the witching potential of computational social science methods, I became somewhat of a zealous, annoying little convert.

I want to thank Gabor Simonovitz for putting up with my excessive enthusiasm when I was making the first, ridiculously unfeasible pitches.

Michael Zeller has been the best advocate for this project, to myself, at a time when I was losing confidence in it.

A debt of gratitude is owed to the faculty and teaching assistants of the Network Science department at CEU, without whose teaching and guidance I would not have even dreamed of designing this project. Particularly, I want to thank Júlia Koltai, Marton Posfai, Gerardo Iñiguez and Gabriela Juncosa.

In my first panicky attempts at working with Python, Arthur Start has calmly and methodically guided me when I was stuck.

My good friend and linguist Steli has done a vertiginously fast proofreading of the paper.

Sorana Constantinescu has always pushed me from behind and there are many things I would not have done, including getting a Master's at CEU, without her brilliant support.

Alex has been my chosen family all throughout the Covid-19 pandemic and before. She was instrumental in making this thesis possible.

Not least, my supervisor, Levente Littvay, has treated this collaboration with the utmost patience, curiosity, and generosity.

Thank you!

Table of contents

Abstractii
Acknowledgements
Table of contentsiv
List of Figures, Tables or Illustrationsvi
1. Introduction1
1.1. Motivation2
1.2. Framings of white supremacy on Stormfront
2. Theoretical Framework
2.1. The Difficulty of Labeling White Supremacy6
2.2. Rhetorical Distance: A Mechanism of Reframing9
2.3. Competition of White Supremacy Frames11
2.3.1. Ideologically covert frames: A Theoretical Contribution and a Choice of
Operationalization
2.3.2. Networked Framing
2.4. Prevalent Methodologies in Studying Frames of White Supremacy16
2.5. Social Media-Appropriate Methods of Studying Frames of White Supremacy: Content and
Networks
2.6. Research Questions and Aims of Study22
3. Methods
3.1. Dataset

3.2. Protocol	25
4. Analysis	31
4.1. Detection of Frames of White Supremacy	31
4.1. 1. Quantitative exploration of the Stormfront corpus	31
4.1.2. Construction of Dictionaries	35
4.2. Networked Frames of White Supremacy	
4.3. Checks	48
5. Discussion and Conclusion	51
5.1. Discussion	51
5.2. Limitations and Further Development	53
5.4. Conclusions	55
List of References	58
Appendices	64
Appendix 1. Replication code for data scraping and cleaning (Python)	64
Appendix 2. Replication code for NLP (R)	65
Appendix 3. Replication code for Social Network Analysis (Python)	66

List of Figures, Tables or Illustrations

Figure 1 Screenshots from Stormfront taken in December 20204
Figure 2 Representation of workflow of data extraction, pre-processing and analysis
Figure 3 Period of activity on the forum and the most active users
Figure 4 Top 50 (chart) and top 100 (word cloud) stemmed unigrams in the entire Stormfront
corpus
Figure 5 Distribution of posts containing the terms "hitler", "jew", "nation" and "white" over
time. The y axis shows the number of word occurrences per post. Each dot is a post
Figure 6 Evaluating LDA topic models
Figure 7 Term loadings for topics detected through LDA 10 topics modeling35
Figure 8 Usage of overt and covert frames of white supremacy over time on Stormfront40
Figure 9 Stormfront users with highest loadings in the Overt and Covert white supremacy
dictionaries41
Figure 10 Full network visualizations for each dictionary. The size of the node = indegree
centrality45
Figure 11 Yearly cross-sections in the overt white supremacy frame. Nodes using this frame are
in red46
Figure 12 Yearly cross-sections in the covert white supremacy frame. Nodes using this frame are
in green46
Figure 13 Regression on the full network
Figure 14 Regression on the network cross sections, y = covert frame
Figure 15 Regression on the network cross sections, y = overt frame
Figure 16 Yearly cross-sectional correlations between three node-level centrality measures and
the two dictionaries

1. Introduction

News cycles and political life in the United States have been increasingly featuring acts of political violence committed by far right actors. In January 2021, as the country was waiting for the Congress to uneventfully formalize Joe Biden's electoral victory, a crowd of protesters broke into the Capitol building, disrupting the congressional session. The infamous riot, dubbed "the storming of Capitol Hill", resulted in five deaths, the second impeachment of President Donald Trump and criminal charges brought against identified participants (BBC News, 2021; Fazio, 2021; Healy, 2021). Notably, the occupation was reportedly organized through online coordination between several groups of far right, white supremacist and conspiracist inclinations, on Parler, Gab, Telegram and other platforms alternative to mainstream social media (Frenkel, 2021; Hoback, 2021).

A body of interdisciplinary social science research has been growing since the 2000s stemming from the recognition that online dissemination of information, opinions and social connections play an important role in social movements (Tufekci, 2014). In the case of far right social movements, online spaces of connection have been identified as crucial, due to a tendency towards decentralization and the cultivation of "lone wolf" actors (Blee & Latif, 2021; Peita, 2021). This project contributes to the field of study investigating the online behavior of far right movements by proposing a computational mixed-methods workflow for accessing and analyzing online data to answer questions about the online activities of far right actors. The application of the methods proposed in this project targets the specific frames of white supremacy used by people spending time on the oldest far right online platform, *Stormfront*. The aim of this application is to contribute a building block, namely that of characterizing the evolution of the frames of white supremacy, to a growing and ambitious body of research asking the very difficult question of

whether lines of influence can be detected between online radicalization and violent political behavior.

1.1. Motivation

The storming of the Capitol continued a trend of far right attacks, which have been steadily increasing in the past decade (CSIS, 2020). During Donald Trump's presidency, white supremacy has often been in the limelight. There have been more far right terrorist attacks worldwide – Christchurch, El Paso, Poway, Hanau, to name only the most recent and resonant ones (Auger, 2020), hate crimes targeting minorities and immigrants (Cai & Landon, 2019), and more white power rallies than ever before. Some of these rallies, although they were regularly organized (Marantz, 2019), barely registered in the mainstream public consciousness before the infamous "Unite the Right" rally in Charlottesville, Virginia.

The darker underbelly of these (anti)social manifestations is thought to lie with online communication, radicalization and, sometimes, even organization of groups across many platforms. These groups can only loosely be brought together under an umbrella term such as "right-wing extremism" or "white supremacy." Far right attackers themselves have been raising the public profile of some of their online communities by posting political manifestos as companion pieces to their attacks or live streaming their violence on platforms such as *Gab*, the now defunct *8chan* and even on *Facebook* and *Reddit*.

With growing public discontent towards social media companies for serving as platforms to hate speech, and both political (Browning, 2020) and commercial (Clayton, 2020) pressure, many of them have updated their community guidelines and have become much swifter in censoring said hate speech and even banning accounts of famous sources of extremist right-wing speech (Rogers, 2020), such as Milo Yiannopoulos and Alex Jones. As a result, these celebrities, who had the role of opinion hubs in their networks, have moved on to alternative platforms which differentiate themselves by not setting any limits to free speech and maintaining very loose or absent community guidelines. Among these, Gab and Parler (Twitter alternatives), MeWe (a Facebook alternative) and Telegram, which accommodates both public and private groups, can be named (Isaac & Browning, 2020; Rogers, 2020).

Findings from studies of political opinion on mainstream social media point towards a phenomenon of polarization generating echo chambers (Balsamo et al., 2019) in which opinion exchange is limited to a circle of like-minded people (Key, 1966). If that is true for platforms hosting many politically diverse groups, what could the case be for politically homogeneous platforms? At least on Twitter, it has been shown that the polarization of opinions and the interaction of users evolve independently in time, such that people who disagree tend to be more polarized in the beginning when reacting to an event, but interact more on the same discordant topic as time goes by (Balsamo et al., 2019). If there is no dissenting cluster in a social network, the possibility of interaction with dissenting opinions is completely taken away. What consequences will this polarization between entire platforms, not just clusters within a platform, have?

1.2. Framings of white supremacy on Stormfront

In the case of white supremacist groups, there has already been, for decades, a prime example of a politically segregated online platform: *Stormfront.* The platform was founded in 1995 by Don Black, a former Ku Klux Klan leader, who continues to be the owner of the platform and sporadically participates in the forum discussions. Stormfront is recorded as the first extremist social platform on the Internet (Peita, 2021) and continues to be the most popular white supremacist online community to date (Southern Poverty Law Center, 2020a). It functions as a forum for discussion, as well as a blogging platform. Stormfront has been not just a popular social platform for white supremacists, but one of the most popular websites on the internet, ranking in the top 1% of existing websites in 2005 (Kim, 2005). After a peak in popularity in 2008, new registrations to the platform have fallen (Hankes & Zhang, n.d.); however, it still has over 363,000 registered users (Stormfront Home Page, 2021), compared to over 155,000 at the end of 2008 (WayBackMachine Internet Archive, 2008). Due to its longevity and popularity, it might be the

best candidate for looking at the effects of the complete segregation of a network from dissimilar opinions. It would be reasonable to assume that the most likely result of such segregation is a progressive advancement of the ideological extremism of discussions. However, just from perusing the discussion threads on Stormfront, this is not evident. Based on personal observations of posts on Stormfront and similar other online communities with ideologies compatible with white supremacy, as well as a close reading of the manifesto published by the perpetrator of the Christchurch terrorist attack of May 2019 (The Independent, 2020), I noticed certain ways of speaking about white supremacist ideas which were effectively erasing and hiding the "regular" or "common" understanding of those ideas. For instance, rather than centering race at the heart of their ideology, some users replace it with whiteness, such that "racist" and "white supremacist" become "pro-white," and any consideration about racial prejudice becomes labeled as "anti-white". *Figure 1* contains several screenshots taken from Stormfront which exemplify this type of linguistic reframing of recognizable white supremacist words and imagery into something charged with an ambiguous positive valence.

"Anti-white" creates a new category

When the word "anti-white" has been put into the minds and onto the tongues of every White person in this world by constant, determined, unceasing repetition, A NEW CATEGORY WILL EXIST:

ACTIONS AND PERSONS THAT ARE INJURIOUS TO WHITES.

That category *does not exist today*. Think how *absolutely necessary* that category is.



Since when has USING the swastika been successful?

I am White. I am pro-White. I am a nationalist. I fully support my country (the USA) and I fully believe in the superiority of the USA over any other country. I believe (no, I KNOW) that the **culture**, scientific achievements and greatness of the USA are due to the efforts of White people.

Figure 1 Screenshots from Stormfront taken in December 2020

The importance of this tendency to reframe white supremacism and transform it from a readily recognizable disgraced extremist ideology to an ambiguous ideology with positive connotations is difficult to miss. The most obvious advantage is that reframing white supremacist ideas in this fashion makes them socially acceptable. Strategies of reframing discussions in online communities may offer some protection against being pigeonholed as a white supremacist group by outsiders. Moreover, this reframing may be conducive to recruitment. "Pro-white" is much more appealing to those who are interested in white identitarian arguments, but are put off by outright racism. Reframing in this fashion, along with other phenomena such as the adoption of an aesthetic and propelling conspiracy theories may contribute to what some researchers see as the "mainstreaming" of far right extremism (Miller-Idriss, 2020).

This research project examines discourse on Stormfront in order to identify whether differences in framing white supremacist ideas exist and what their evolution is in time.

2. Theoretical Framework

2.1. The Difficulty of Labeling White Supremacy

Scholars of white supremacist groups, movements and terrorism disagree on the terminology used to refer to the ideological constructs equivalent to white supremacy. How to name and how to define these groups is an intensely debated issue (Blee, 2017, p. 9), which generates confusion in the field. While for some, the claims that the groups themselves make about following or not following a white supremacist ideology count, for others, these claims are mere deceptive rhetoric. Definitional clarity about what makes a white supremacist ideology is paramount when studying its discourses.

Some researchers do not make a special distinction for white supremacist groups and, instead, prefer to refer to them with umbrella terms such as "the far right", which captures everything from conservative political parties to the right of mainstream parliamentary politics to anti-abortion acts of terror, to the Ku Klux Klan. This approach simply highlights the shared conservative ideology (Blee, 2017) and ignores differences on other axes such as religious orientation, attitudes towards the environment, attitudes towards race, or what these groups claim to believe.

One exclusionary step in the specialized terminology used is "right-wing extremism." The *extremist* character is defined as anything that presents procedurally and normatively as antidemocratic (Klandermans and Mayer 2006 as cited by Blee 2017). "Extremism," however, is by no means an uncontested term, and some prefer to avoid it because of its own normative load and dependency on the politics of "who is doing the labeling" (McNeil-Willson et al., 2019).

Another generalization, however, which is slightly more targeted is "white supremacy." Blee synthesizes its definition as: "extreme rightist groups and networks that advance explicit goals of white superiority and white power" (2017, p. 9). This understanding of social movements research is in line with the sociological definition of white supremacy (Rodríguez, 2017). This definition is usually operational for "classical" white supremacist groups like the Ku Klux Klan, Christian Identity, neo-Nazis, and white power skinheads, but some contend that it becomes insufficient when labeling newer groups which deny that racism and domination are definitional characteristics of their identity. Some examples are groups self-categorized as identitarian or "altright," which vehemently deny accusations of racism or white supremacy, and may even take pains to distance themselves as much as possible from those who openly express racist beliefs. Journalist Andrew Marantz documented the efforts that certain factions within the alt-right took to distance their flavor of extremism, "civic nationalism," "identitarian nationalism" or "American nationalism" from that of the likes of Richard Spencer and Mike Enoch, leaders of the "white nationalist" alt-right (Marantz, 2019, Chapter 25). I will presently refer to these groups as identitarian, by distinction from "(classical) white supremacist."

At the core of the professed ideology of identitarian groups is a concept of threat to identity. They claim that they are not gratuitously hating others, people of color or immigrants because of their difference, but, instead, believe that the special character of their own white identity is being threatened by immigration and the liberal understanding of multiculturalism as mingled diversity. Therefore, in order to preserve "true diversity," people of different colors and cultures should live in separate nations. Such a policy, in their opinion, would preserve "true diversity," which liberal multiculturalism cannot handle. Common sense therefore dictates that it is not necessary for them to believe that whiteness is superior, just to have commitment to "true diversity." Some researchers see in these claims enough reason to apply a new terminology to these groups, namely "(white) identitarian," "white nationalist" and "the New Right" to operationally distinguish them from white supremacist groups which center hate and domination by virtue of a constructed superiority of the white race, a.k.a. "the Old Right" (Guenther et al., 2020; McNeil-Willson, 2020).

In line with another group of researchers, I argue that this distinction should not be applied when studying the ideology of these groups. Instead, the approach of Blee (2017), Hartzell (2018, 2020) and Sanchez (2018), among others, is to continue to apply the label "white supremacy" to these groups as well, despite acknowledging that differences between older and newer far right groups exist. The reasoning is that these differences merely reflect a rhetorical, deceptive, or strategic framing of white supremacist beliefs as a way to flip and abscond them, in order to make them more socially acceptable, but their ideological core is still shared with the likes of the Ku Klux Klan (Blee, 2017; Hartzell, 2018, 2020; Sanchez, 2018).

I contend that the practice of labeling both older and newer extremist groups as white supremacist, as long as their ideological system rests on race or white identity, has a major advantage. Namely, this approach renders useless these groups' efforts to obfuscate the importance of the racial superiority of whiteness in their belief systems. Instead, it makes plainly visible that there is an ideological equivalence between old and new versions of white supremacist ideologies, despite surface rhetorical differences. Otherwise, social movement scholars may be susceptible to be commonsensically blinded to white supremacy, much like Phillips and Yi (2018) have been when they have non-ironically claimed that Richard Spencer was, in fact, "not a violent supremacist," because he presented as an activist for "a peaceful approach to preserving white national identity and returning to traditional Western values" (Phillips & Yi, 2018, p. 223). The two researchers made the mistake of taking "national identity" and "Western values" at face value, completely ignoring the contextual cues like calling for a racial European white empire (Spencer, 2016), among others, that would surely label Richard Spencer as an activist for a violent, racist ideology.

One caveat should be stressed however. This encompassing labeling of white supremacy should not be done as a generalization, erasing the differences between groups, and sweeping them under the same banner. The usefulness of the "white supremacist" label comes from its functioning as a reduction to the ideological core of the groups' ideology, a minimalist definition. It should still allow for a sub-classification based on other traits of the groups. This allowance is important because of the high fragmentation of groups and beliefs on the far right, as well as the characteristic decentralization augmented through internet-mediated communication. In the words of Marantz, "not every bad guy on the internet was bad in the same way" (Marantz, 2019, Chapter 23).

2.2. Rhetorical Distance: A Mechanism of Reframing

To support the assertion that both identitarian far right ideologies and racist extremism share a white supremacist core, I introduce a theoretical distinction coming from the field of communication studies, namely the concept of *rhetorical distance* as formulated by Hartzell in two excellent analyses of white supremacy within the alt-right and on Stormfront (Hartzell, 2018, 2020). Rhetorical distance explains my observations presented in the introductory section, it affirms the strategic value of the identitarian New Right's approach to white supremacy, and it can be used to guide the distinction and definition of different white supremacist groups and ideologies.

In Hartzell's theory (2018, 2020) rhetorical distance, as used within white supremacist groups and online platforms, is the practice of *rhetorically de-emphasizing racially motivated ideas* through discursive strategies in order to *substantively promote racial prejudice*. This is made possible by a feature of commonsense liberal conceptions of racism which assume that discourse is not racist if it is race-neutral.

The theorized mechanism of how rhetorical distance from white supremacy can paradoxically promote it involves the mobilization of the aforementioned common sense, along with affective processes, and an alleviation of discomfort when faced with racist ideas. Engaging these three processes in discourse can distance whatever white supremacist ideas are presented from the label of white supremacy and replace it with labels which sound more innocuous, such as "race realism," "identitarianism," "white nationalism," "alt-right" and others. A straightforward example of this would be an opinion such as: *the KKK are violent and irrational, they bate people of color and want to dominate them, whereas people in the alt-right simply want to be allowed to have a white identity, with a heritage that they can feel proud of, and preserve this identity through rational, non-violent means which protect white interests.* It follows from this sample reasoning that the altright has nothing in common with the KKK, which means that it is not white supremacist. Following the described mechanism, this argument can mobilize an affect in the discourse (the pride of belonging to a shared identity) which becomes an emotion when someone does decide to join an alt-right group (belonging to a community which is proud of their shared identity). It also contains a judgment of common sense (e.g., liberal multiculturalism allows every minority to be proud of their heritage, therefore there is nothing wrong or extreme about white people who also feel proud of their own heritage) which, in turn, alleviates the discomfort that someone might feel at contemplating the thought of taking pride in being white.

In summary, hate of others becomes love of one's own identity, and guilt about the legacy of slavery, imperialism, colonialism, structural privilege become pride of culture and traditions in what Hartzell (2020), building on Ahmed's theory of the role of affect in relating to ingroup and outgroup (Ahmed, 2004), has identified as the white supremacist "affective slide." As Arviv et al. (2020) poignantly put it, "it's a thin line between love and hate" when it comes to online white supremacist speech.

The mechanism of affective mobilization may also involve cultivating a community. When looking at Stormfront at least, researchers have found that interactions on the site have the characteristics of a community and instill, in some members, a feeling of belonging (Bowman-Grieve, 2009; Caren et al., 2012; De Koster & Houtman, 2008; Hartzell, 2020; Thompson, 2001).

The avoidance of shocking, readily recognizable white supremacist language and imagery is conducive to the management of discomfort. Instead of using what can be easily identified by human readers and social media algorithms alike as hate speech or racist phrases, imagery and symbols, groups which want to escape the label of white supremacy refrain from using the shock value of their ideas and adopt a civilized and calm language. Online, this is achieved through the use of moderators or through automation. Stormfront goes so far as to censor posts which use grammatically incorrect language and promotes good writing in its community rules. It also automatically censors the most obvious slurs (Hartzell, 2020).

2.3. Competition of White Supremacy Frames

White supremacy, just by virtue of being an extremist ideology, is generally and commonsensically thought to be direct, overt. My observations and previous research presented so far show that overt white supremacy may not be the dominating strategy of communication within these groups. How effective covert white supremacist online discourse can be at making white supremacist ideas more palatable when compared to classic overt white supremacist speech is an important question. The effectiveness of covert white supremacy would depend on whether there was a competition between covert and overt white supremacy, and which one dominated the discourse in white supremacist online communities.

2.3.1. *Ideologically covert frames:* A Theoretical Contribution and a Choice of Operationalization

Investigating this issue requires, first, a way to operationalize rhetorically distant / covert white supremacist speech and distinguish it from direct / overt white supremacist speech. The method most fit for this purpose seems to be frame analysis and theoretical framework, which has been preferred by researchers looking at the online discursive practices of white supremacist groups (e.g. Cabrera, 2014; Guenther et al., 2020; Jackson, 2019; Knüpfer et al., 2020; Wright, 2009), who refer to the phenomenon of covert white supremacy as "strategic framing."

Indeed, much of the discursive work of rhetorical distance can be achieved through reframing by shifting the semantic field used. Framing analysis in social movements research has been used to explain how ideas of social movement members become aligned with the social movement's perspective through discourse practices, a phenomenon referred to as frame alignment processes (Snow et al., 2014). Rhetorical distance as practiced in white supremacist discourse can be considered a type of alignment process in the sense that it bridges the white supremacist ideology of an existing group with the common sense of newcomers by deliberately showcasing some aspects of white supremacy and obscuring others (Entman, 1993).

The framing strategies of covert white supremacy can replace what is stereotypically recognizable as the semantic field of white supremacy, nationalism, and racism with dog-whistles.

Dog-whistling has long been studied in relation to how race is latently constructed and how elites can cue latent racist attitudes through political rhetoric (see Wetts and Willer 2019 for a short review). Unfortunately, much less attention has been given to the usage of dog-whistling frames in relation to white supremacist discourse. Some examples of white supremacist dogwhistles come from studies looking at how the alt-right use language to escape censorship on social media (Bhat & Klein, 2020) and how certain imagery, internet memes and hand gestures have been coded to signal supremacist ideas (Drainville & Saul, 2020). To illustrate, Bhat and Klein (2020) found that the word "skittles" came to be used as a euphemism for "Muslims" after a tweet made by Donald Trump Jr. using the phrase "bowl of skittles" as an analogy to the migrant crisis. Likewise, the number "1488" is used as an acronym for the slogan "we must secure the existence of our people and a future for our white children" (14) and "Heil Hitler" (88) (Bhat & Klein, 2020). These are classic cases of dog-whistling as coded language that simultaneously hides its meaning to one group of people and provokes strong reactions in another group (López, 2015).

However, there is a crucial difference between using signifiers as codes (e.g. the case of "skittles" and "1488") and using strategic frames of white supremacy. Codes still express the directly white supremacist meanings, whereas strategic frames obscure not the meaning of the words, but the ideology of the ideas expressed. For instance, the meaning of "1488" is still in line with racial hate and fascism ideologically, even though its meaning may be obscured to the uninitiated. However, the expression "ancestral homeland" (Bhat & Klein, 2020) entirely replaces racial hate with an emotionally positive meaning. It therefore obscures the ideological loading of

the white supremacist construct that white people are entitled to take back an "ancestral homeland" which is now "endangered" by, say, immigration. Instead of racial hate, an imagined aspect of whiteness is being elated (Bhat & Klein, 2020). The current project introduces this important theoretical distinction that is usually ignored in the literature. Strategic white supremacist frames (obscuring the white supremacist ideology) should be operationally distinguished from dog-whistles (obscuring the white supremacist meaning).

To further clarify this distinction, consider the example of framing racist hate as "protecting traditional Western values." This still functions like a dog-whistle, but it is related to the ideology linked to this expression, not to the immediate meaning of the words. There are those who understand that this means excluding non-whites, non-Christians, non-cis people, nonheterosexuals, but, at the same time, it comes with the advantages of sounding nice to those who do not understand the specific construct of "Western values" in context, for whom the exclusion implied by the expression is hidden.

This dynamic of hiding and highlighting is also consistent with Entman's theoretical elaboration in communication studies of how framing works (1993). According to Entman, framing operates through selection, highlighting one aspect of reality or several bits of information to guide the interpretation of an issue, and salience, which deems only the selected aspects as being important and, therefore, it deliberately hides all others. Confusingly, both white supremacist strategic framing and classic dog-whistles operate though selection and hiding. For the sake of clarity when working with the two as distinctive concepts, I will refer to white supremacist strategic ideological framing as *ideologically covert frames* or *rhetorically distant frames*, while keeping the term *dog-whistles* to mean semantically covert white supremacist language. Furthermore, because dog-whistles still directly express racial hate, for the purposes of this project, I will include them as an operationalization of ideologically overt white supremacy. Ideologically overt and ideologically covert frames are the two framing strategies for which this project aims to design a method of detection.

The phrases mentioned before, "Western values" and "ancestral homeland" (Bhat & Klein, 2020), along with "European values," "White civilization," "White culture," "heritage," "White genocide," "the great replacement," "anti-white," "invasion" (Berger, 2016; de Gibert et al., 2018; McNeil-Willson, 2020; Sanchez, 2018; Saresma, 2020) are also part of the ideologically covert white supremacist frame. As explained before, these hide the prejudiced exclusion at the heart of the ideology they express and instead highlight the value of whiteness. These establish distance from classical white supremacy and offer the possibility to superficially replace it with a different claim, which may be identitarianism, racial realism, national security, white civil rights, and others.

Meanwhile, what semantic white supremacist dog-whistles achieve has more to do with thwarting hate speech detection algorithms and hiding in plain sight.

The semantic field of classic overt white supremacist language (expressing racial hate without dog-whistles) along with white supremacist dog-whistles can be identified from studies looking at expressions of racialized hate speech (such as Alatawi et al., 2020; Association for Computational Linguistics et al., 2018; Brown, 2009; de Gibert et al., 2018; Gerstenfeld et al., 2003.; Wahlström et al., 2020), including those reviewed by Arviv et al. (2020). The Methods chapter will detail the chosen operationalization of ideologically covert and overt white supremacy frames, specifying the terminology included along with the literature sources upon which the terminological selection was made.

2.3.2. Networked Framing

Framing theory and analysis operates with two distinctive and, sometimes, complementary concepts of frames (Benford & Snow, 2005; Oliver & Johnston, 2006). On the one hand, frames are understood as fixed cognitive maps, schemas or scripts which enable people to understand many different new iterations of a situation, and offer ways of interpreting the multitude of content they encounter every day. On the other hand, frames can be understood as emergent interpretative constructs which are formed through interaction.

When strategic framing is discussed in studies of social movements, it can be approached from both angles (Benford & Snow, 2000, 2005). The concept of emergent frames is needed when discussing frame-alignment processes: a movement may strategically propose frames which are likely to modify the way newcomers or opponents think about an issue. The emergent quality of strategic framing may also consist in a negotiation of frames internal to the social movement. However, when prominent sources of opinion in a movement disseminate strategic frames in a top-down fashion, even when adopted, frames are no longer negotiated, they are given, static maps of interpretation (Benford & Snow, 2000, 2005).

At the level of analysis of frames, when operationalizing them in discourse or content analysis, researchers have no choice but to reify the semantic boundaries of frames in order to identify, characterize and compare them (J. M. Jasper, 2017). This usually also implies temporal limits to when the studied frame occurs.

If the time-window of the occurrence of a certain frame is enlarged, then it is almost certain that processes of negotiation and contestation of the frame become observable, because of the social, political and ideological contingence of frames (Benford & Snow, 2005). Therefore, any type of frame analysis which takes these contingencies into account must be methodologically finetuned in order to be able to capture the dynamism and change of negotiated frames of interpretation.

One way of achieving such fine-tuning is the networked framing approach (Meraz & Papacharissi, 2013; Stewart et al., 2017). According to Meraz and Papacharissi (2013), the proponents of this method, social network analysis can assist in illuminating the dynamism and contention of frames because it is able to pick up on the mechanisms of frame diffusion in a certain network. Of course, such an approach is only possible where researchers can collect network data surrounding a frame, which, in the case of working with social media data, can be accessed relatively easily through metadata (Peita, 2021).

Using network data, this approach can trace the movement and evolution of frames within an online network (Meraz & Papacharissi, 2013; Stewart et al., 2017). With Twitter data, for instance, hashtags can function as proxies for the employment of a certain frame by the author of a tweet. At network level, the popularity of the hashtag on Twitter gives an indication of the prevalence of the frame, and when different times in the life of the network are selected, hashtag use indicates whether the frame has a strong diffusion, whether it is in competition with another frame or whether it has overcome echo chambers and protruded into other groups in the network than the ones in which the studied frame originated.

2.4. Prevalent Methodologies in Studying Frames of White Supremacy

Most of the studies using the concepts of strategic framing, dog-whistles, covert white supremacy and rhetorical distance, however, evade the possibility of capturing the dynamism of evolving frames. This is because they employ some form of qualitative content analysis on specific, static case studies to demonstrate the strategic role of distancing white supremacist ideas from their ideological background. This approach binds the studied movements to a narrow temporal window and a static formulation of frames.

Nevertheless, this body of research does contain important findings which can inform the methodological approach proposed by this work.

Hartzell (2020) shows that Stormfront's moderators push for the language and content on the website to appeal to mainstream white audiences through introductory posts which set rules for using the website, meant to be read by all new visitors and members: "Introduction and FAQ," "Guidelines for Posting" and "Questions about this Board." The study shows that the intent of moderators is to associate what they call "white supremacy" with negative affect from which they distance "white nationalism," associating it with positive affect (Hartzell, 2020). Unfortunately, the study stops at three posts of moderator's content and does not go further to examine whether the intent of the moderators is actually complied with by other members of the board. More aggressive policing of far right extremist groups prompted by the Oklahoma City bombing of 1995 perpetrated by white supremacist anti-government domestic terrorists (Southern Poverty Law Center, 2020b) may have contributed to the decline in organized white supremacist groups offline, and the increase in the online dissemination of white supremacist ideas (Wright, 2009). Wright (2009) contends that strategic framing of white supremacist ideas online was an important factor in the resilience of far right movements and groups post-2000.

Another study (Sanchez, 2018) compares the meaning of the words "patriotism," "heritage" and "security" in a propaganda campaign by a KKK division which took place in 2017 in a small town in Texas, with declarations made by Donald Trump in his 2016 presidential campaign. The comparison shows that the same words conveyed explicitly racist ideas in KKK's propaganda and were used as racist dog-whistles in Trump's rhetoric. Here, the "versatility of white supremacist rhetoric," as the author puts it (Sanchez, 2018), operates a slide from the fringe (the KKK) to the mainstream (a presidential campaign).

Åkerlund's approach (2021) is somewhat more systemic, in the sense that it tracks the occurrence of a coded white supremacist phrase, "culture enricher" (somewhat of a Swedish equivalent to "social justice warrior" or "snowflake" in American online culture), from its first appearances in the 1990s until today. The study manages to show that this coded expression contributed to the "mainstreaming" of far right discourse online (Åkerlund, 2021).

The accumulation of findings from qualitative research alone can increase confidence in the idea that some white supremacist actors, in some contexts, do use strategic framing in order to move their ideas more towards the mainstream. However, more systematic approaches are needed to generalize these findings, both when it comes to the size of the white supremacist networks and the size of the corpora of content produced by them which are being analyzed. What is lost when analyzing only small corpora of content produced by small groups of actors is a sense of the scale of the phenomenon of white supremacist strategic framing, and an image of how these frames evolve in time through internal negotiation or outside contestation. Is it a practice adopted by enough members of a white supremacist group or an online network that it may become consequential for recruitment or the social acceptability of the group? Do members of Stormfront buy into the discursive strategy pushed, as shown in previous research (Hartzell, 2020), by the board moderators, or is it just a top-down direction that is ignored by most rank-and-file registered members of Stormfront?

2.5. Social Media-Appropriate Methods of Studying Frames of White Supremacy: Content and Networks

At a time when the computational capacities and the methodologies of dealing with large quantities of data have been developed to unprecedented standards, it would be regrettable if scholars looking at political extremism and movements on the far right were to miss taking advantage of these new capacities of analysis. This is especially the case since, when it comes to dealing with online social networks, large quantities of data are being produced every day, and an important amount of these data can be accessed at almost no cost to the researcher. I believe that big data and social media analytics can significantly expand our understanding of how white supremacist communities communicate online and how their online behavior spills into the offline.

There are already some promising results obtained from big social media data dealing with strategies of communication in white supremacist networks. Unsurprisingly, most of them come from fields like computer science and software engineering, and generally make less use of established communication and social movement concepts, such as framing, to contextualize their findings. Studies coming from social sciences, on the other hand, tend to be less methodologically elaborate and analyze data at a smaller scale.

Most often, studies taking the computational social science approach to the far right apply either some form of automated content analysis alone or mixed with discourse analysis (Bhat & Klein, 2020; Guenther et al., 2020; M. Jasper et al., 2021; Peita, 2021; Than et al., 2020), social network analysis alone (Chau & Xu, 2007), or a mix of the social network analysis and natural language processing methodologies (Berger, 2016; Rauchfleisch & Kaiser, 2020). In some cases, both have been combined with an experimental design as well, like the fascinating work of Kruglanski et al. (2014), worth mentioning despite being situated somewhat outside of the scope of this literature review. Such analyses have been applied to social media data obtained either through a custom algorithm or an archive (Peita, 2021). Even when opting for content analysis alone, the size of the data allows for some form of quantitative content analysis which, albeit less precise than analyses based on hand-coding or discourse analysis, has the advantage of demonstrating a wider, maybe even platform-wide tendency. Of note are some researchers' efforts to develop algorithms which can not only download but also automatically identify extremist online content (Bouchard et al., 2013).

Generally, Social Network Analysis is able to significantly enrich study design when coupled with content analysis including Natural Language Processing compared to using content analysis alone. The combination of the two can also broaden the scope of possible research questions from inquiries into what themes, frames and communication strategies can be identified in the corpora to questions of how these elements of discourse are being disseminated through a network of social media users or which of the identified elements of discourse are more effective for a specific purpose. As mentioned earlier in this chapter, network analysis coupled with frame analysis has the capacity to shed light on the processes of contestation and negotiation of frames over time, in the networked framing approach (Meraz & Papacharissi, 2013).

Examples of procedures of NLP which can be used on their own or mixed with SNA are topic extraction, sentiment analysis, and even the ingenious networked content analysis, which is able to construct bipartite networks of authors / documents (of say, tweets or Facebook comments) and the semantic fields they use. In some cases, frame detection relies on detecting groups of terms and communities of authors in such networks (Radicioni et al., 2021 offer an intreaguing application of this methodology). SNA may rely on analyses based on ego networks which map the networks of only a few individuals relevant to the study. In this case, as pointed out by Peita (2021), the difficulty is that the researcher would have to already know who the important white supremacist actors are whose networks are worth mapping. However, a simpler, but possibly more computationally intensive approach is to map the entirety of a known white supremacist network or to select a significant sub-network, or sample it. Examples of useful SNA methods which may be applied to such networks are network and node-level characteristics (Panizo-LLedot et al., 2019), especially when used as predictors or effects for an observed variable outside the structure of the network.

There are some studies which attempt such designs to investigate frames of white supremacy. A good effort is a George Washington University report (Berger, 2016) employing mixed methods on Twitter data before Twitter banned a series of accounts pushing far right ideas. The report found that the most discussed issue on far right Twitter was that of "white genocide," a concept which squarely fits into the reframing from white supremacy to a sense of threat to identity posed by such contentious issues as immigration and multicultural policies. Interestingly, the report also found that the white supremacist Twitter network was more effective in recruiting others than the ISIS Twitter network (as interpreted by Peita, 2021). Recruitment relied heavily on elaboration and reinforcement of the concept of "white genocide" and accusations that non-racist or anti-racist content was actually "anti-white" (Berger, 2016). A Twitter dataset containing only alt-right accounts during the 2016 US presidential election contained, in addition to the use of "#WhiteGenocide," the hashtag "#ItsOkToBeWhite" (Panizo-LLedot et al., 2019 as cited by Thorburn et al., 2018). Findings like these have led some to conclude that the concept of existential threat to whiteness is indeed a way of normalizing a white supremacist ideology (Miller-Idriss, 2020). An examination (van der Vegt et al., 2020) of the language used by political content creators on YouTube to refer to far right topics before and after the "Unite the Right" rally in Charlottesville found that alt-right videos preferred the term "white nationalism," whereas leftist videos used the term "white supremacism" as an equivalent.

Research in this area has also found that a pipeline of recruitment for far right extremist online communities is constituted by the so-called "Manosphere" on YouTube and Reddit, namely anti-feminist online communities discussing topics such as men's rights and pick-up artistry (Mamié et al., 2021). Strong evidence for the existence of a far right radicalization pipeline on YouTube was brought by Ribeiro et al. (2019). The authors (Ribeiro et al., 2019) examined an impressively large amount of videos and comments from the platform and found that channels with moderately conservative content categorized as "Alt-lite" (promoting "civil nationalism" as a distinct construct from "racial nationalism") and "Intellectual Dark Web" (promoting ideas by "iconoclastic academics" such as Sam Harris, Jordan Peterson, Ben Shapiro and others) channel their consumers towards fringe alt-right channels.

A particularly interesting project (McIlroy-Young & Anderson, 2019) using fairly simple word and topic frequency statistics over time asks the question whether the online activity on the social media platform GAB of the Pittsburgh Synagogue shooter could have been used to predict his violent political act. The authors found that the shooter was one of the most extreme users of the platform, both in terms of how much he used the platform, being on the higher tail of number of posts and reports, and in terms of the language used (McIlroy-Young & Anderson, 2019). However, there were significantly more users on Gab at the time who had a higher ratio of anti-Semitic posts than the shooter (McIlroy-Young & Anderson, 2019). Consequently, the mere quantity of extremist language used online seems to be a poor predictor of extremist violent behavior.

Most of the studies included in this short review have used computational social science methodologies to study far right online communities on English language social media networks in the United States. There are fewer examples of such designs applied to online manifestations of white supremacist discourse from Europe. Among them, an interesting finding for the topic of this research project came from connecting the content of YouTube comments to the typology of videos published by far right content creators in Germany (Rauchfleisch & Kaiser, 2020). The study showed, at least when it comes to German language channels, that the audience of "New Right" type of channels are far more concerned with identity issues, while the audience of "Old Right" themed channels are more prone to use insulting language in the comments sections of videos (Rauchfleisch & Kaiser, 2020). Twitter accounts of the far right in Germany and Austria have demonstrated a strategy of hijacking political frames initiated by the left, detected through a networked framing analysis approach (Knüpfer et al., 2020).

The studies selected in this section illustrate the breadth of questions that can be asked by including computational methods like NLP and SNA when studying white supremacist online groups. I would stress that designs like these are clearly more appropriate for application on large-scale social media data. However, the majority of studies found were undertaken by information technology scientists, which indicates that social scientists may have to deal with a training gap in order to bring computational methodologies on an equal footing with qualitative methodologies in the study of the online far right in their fields.

2.6. Research Questions and Aims of Study

The construction of rhetorical distance from white supremacy may have been a contributor to far right online radicalization and the uptick in racially motivated political violence from the past decade. While drawing a direct line of influence based on analysis of the online environment would be unwise, as the case of the Pittsburgh Synagogue shooter demonstrates (McIlroy-Young & Anderson, 2019), determining how white supremacist ideas are framed online does have intrinsic value.

The questions guiding this study are:

a) Over the past decade, which of the two frames of white supremacy theorized – ideologically covert and ideologically overt white supremacy – was the dominant way of expressing white supremacist ideas on Stormfront? Was there a transition identifiable from one to the other over time?

b) Is the importance of users on Stormfront a good predictor for the use of either of the two frames?

Based on previous findings reviewed in the above sections of the paper, I have the following hypothesized expectations:

- 1. A transition from more overt ways of expressing white supremacist ideas earlier in the life of Stormfront to more rhetorically distant ways in recent time is expected;
- 2. The more important users on Stormfront push either the ideologically overt white supremacy frame (more in the distant past) or the ideologically covert frame (more in the recent past and in the present).

To answer these questions, there is no better online source of white supremacist ideas than Stormfront, the oldest, most popular, and most established platform of its kind. By selecting Stormfront, I am also stacking the cards against my expectations of finding rhetorically covert white supremacy frames. The reason is that Stormfront is likely to have many users who have been on the platform for many years and who may be more familiar with the ways of the Old Right than the newer ideas of the likes of the alt-right. My reasoning is that, if the study can indeed find rhetorically distant white supremacist framing on Stormfront, a bastion of online white supremacy, then there is a higher chance that this type of strategic framing is being used in other far right online communities as well, especially those that may be more open and frequented by people who do not embrace a far-right ideology.

Along with addressing the question of the types, prevalence, and evolution of framing white supremacist ideas on Stormfront, this study intends to develop a computational social science mixed-methods research workflow for studying white supremacist online communities. This is meant to function as an example and to provide the possibility to replicate this workflow using data from different social networks.

The project illustrates a good practices workflow from extracting online data to data transformation, to applying a mixed methodology based on NLP and SNA. The study will also be

accompanied by full replication material in the hopes that it can also contribute to closing the gap between social scientists and computational scientists when it comes to working with social media data.

Since Stormfront does not have hashtags like Twitter or Twitter-like platforms, the approach to networked framing put forth by Meraz & Papacharissi (2013) will not be viable in this case. In fact, the use of hashtags to detect frames on social media data is pervasive in the literature to the detriment of other methods. For the alternative in which hashtags are absent, other quantitative frame detection methods have been explored.

3. Methods

3.1. Dataset

Data used to answer the proposed research questions were extracted from Stormfront between March 22nd and April 24th, 2021, using a custom-made Python script. Data downloaded included both metadata (the date when a message was posted to the forum, username, whether the message is a reply to another user) and the text of each message. Images were excluded, as only text was needed for the proposed analysis.

Only the sub-forum "Ideology and Philosophy" of Stormfront was selected to be scraped, because this is where most of the politically themed discussions take place on the platform. To reduce the size of the resulting corpus and network, selection was further applied to the discussion threads featured under the Ideology and Philosophy sub-forum: only threads with more than 6000 views were included, and these were scraped in their entirety, that is, from when the thread was opened to the last post included at the time of scraping. This selection ensured that the data are within the scope of the study – popular political discussions on Stormfront – while also resulting in a computationally manageable dataset.

The resulting dataset contained 21,236 individual comments, posted by 2,653 users, across 256 threads, between 2005 and 2021, covering a period of 17 years.

3.2. Protocol

This section details the steps undertaken as part of data extraction, pre-processing and analysis.

1. *Data scraping.* The first step was to develop an HTML crawler in Python which would download the needed data from *Stormfront.org.* Data obtained this way were saved locally in their raw format. The script was tested on March 22nd, 2021 by extracting two of the most popular topic threads of the Ideology and Philosophy forum. The rest of the selected threads, filtered to only those which had amassed over 6,000 views at the time of scraping according to the website

statistics, were extracted on April 24th, 2021. The code for the scraper can be found in *Annex 1*.

A note on the ethics of accessing these data is that only threads which were publicly available were accessed. All the names of the users are already anonymized by the website through the use of usernames. Other metadata possibly tracing back to the identity of the user, such as location (also publicly available), were deleted. No internet addresses were accessible through the scraper. I have searched, to the best of my ability, for terms and conditions of the website prohibiting the automated extraction of data from Stormfront. None were found. Therefore, the risk of injury to any of the users included in the dataset is null.

- 2. *Data preprocessing*. Following data extraction, data cleaning steps, including some specific to bagof-words automated text analysis methods, were applied to ensure correct analysis: removal of any lingering HTML tags and unwanted metadata, removal of links and special characters, ensuring that date formats are readable to the analysis algorithms, tokenization, and removal of stop words. See *Annex 1* for the specifics of each data cleaning step.
- 3. Frame detection. Of interest to this study are two frames which were theorized in the previous chapter as possible ways of expressing white supremacist ideas online: ideologically distant / covert white supremacy and ideologically overt / direct white supremacy. Henceforth, they will be referred to as covert and overt frames.

Stormfront is a discussion board and does not have any feature similar to hashtags like other social media platforms. Hashtags have often been used to detect the frame of a comment. Because of their absence, a different strategy had to be articulated. The next best procedure was to apply a content analysis step based on categorization of the language in the Stormfront comments as overtly white supremacist, covertly white supremacist or neither. One of the most straightforward ways to achieve this is though dictionary-based text analysis. The task could also be done through either a supervised or unsupervised machine learning classifier, which might yield more precise results compared to a bag-of-words dictionary-based analysis. However, these methods were excluded due to feasibility concerns, in favor of the simplicity of the dictionary-based method. A future development of this analysis workflow should include a machine learning classifier to enhance the precision of the frame detection procedure, or as an alternative comparison method to dictionary-based analysis.

There are already semantic fields associated with either overt or covert white supremacy identified in the literature of online hate speech and white supremacist discourse, as reviewed in the theoretical framework chapter. A dictionary analysis allows for the integration of terms and dog-whistles from the semantic fields from previous findings, as well as the adaptation of the method to the particular language of Stormfront. For instance, many of the terms associated with overt white supremacy include racial slurs. However, I already know that Stormfront uses both algorithms and moderation to weed out the most obvious slurs so they are not likely to appear in any of the published comments. As such, racial slurs cannot be used in this corpus, but that does not mean that overt white supremacist ideas are absent from Stormfront. This is where an inductive approach to finding the particular terms used to express racist ideas directly was added.

Therefore, a deductive approach to dictionary construction was taken by extracting terms associated with online white supremacy from the literature. This was followed by an inductive stage of hand-coding a random sample of 1% (n = 212) of the posts in the corpus and extracting words which seemed, at face value, to express either covert or overt white supremacist ideas.

At this exploratory stage, a topic modeling technique based on Latent Dirichlet Allocation (LDA) was also applied, in order to get a broader picture of the subjects of discussion on the forum and identify other words highly associated with racialized topics, which may not have been found from the literature review or the hand-coding stages.

Lastly, two dictionaries were constructed, one for covert framing of white supremacy and one for overt framing. The resulting content of the dictionaries is available in the Analysis section, under the Construction of dictionaries subtitle. Applying a frequency-based analysis to the corpus using the two dictionaries yielded a score equivalent to the number of words found in a dictionary for each post to the forum.

In summary, the frame detection procedure included the following steps:

- a) Quantitative exploration of the corpus including LDA topic modelling;
- b) Selection of dictionary terms from the literature;
- c) Selection of terms from hand-coding;
- d) Construction of dictionaries and dictionary analysis.

4. Social Network Analysis. Following the dictionary-based text analysis step, the dictionary scores of posts were aggregated at the level of the forum user, such that each user in the Stormfront network was scored on their usage of ideologically overt and covert white supremacy. These scores were imported into a network analysis tool, *NetworkX* implemented in Python (Hagberg et al., 2008), and used as node-level attributes, which allowed for the study of dictionary scores in relation to network-specific indicators.

Links between users (referred to as ties going forward to comply with the sociological paradigm of SNA) were based on quotes between site users. Quotes can be considered direct replies to the quoted users. Surely, these are not the only types of ties in the Stormfront network. Even simply going on a thread and reading a user's post can be considered a tie, in the sense that the reader is being exposed to someone's ideas or is being communicated to. However, the website does not keep any such records. Similarly, quotes are not the only types of replies. It is possible that someone may reply by mentioning another user's handle without actually quoting anything that the target user has said before. In this respect, basing ties on the quote feature is likely to underestimate both the total number of ties existent in this network by only including one type of link and the total number of reply type links. However, it is the only kind of link that the features of the website offer a record of. Quotes are also the surest way to ascertain whether a user was exposed to someone else's ideas on the platform, and is a

measure of engagement with these ideas. Quotes are, therefore, the fittest indicators for how the studied frames move in the network from quoted users to users engaging with their ideas.

In order to address the evolution of the overt and covert dictionaries in the network longitudinally, and also, due to the fact that the data covered a period of 17 years' worth of posts, an analysis based only on the aggregated network would have been inadequate. Instead, network indicators were calculated longitudinally, across static cross-sections of the network recreated from all the posts published during each of the 17 years. This resulted in 17 crosssections. The size of the one-year temporal window is arbitrary and artificial, and there are more involved ways of longitudinal network analysis with rolling windows, the size of which can be determined from the structure of the data. However, there is no one method of selecting the size of the time-frames or the type of longitudinal network analysis that has been recognized in the literature as superior to others, and the choice is often dependent on the dataset (Uddin et al., 2017). Yearly cross-sections have the advantage of simplicity, and I argue that they are fit for this dataset because they essentially reconstruct the network of active users on the website during each year. In effect, this yields, for instance, network centrality indicators about which users were most active or most important during a year, even though they may not have even been registered to the website the previous year or stopped posting the next year. This approach can render what the network looked like at each point in time, whereas, by collapsing all 17 years into one network, it would not be possible to capture change in the network structure.

Network indicators were taken at the level of each cross section and compared to the aggregated network. Node-level centrality indicators were also taken and used as predictors in simple linear regression models for the usage of the two frames.

4. *Sanity check*. Lastly, to ensure that the results obtained from predicting dictionary use with nodelevel centrality measures were not confounded by an inadequate dictionary measurement, the
overlap of the two dictionaries within the network was calculated based on the percentages of nodes using both dictionaries.

For clarity, the following figure (*Figure 2*) offers a visual representation of the workflow described thus far.



Figure 2 Representation of workflow of data extraction, pre-processing and analysis

4. Analysis

4.1. Detection of Frames of White Supremacy

4.1. 1. Quantitative exploration of the Stormfront corpus

Before attempting to detect and then capture the evolution of the usage of overt and covert frames of white supremacy on the Ideology and Philosophy sub-forum of Stormfront, it was important to understand its level of activity over time. There is a definite increase in the number of posts submitted to the subforum since 2016, as shown in the first part of Figure 3, which may be related to the ascent of President Donald Trump and the contested election of 2020. There may be a connection between the levels of activity on the platform and controversial or turning-point presidential election years, as the number of posts peaked in 2008 to close to three thousand, and, in 2020, to close to five thousand posts. Between 2012 and 2016, the activity on the forum lulled to less than one thousand posts per year.



Figure 3 Period of activity on the forum and the most active users

During the seventeen years, the ten most prolific members of the forum, shown in *Figure* 3 with their screen names, posted between 200 and 500 comments each.

Unsurprisingly, the language used by the Ideology and Philosophy members and visitors is strongly centered around notions of white race, white nations and white people. "White" is, by far, the word with the highest frequency overall, with close to twenty thousand appearances. "Race," "Jew," "Hitler," "German" all feature among the most frequent words, giving a strong indication that people talk openly about national socialism and race-related issues on this platform, which, at first glance, would fall into the ideologically overt white supremacy frame.



Figure 4 Top 50 (chart) and top 100 (word cloud) stemmed unigrams in the entire Stormfront corpus

At face value, and based on the conclusions that can be drawn from previous qualitative analyses of white supremacist language as reviewed in the first chapter, some of the most indicative tokens for the two frames might be "hitler" and "jew" for the overt frame, and "nation" and "white" for the covert frame. "Nation" and "hitler" are densely distributed during the 2008 – 2010 period, which points to a possible weakness of bag-of-words methods for modeling the two frames. Indeed, if, during that time, a thread focused on national socialism was particularly active, it makes sense that the token "nation" would frequently appear in the construction "national socialism." Certainly, when looking at the results of an LDA topic modeling analysis (*Figure 7*), the

token "nation" is one of the highest loading words in three estimated topics, not just one. The three topics that the word contributes to can be broadly categorized as *government and business*, *national socialism* and *nationalism and race*. When contributing to the national socialism topic, the term is obviously part of an overt white supremacist frame, whereas, in the other two topics, it has a much higher chance of appearing in constructions of covert white supremacy.

The possibility to contextually distinguish the occurrence of the stem "nation" in covert framing expressions such as "white nation" is, therefore, weakened. However, the dictionary-based method should still be able to give a rough estimate of the appearance of the two frames discussions on Stormfront. A higher margin of error in the interpretation of these results should be taken and, at a later stage of development, this margin of error could be reduced by testing more contextually sensitive supervised NLP classifiers.



Figure 5 Distribution of posts containing the terms "hitler", "jew", "nation" and "white" over time. The y axis shows the number of word occurrences per post. Each dot is a post.

As an unsupervised exploration of the lengthy Stormfront corpus of comments, LDA topic modeling (Blei et al., 2003) is a good place to start distinguishing which constellations of words are likely to appear together across different comments and, therefore, can be categorized under a specific topic. As is the case with LDA, I had to arbitrarily select an appropriate number of topics. For this, I calculated perplexity scores for models with different numbers of topics, arriving at 10 topics as a good balance between a lower perplexity score, which indicates a better

generalization of the model and is a measure of fitness (Blei et al., 2003), and a comprehensible distribution of words across topics for semantic validity (see *Figure 6* below).



Figure 6 Evaluating LDA topic models

From the LDA topic modeling, I was interested in discovering which words tend to appear together, if there are any topics that can be assigned to either of the two frames and, generally, what the main preoccupations of the users of Stormfront are. *Figure 7* plots the beta values of the topics for the top loading words, and the figure caption contains a possible categorization of the topics detected. The concern with political issues is evident across almost all topics consistent with the platform categorization of the forum as "Ideology and Philosophy." One topic belongs, right off the bat, to the overt white supremacy frame, that is, topic 3, which I have deemed as discussing "national socialism." Topic 9, "identity," sits at the other extreme, clearly indicating a covert frame, resting on words like "human," "origin," "gene(tic)," which appear together with "white," "europe," "culture." Topics 10, "nationalism and race," 1, "history," 2, "government and business," and even 6, "values and society," are harder to categorize without context, as they may fit both frames.



Figure 7 Term loadings for topics detected through LDA 10 topics modeling. Topic 1: history; Topic 2: government and business; Topic 3: national socialism; Topic 4: too ambiguous; Topic 5: interacting with the forum; Topic 6: values and society; Topic 7: gender; Topic 8: too ambiguous; Topic 9: identity; Topic 10: nationalism and race.

The easy and evident differentiation between topic 9 ("identity") and topic 3 ("national socialism") confirms many of the findings reviewed before about the words usually associated with a covert versus overt type of white supremacist language. The words associated with topic 9 especially, "white," "european," "culture," "human," "origin," "hope" can be found among those singled out by other researchers as being indicative of strategic framing of white supremacist ideology. Likewise, the terms forming topic 3, national socialism, are also very recognizable as an open, direct way of expressing white supremacist ideas.

4.1.2. Construction of Dictionaries

Having had a type of confirmation from the quantitative exploration of the Stormfront corpus that the terms found in the literature on strategic framing are, most likely, fit for the language used on Stormfront too, I moved on to constructing the ideologically covert and overt

white supremacy dictionaries.

1. Selection of dictionary terms from the literature.

Table 1 below contains terms found in the literature along with their sources.

Term	Frame	Note on meaning	Source	
google*	ideologically	dog-whistle for "black people"	(Bhat & Klein, 2020)	
0 0	overt white	~ 1 1		
	supremacy			
skype*		dog-whistle for "Jewish"		
skittle*		dog-whistle for "Muslims"		
bing*		dog-whistle for "Chinese"		
butterfl*		dog-whistle for "LGBTQ+"		
yahoo*		dog-whistle for "Mexican"		
nuttin*		part of the dog whistle "Dindu		
		Nuttins" for "black people"		
goy		non-Jews who conspire with		
120/		Jewish plots		
1370 1488 or 14/00		"we must secure the evistence of		
1400 01 14/ 00		our people and a future for our		
		white children" and "Heil Hitler"		
cuck*		liberal or establishment		
euen		conservative		
beta*		used in "beta male" for liberals		
snowflake*, SJW*		liberals		
ape*			(de Gibert et al., 2018)	
parasite*				
fag*				
scum				
savage*				
filth*				
mud				
homosexual*				
beast*				
monke*				
libtard*				
coon [↑]				
nigiet				
pani guinea*				
enrechaun*				
inngle			(Brown, 2009)	
bastard*			(210 ,, 2007)	
negro*				
subhuman*			(Gerstenfeld et al., 2003)	
swine			(Wahlström et al., 2020b)	
carrion				
scum				
dregs				
vermin				

Table 1. Dictionary terms from the literature

barbarian*

ancestr*	ideologically covert	part of phrases like "ancestral homeland"	(Bhat & Klein, 2020)	
	white			
	supremacy			
Vinland	The second	expresses the right of whites to		
		an imagined historical claim over		
		North America		
Deus		part of "Deus Vult" ("God will		
		it" in Latin), expresses pride in		
		Christian Western culture		
proud			(Berger, 2016)	
genocid*		usually in "white genocide",	(McNeil-Willson, 2020)	
		"cultural genocide" or "genocide		
		against our people"		
western		"Western		
		values/culture/civilization"		
patriot*			(Sanchez, 2018)	
heritage				
secur*				
assimilat*				
protect*				
honor*				
safe*				
civil*				
pacesetter*			(Brown, 2009)	
doer*				
conquer*				
divers*			(Gerstenfeld et al., 2003)	
pride				

2. Selection of terms from LDA topic modeling

Table 2 adds some terms for which the LDA topic model has given strong confirmation that they may indeed belong to the covert frame, despite the fact that their usage is highly contextdependent in non-white supremacist settings.

Table 2. Dictionary terms selected from LDA topic models			
Term	Frame	Source	
european	ideologically covert white	LDA topic model ($k = 10$)	
	supremacy	applied on the Stormfront	
cultur*		corpus	

le de l'Alexand DA de l'e e

3. Selection of terms from hand-coding

Table 3 contains terms found from the last source of the two dictionaries, namely the Stormfront data itself. A random 1% sample of comments from the whole corpus was selected and submitted to close reading. Next, comments were labeled as overt, covert or neither, when it comes to the white supremacist ideology expressed. The main criterion was that of face value validity based on the premises of the two frames discussed in the theoretical part of this work.

Term	Frame	Source
genocide	ideologically covert white supremacy	1% random sample of comments from the Stormfront corpus
dominat*		1
resistance		
extinct*		
pur*		
sacred		
volk		_
arya*	ideologically overt white supremacy	
rude		
ghetto		
disrespectful		
mutilate		
hitler*		
reich		
88		
kampf		
raven		
cauldron		
mongrel*		
gyps*		
smelly		
sand		
promiscuous*		
pervert*		

Table 3. Dictionary terms selected from coded comments

To exemplify some of the choices made during hand-coding, a few of the comments selected in the random sample are shown below. The word "sacred" was found in comments discussing cultural and religious values, as well as gender roles. A good illustration is this comment from a discussion on the role of women in society: "The roles are sacred. Family is sacred.. it is re-defining served agenda. Specifically the re-defining constituted a devaluation of the role. It could be said that there is nothing sacred to those who do the (re)defining except for their agenda and the pursuing of it.."

While avoiding to openly express that women are in some way inferior, this user seems to

want to resist any gender role changes based on the "sacred" value of family.

The term "pure" was also used in comments expressing racial pride instead of racial hate,

like in this comment:

"it means, and its true, that we should care for our skin since we are born whites and we do not need anything other on our skin than our pureness..also means, that being white by birth, with white skin, but not white in mind its not good enough to be true white..i know there are lots of you with tattoos, smoking, drinking and you are all honest whites, but in my opinion i would try to keep my skin clear from everything that is not natural..being a white is having a white skin, but we need to be pure whites also with our mind and awareness about our race and who we are."

On the flip side, the decision to add a word to the overt dictionary came from reading a

comment as openly racist and trying to find the most unequivocally racist terms which had not

already been included from the review of the literature. One comment read in this key was:

"I was not taught to be racist. Elementary and jr high I never knew a black person. Once I got to high school we were bussed inside the city limits and it was awful!! Full of black ppl who were rude, thieves, they stunk, loud, disrespectful, just awful!! They are extremely jealous and hateful to white ppl. I have zero use for them!! I admit that there are occassional decent black ppl BUT, they have friends and relatives who are of the ghetto type and can and will turn to ghetto trash themselves!! Be polite but never bring one home with you!!"

4.2. Networked Frames of White Supremacy

Following the dictionary construction, the NLP part of the analysis was continued by scoring each post and each user with a covert and an overt dictionary value constituted from the number of words for each post and for each user which could be found in the dictionaries, counting every occurrence.

From 2005 to 2021, the usage of the two frames on the sub-forum reflects the periods of activity of the platform, with higher activity reflecting in higher frequencies of the dictionaries. Interestingly, between 2017 and 2018, there is a significant spike in the use of covert language not

accompanied by a similar trend in the overt dictionary. However, in 2018, this unusual spike dropped back in line with the overt dictionary rather than continuing to diverge.

There is a positive correlation between the two dictionaries and it varies over time between coefficient of r = 0.1 and r = 0.6 (Pearson) (*Figure 8*), meaning that their usage is associated, instead of the expected scenario in which when one would be on the rise, the other would fall.



Figure 8 Usage of overt and covert frames of white supremacy over time on Stormfront

Since every user was scored on both dictionaries, there can be overlaps between the two, which may explain the correlation between them. To check whether this overlap is substantively significant, *Figure 9* plots the top highest loading users on the overt dictionary (top half of the figure) and the covert dictionary (bottom half of the figure), along with their scores on the opposite dictionary plotted on the opposite side of the figure. When it comes to the most fervent users of one or the other of the frames, while there is some overlap, with the same three usernames being found in both top tens, generally, those with the highest scores on one dictionary tend to have lower scores on the opposite dictionary, but this relation is not linear or consistent by any means.

Top 10 Users with Highest Dictionary Loadings



Figure 9 Stormfront users with highest loadings in the Overt and Covert white supremacy dictionaries

At this point, a tentative answer can be drawn for the first set of research questions proposed, which is: Over the past decade, which of the two frames of white supremacy theorized – ideologically covert and ideologically overt white supremacy – was the dominant way of expressing white supremacist ideas on Stormfront? Was there a transition identifiable from one to the other over time?

While there are fluctuations over time in the amount of usage of ideologically overt and covert frames of white supremacy, for the most part, the two trends evolved parallelly. The largest gap between the two (between 2016 and 2019) did not seem to be sustained. From 2019 onwards, the two frames resume a parallel evolution, positively correlated. In most years, the covert frame

is slightly more used than the overt one, but the difference between the two is so narrow that it might easily fall within the margin of error of the dictionary-based frame detection method.

To add the networked aspect of the frame analysis, the distribution of dictionary scores was examined in relation to several network characteristics.

Following the reconstruction of the yearly networks of the Ideology and Philosophy subforum, seventeen network cross-sections were obtained. The last one, containing the users active during 2021, is incomplete, as the dataset only included posts published up to the date of data extraction. Additionally, the full network was also examined. The full network contains all the users active in the 2005 – 2021 period, and it includes cross-yearly links which are ignored in the yearly cross-sections. The following table shows relevant network-level metrics for the seventeen cross-sections and the aggregated network.

Network	Number of nodes	Number of edges	Density	Number of weakly connected components	Number of strongly connected components	Average clustering coefficient
Full network	2796	5974	0.000764	967	2071	0.001812
2005	99	138	0.014224	27	73	0.027082
2006	70	65	0.013458	23	56	0.026651
2007	181	198	0.006077	66	154	0.019982
2008	330	715	0.006586	81	249	0.008764
2009	317	730	0.007287	99	232	0.008637
2010	228	360	0.006956	79	173	0.011509
2011	240	291	0.005073	101	192	0.008644
2012	184	110	0.003267	102	177	0.006809
2013	256	144	0.002206	139	243	0.004355
2014	229	197	0.003773	105	196	0.004962
2015	241	231	0.003994	101	206	0.00802
2016	212	153	0.003420	101	188	0.010196
2017	370	667	0.004885	138	269	0.006452
2018	192	292	0.007962	60	132	0.001558
2019	281	404	0.005135	90	217	0.001336
2020	407	1158	0.007008	94	244	0.003140
2021	158	423	0.017052	49	91	0.002713

Table 4. Network-level metrics

The type of metadata gathered from the Stormfront website allowed for the creation of directed networks based on information about "quotes," indicating whether or not a post is a response to someone else's post. The quotes, therefore, are represented by directed edges between nodes. Moreover, edges are weighted with a value equal to the number of times user A (the source) quoted user B (the target) during a year.

The number of nodes increases in those years that showed peaks in the number of posts submitted to the forum (see Figure 3), which means that peak activity is explained by more people coming to Stormfront and engaging during those periods rather than a stable, loyal amount of platform users increasing their activity.

The Stormfront network is very loosely connected. As *Table 4* indicates, the density of both the full network and the cross-sections is consistently very low. Density is a measure of how connected the network is by looking at the proportion of existing ties to the maximum number of possible ties given the number of nodes (Barabasi, 2016). Along with the numerous connected components, that is, numerous isolated divisions, groups of nodes connected among themselves but not to the larger network, these measures indicate that the level of engagement with each other's ideas is rather low on Stormfront. It is worth stressing, however, that, as motioned in the methods section, quotes likely underestimate what they measure, namely the level of engagement with exposed ideas and the level of connectivity of the networks. The clustering coefficients, measuring the density of triangles in the networks (Barabasi, 2016; Newman, 2003), are also consistently low, indicating that nodes do not tend to form closely knitted groups or cliques (groups of nodes with all possible ties present).

The second research question asked was: Is the importance of users on Stormfront a good predictor for the use of either of the two frames?

Structural importance of nodes in the network was proposed as a way to ascertain which of the two frames is more likely to be spread throughout the network. The nodes with higher levels of importance in the network, based on how much other nodes engage with their ideas via the quoting feature, are more visible. Their words are being repeated in quotes and they are receiving more answers than others. This way, even though there is no way of measuring exactly how many people read their specific posts, it would make sense that their ideas have a higher exposure rate than the ideas of people who are being quoted less.

For directed graphs, there is a node-level metric that directly measures how much engagement a node gets, which is indegree centrality. While the degree centrality of a node measures the total number of ties connected to that node (Newman, 2003), for directed graphs, degree can be split into indegree (the number of ties for which the node is the target – incoming ties) and outdegree (the number of ties for which the node is the source – outgoing ties). In this case, indegree measures the number of quotes a user gets, and outdegree measures the number of times the user quotes others. Indegree is of particular interest here because it is the most direct measure of how important or how visible the ideas of a user might be, therefore, how users with high indegree frame white supremacy is of great importance.

First off, the full network and the cross sections were visualized separately for each frame. In the Figures 10 to 12, the color of the nodes indicates whether they have scored in a dictionary (red color for the overt dictionary, and green for the covert dictionary), and the size of the nodes is linearly related to their indegree centrality (i.e. how many times they were quoted). Larger nodes indicate having a larger indegree.

What becomes visible right away is that most nodes with larger and medium indegree are colored, either according to the covert or the overt frame. A good part of them may be using both frames. There are fewer nodes with a high indegree which did not register in either of the two dictionaries.

Since the visualizations of the networks can offer only an intuitive answer to the question of the relationship between indegree and the frames of white supremacy, a more precise measure of this relationship was needed.





Figure 10 Full network visualizations for each dictionary. The size of the node = indegree centrality

Yearly Cross-Sections in the Stormfront Network Longitudinal Visualization of Nodes Using Overt Framin



Figure 11 Yearly cross-sections in the overt white supremacy frame. Nodes using this frame are in red.

Yearly Cross-Sections in the Stormfront Network Longitudinal Visualization of Nodes Using Covert Framing of White Supremac



Figure 12 Yearly cross-sections in the covert white supremacy frame. Nodes using this frame are in green.

I opted for regressing the two dictionaries separately on the indegree centrality measure. This would give an indication of the predictive power of the importance of nodes over their practices of framing white supremacy. The representation of the results of these simple OLS linear regression analyses can be found in figures 13 to 15. The regression coefficients, R-squared and other relevant outputs can be found in *Annex 3*, included in the output of the coding notebook.



Figure 13 Regression on the full network



Figure 14 Regression on the network cross sections, y = covert frame



Figure 15 Regression on the network cross sections, y = overt frame

These regression analyses indicate that it is definitely the case that a high indegree is a good predictor of either using a covert or an overt frame when expressing white supremacist ideas. The regression coefficients are not, however, stronger in the case of the covert frame by comparison with the overt frame, and there is no discernible trend in the predictive power of indegree over time.

4.3. Checks

There is a plethora of other measures of the importance of nodes within a network available in the toolbox of SNA. To confirm that indegree centrality is a good indicator for the importance of those who propagate the two frames in the network, two other node centralities were calculated and checked.

Betweenness centrality is an indicator of the presence of so-called bridges. Bridges are nodes which structurally connect parts of the network which would otherwise be disconnected (Everett & Valente, 2016). Considering how fragmentary the Stormfront network is, bridges using one frame or the other are important because they get to control or broker the flow of information from one otherwise disconnected part of the larger network to the other. Similarly to indegree and betweenness, Eigenvector centrality measures how influential a node is in the network, but it is based on the prestige of the node, that is, how well-connected the node is to other well-connected others. This is what Page-Rank centrality is based on, a measure used by internet search engines to rank websites. In this case, Eigenvector centrality would raise the importance of those who quote and are being quoted by other users who are also well quoted.

The two centralities, along with indegree, were correlated with the two dictionaries as shown below, in order to provide a more straightforward way of comparing them.



Figure 16 Yearly cross-sectional correlations between three node-level centrality measures and the two dictionaries

Betweenness centrality and Eigenvector centrality follow a very similar trend to each other, which is not surprising because network centrality measures do tend to be correlated because they are all measuring different aspects of the same phenomenon – the structural importance of nodes in the network. Across all three centralities, the usage of both dictionaries tends to increase as centralities rise, but there is significant yearly variation. Mainly, when the network is more active in general, it has more posts and it is frequented by more users, that is when these centralities become more positively associated with the two dictionaries.

A pertinent question that arises is whether this consistent association between the two frames is actually explained by a poor measurement design of the dictionary-based NLP method, leading to both dictionaries in fact measuring the same frame.

To test this concern, I looked at the overlap between those who use both frames. If the two dictionaries actually measure the same dimension, there should be a high overlap, with those scoring high in one dictionary gaining similar scores in the other. Already from plotting the top ten highest scoring users for each dictionary, there is some evidence that overlap may not be an issue. However, a picture including all the users was needed, not only the most fervent adopters of the two frames.

Among those who have a high dictionary score (over 50 words used during the entire period) in both covert and overt framing, there is an overlap of 23 percent. However, only 1 percent of users have at least one word in both dictionaries. This means that those with very high scores in one dictionary may be only somewhat more inclined to have a high score in the other dictionary as well, but the majority of users tend to stick to one or the other frame, or use none of the two frames. Moreover, among the whole population, 31 percent use more words from the covert frame, 18 percent use more words from the overt frame, and 77 percent use no frame at all.

5. Discussion and Conclusion

5.1. Discussion

The results described in the previous chapter are not in line with the expectations formulated before running the analysis. The first hypothesized expectation was that:

1. A transition from more overt ways of expressing white supremacist ideas earlier in the life of Stormfront to more rhetorically distant ways in recent time is expected;

Contrary to this statement, no significant discrepancy between the usage of ideologically covert and overt white supremacy frames was detected. Over time, they tended to evolve in parallel trends.

This result is at odds with findings from qualitative analyses of speech on Stormfront and demonstrates that, when looking only at a small portion of white supremacist speech online, the conclusions drawn may not even be generalizable to the website from which the analyzed speech was extracted, let alone other websites or social media platforms. For instance, Hartzell (2020) has very recently found that the most important posts on the discussion board, meant to be read by every member and visitor because they contain rules of conduct and give an explanation for the purposes of the website, strongly push for an ideologically covert understanding of white supremacy. These posts use language expressing pride in one's white identity rather than hate of racial others, they strive to create a sense of community, they appeal to common sense, with a cumulative effect of reframing racist ideas in order to make them capable of evoking positive affects Hartzell (2020).

What the present results, based on a much larger portion of speech on Stormfront, found, however, was that not all forum users comply with this ideological project of reframing claimed by the site moderators. Some users (about 23 percent of the selected sample) intensively use language typical for both frames. While it is true that more users prefer the covert frame, there has been no important movement away from direct, abrasive language about race and toward the positive, covert frame. There are some discrepancies between the two frames in years when activity on the website increases, when also more users are present than usual, but as soon as these extra users stop engaging with Stormfront, these discrepancies disappear.

The second expectation was that:

2. The more important users on Stormfront push either the ideologically overt white supremacy frame (more in the distant past) or the ideologically covert frame (more in the recent past and in the present).

Looking at the results of the regression analyses of dictionaries on indegree centrality, there is no identifiable trend of more important users in the network choosing to engage more with the covert frame than the overt frame over time. It can, however, be confirmed that the users who are quoted most by other people do tend to present their ideas in either one ideological frame or the other. Among the possible factors which might contribute to someone being quoted more than others are how old and established the account of the respective user is on Stormfront, whether they are a site moderator, whether they manage to express controversial or engaging ideas. This may explain why structurally central users do fall well within the group that uses one of the two frames.

When it comes to the other aim of the current research project, besides answering specific questions about the ideological framing used on Stormfront, the present approach does not come short. The aim was that of developing or exemplifying a possible workflow for analyzing white supremacist online speech using computational social science methods, starting with a way to extract data and finishing with a mixed-methods design, complementarily using NLP, SNA combined with a social sciences theoretical basis.

While the methods applied are somewhat barebones, and the bag-of-words, unigram, dictionary-based analysis might not be the most precise in detecting complex ideological framing,

they are simple enough to form a starting point for working with big social media data from the social scientists without a background in information technology. The major advantage of this estimate of language use on a white supremacist website is that, despite its roughness and possible margin of error, it is still able to give an idea about what is happening on these websites that has far larger generalizability to the entire online community compared to a qualitative analysis of only a handful of posts which fail to reflect in all fragmentary corners of the network. I would argue that a rougher, but granular enough idea about how people engage with white supremacy over a larger period of time and in different parts of their community is more useful for researchers of online radicalization and policy makers trying to curb it than a highly accurate but narrow, non-generalizable slice from the life of the online community.

Moreover, the project has managed to produce a deliverable that can be used by other social scientists to replicate a similar design on Stormfront or other white supremacist online spaces, in the form of the replication code provided access to in the annexes.

5.2. Limitations and Further Development

As mentioned before, the methods employed here may have some downsides. Most importantly, the dictionary-based analysis used to detect the two frames is not particularly sensitive to context. That is because it relied on unigrams when a good part of the words found in previous literature to express either of the two frames are actually phrases or compounded expressions. Likewise, sometimes, racist dog-whistles, which, in accord with the theoretical distinction between ideological and linguistic dog-whistles, were included in the ideologically overt white supremacist frame, are expressed through symbols, not words. A famous case is that of the so-called "echo," meaning the three parentheses – "(((...)))" – surrounding a person's name which are meant to signal that the name in question belongs to a Jewish person or to mock their being Jewish (Arviv et al., 2020; Bhat & Klein, 2020). Controversially, the developers of the now-defunct *8chan*, another

bastion of online white supremacy, went so far as to code the echo into the website features, allowing an easier and stylistically evident use of it (Hoback, 2021).

Higher precision might be possible by replacing the dictionary-based NLP part of the analysis with a machine learning classification method trained on a set of Stormfront comments classified by human coders. Another approach, still based on a bag-of-words, unigram model but with possibly better accuracy is a deep learning model developed by Alatawi and colleagues (2020), using domain-specific word embeddings. While a model based on human classification would be more sensitive to context, the white supremacy word embeddings model has the advantage of being able to recognize the same concept expressed in different words or in different forms of the same word. Further research or even a further development of the current project should compare the accuracy of several NLP models to bring evidence of which approach would be fittest.

Another limitation of the current design is the simplicity of the quasi-longitudinal SNA based on network cross-sections. Ultimately, this is a static approach comparing snapshots of the network in each arbitrarily determined time-window. There are alternatives of dealing with longitudinal network analysis with potentially better accuracy, but they have the disadvantage of not being accessible to the social scientists with a non-computational background. By comparison, the serial cross-sectional approach was feasible and constitutes a more accessible starting point.

As stressed multiple times throughout this paper, Stormfront is an exceptional online space for far right communities because it is the oldest and the largest around. It is therefore likely that it attracts people holding many gradients of extremist beliefs including those with white supremacist beliefs who also value free speech to such a great extent that they could care less for moderating their speech after the covert framing template pushed by Stormfront moderators. Likewise, despite the largest part of the platform being public, Stormfront can still be thought of as a closed community. People who do not already subscribe to far right white supremacist ideas, have a low chance to find the website by chance, unless they already have a curiosity towards the ideologies they might find on the website. For these reasons, as mentioned before, speech on Stormfront is, theoretically, much more prone to contain the overt frame rather than the covert one, compared to other websites that have far tight communities but also host groups with other ideologies. What this means for the current project is that it only constitutes an initial step in the assessment of the prevalence of the two frames in online white supremacist spaces. A second necessary development is to compare the results obtained on stormfront with those from far right communities on Twitter, Facebook, Reddit or a similar platform which is much more accessible to a larger public and where, potentially, recruitment efforts by far right communities might be intensified.

5.4. Conclusions

The motivation for this research project came from three main sources.

Firstly, over the past decade, there has been a very visible increase in far right-related political violence. The question that has been on the minds of many researchers and policy makers is whether or not there is a connection between offline far-right violence and online communication, coordination, community-making and, most importantly, radicalization. One of the most extreme ideological domains of the far right is that of white supremacy, which has been chosen as the subject of study here.

Secondly, from exposure to white supremacist online content, a puzzle became apparent, namely, if common knowledge about online white supremacist speech is in line with the idea that it is a breeding ground for racialized hate speech, how come much of the content observed argued against expressing hate and was instead simply exalting white identity and constructs of white culture and heritage (examples from these preliminary observations were given in *Figure 1*).

Thirdly, reviewing the social science literature relevant to online radicalization, especially in the fields of political science and sociology, left me with the uneasy idea that the majority of the methods employed, qualitative discourse analysis and NLP on small corpora and case studies, may not entirely be fit for purpose. Or, more accurately, that they are missing study opportunities inherent to large-scale digital-trace data. Correspondingly, the studies using computational methods applied to study the online far right often have only a thin conceptualization and miss building upon social science theory. From these observations, the second motivation was to follow in the steps of those few studies reviewed under subtitle 2.5 ("Social Media-Appropriate Methods of Studying Frames of White Supremacy"), which attempt to be rooted in both social science theory and computational methods fitter for the type of data coming from websites and social media.

The study rests on several central theoretical distinctions.

The guiding theoretical decision was to adopt a minimal definition of white supremacy as any ideology in which racist ideas can be detected, even in cases where the adopters avoid commonsensically recognizable racist speech.

From this follows another concept, that of rhetorical distance as defined by Hartzell (2018, 2020). Through its mechanisms, appeal to common sense, the affective slide from racial hate to self-love, and the alleviation of discomfort associated with being exposed to racist ideas, rhetorical distance leads to a positive re-framing of white supremacy, or, as I have termed it, an ideologically covert framing.

Additionally, a distinction has been drawn between ideologically covert framing and dogwhistles. The former does not only use codes that still express racial hate (like the echo ("(((...)))") or the word "skittles"), it also strives to entirely erase the notion of racial hate from the equation, replacing it with other things, first of which being taking pride in being white.

Methodologically, the study took these theoretical foundations and tried to characterize the dynamic of ideologically overt and covert white supremacy on the oldest online platform for white supremacy, Stormfront. The expectations formulated based on previous findings were not sustained by the results of the analysis, but valuable outputs have been gained from this effort.

The literature focused on the covert ways in which white supremacist ideas are being hidden in online or other types of speech of far right groups and platforms tends to cherry-pick the most interesting cases of rhetorically distant or strategic framing of racism. The accumulation of these cases creates the possibly false impression that rhetorically distant framing in white supremacy is, if not an obvious and popular strategic choice, then, at least, on the rise. This study disputes this impression. The evidence brought here, due to the possibly high margin of error of the frame detection method employed, is not definitive, but it is a rough indication that the dynamic of the two framing options is more complicated. It certainly points towards the idea that strategic framing may not always be desirable for white supremacist groups.

Even if it is only a partial result obtained from a community already prone to speaking openly about white supremacist ideas, the fact that the covert frame is not as prevalent on Stormfront as thought is good news. The functional importance of the covert re-framing of white supremacy is that it is strategic. It hides the violence of the ideology and it promotes it through a positive affective valence. Certainly, if the established users of Stormfront do not overwhelmingly embrace it, it also means that they cannot use the overt frame's power of attraction to recruit new followers.

When compared to qualitative designs, an important conclusion that can be drawn from this effort is that computational methods have more of a chance of producing results which can be generalized, if not to all online white supremacist communities, then at least to a website, something that qualitative methods cannot claim with any measure of certainty.

Additionally, the workflow developed has created visible ways to improve a mixedmethods design like this by enhancing the accuracy of the NLP methods and the finesse of longitudinal network analysis.

Lastly, the project resulted in a valuable deliverable in the form of notebooks with coding solutions which will hopefully be used by other researchers of the far right with a background in social sciences, to augment the kinds of research designs feasible for them and the kinds of research questions that can be asked about online spaces of white supremacy.

List of References

- Ahmed, S. (2004). Affective economies. Social Text, 22(2), 117–139.
- Åkerlund, M. (2021). Dog whistling far-right code words: The case of 'culture enricher' on the Swedish web. *Information, Communication & Society*, 1–18. https://doi.org/10.1080/1369118X.2021.1889639
- Alatawi, H. S., Alhothali, A. M., & Moria, K. M. (2020). Detecting White Supremacist Hate Speech using Domain Specific Word Embedding with Deep Learning and BERT. *ArXiv:2010.00357 [Cs]*. http://arxiv.org/abs/2010.00357
- Arviv, E., Hanouna, S., & Tsur, O. (2020). It's a Thin Line Between Love and Hate: Using the Echo in Modeling Dynamics of Racist Online Communities. *ArXiv:2012.01133 [Cs]*. http://arxiv.org/abs/2012.01133
- Association for Computational Linguistics, Fišer, D., & Association for Computational Linguistics. (2018). *Proceedings of the 2nd Workshop on Abusive Language Online*. ACL. http://aclweb.org/anthology/W18-51
- Auger, V. A. (2020). Right-Wing Terror: A Fifth Global Wave? Perspectives on Terrorism, 14(3), 87– 97.
- Balsamo, D., Gelardi, V., Han, C., Rama, D., Samantray, A., Zucca, C., & Starnini, M. (2019). Inside the Echo Chamber: Disentangling network dynamics from polarization. *ArXiv:1906.09076 [Physics]*. http://arxiv.org/abs/1906.09076
- Barabasi, A.-L. (2016). Network Science (1 edition). Cambridge University Press.
- BBC News. (2021, January 14). Trump impeached for "inciting" US Capitol riot in historic second charge. *BBC News*. https://www.bbc.com/news/world-us-canada-55656385
- Benford, R. D., & Snow, D. A. (2000). Framing Processes and Social Movements: An Overview and Assessment. *Annual Review of Sociology*, 26, 611–639.
- Benford, R. D., & Snow, D. A. (2005). Clarifying the relationship between framing and ideology. Frames of Protest: Social Movements and the Framing Process. Oxford: Rowman and Littlefield, 205– 212.
- Berger, J. M. (2016). Nazis vs. ISIS on Twitter: A Comparative Study of White Nationalist and ISIS Online Social Media Networks (p. 32). George Washington University, Program on Extremism.
- Bhat, P., & Klein, O. (2020). Covert Hate Speech: White Nationalists and Dog Whistle Communication on Twitter. In G. Bouvier & J. E. Rosenbaum (Eds.), *Twitter, the Public Sphere, and the Chaos of Online Deliberation* (pp. 151–172). Springer International Publishing. https://doi.org/10.1007/978-3-030-41421-4_7
- Blee, K. M. (2017). HOW THE STUDY OF WHITE SUPREMACISM IS HELPED AND HINDERED BY SOCIAL MOVEMENT RESEARCH*. *Mobilization: An International Quarterly*, 22(1), 1–15. https://doi.org/10.17813/1086-671X-22-1-1
- Blee, K. M., & Latif, M. (2021). Sociological Survey of the Far Right. In S. D. Ashe, J. Busher, G. Macklin, & A. Winter (Eds.), *Researching the Far Right: Theory, Method and Practice* (pp. 45–59). Routledge.

- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 993–1022.
- Bouchard, M., Joffres, K., & Frank, R. (2013). Preliminary Analytical Considerations In Designing A Terrorism And Extremism Online Network Extractor. In V. K. Mago & V. Dabbaghian (Eds.), *Computational Models of Complex Systems* (pp. 171–184). Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-319-01285-8_11
- Bowman-Grieve, L. (2009). Exploring "Stormfront": A Virtual Community of the Radical Right. *Studies in Conflict & Terrorism*, 32(11), 989–1007. https://doi.org/10.1080/10576100903259951
- Brown, C. (2009). WWW.HATE.COM: White Supremacist Discourse on the Internet and the Construction of Whiteness Ideology. *Howard Journal of Communications*, 20(2), 189–208. https://doi.org/10.1080/10646170902869544
- Browning, K. (2020, November 17). Zuckerberg and Dorsey Face Harsh Questioning From Lawmakers. *The New York Times*.

https://www.nytimes.com/live/2020/11/17/technology/twitter-facebook-hearings

Cabrera, N. L. (2014). "But I'm oppressed too": White male college students framing racial emotions as facts and recreating racism. *International Journal of Qualitative Studies in Education*, 27(6), 768–784. https://doi.org/10.1080/09518398.2014.901574

- Cai, W., & Landon, S. (2019, April 3). Attacks by White Extremists Are Growing. So Are Their Connections. (Published 2019). *The New York Times*. https://www.nytimes.com/interactive/2019/04/03/world/white-extremist-terrorismchristchurch.html, https://www.nytimes.com/interactive/2019/04/03/world/whiteextremist-terrorism-christchurch.html
- Caren, N., Jowers, K., & Gaby, S. (2012). A Social Movement Online Community: Stormfront and the White Nationalist Movement. *Research in Social Movements, Conflicts and Change*, 33, 163–193. https://doi.org/10.1108/S0163-786X(2012)0000033010
- Chau, M., & Xu, J. (2007). Mining communities and their relationships in blogs: A study of online hate groups. *International Journal of Human-Computer Studies*, 65(1), 57–70. https://doi.org/10.1016/j.ijhcs.2006.08.009
- Clayton, J. (2020, June 29). Could a boycott kill Facebook? *BBC News*. https://www.bbc.com/news/technology-53225139
- CSIS. (2020). The Escalating Terrorism Problem in the United States. https://www.csis.org/analysis/escalating-terrorism-problem-united-states
- de Gibert, O., Perez, N., García-Pablos, A., & Cuadros, M. (2018). Hate Speech Dataset from a White Supremacy Forum. *ArXiv:1809.04444 [Cs]*. http://arxiv.org/abs/1809.04444
- De Koster, W., & Houtman, D. (2008). 'STORMFRONT IS LIKE A SECOND HOME TO ME': On virtual community formation by right-wing extremists. *Information, Communication & Society*, 11(8), 1155–1176. https://doi.org/10.1080/13691180802266665
- Drainville, R., & Saul, J. (2020). Visual and Linguistic Dogwhistles. In Oxford Handbook of Applied Philosophy of Language, upcoming.
- Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x
- Everett, M. G., & Valente, T. W. (2016). Bridging, brokerage and betweenness. *Social Networks*, 44, 202–208. https://doi.org/10.1016/j.socnet.2015.09.001

- Fazio, M. (2021, January 10). Notable Arrests After the Riot at the Capitol. *The New York Times*. https://www.nytimes.com/2021/01/10/us/politics/capitol-arrests.html
- Frenkel, S. (2021, January 6). The storming of Capitol Hill was organized on social media. The New York Times. https://www.nytimes.com/2021/01/06/us/politics/protesters-stormcapitol-hill-building.html
- Gerstenfeld, P. B., Grant, D. R., & Chiang, C.-P. (2003). Hate Online: A Content Analysis of Extremist Internet Sites. 18.
- Guenther, L., Ruhrmann, G., Bischoff, J., Penzel, T., & Weber, A. (2020). Strategic Framing and Social Media Engagement: Analyzing Memes Posted by the German Identitarian Movement on Facebook. *Social Media* + *Society*, 6(1), 2056305119898777. https://doi.org/10.1177/2056305119898777
- Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. *Proceedings of the Python in Science Conference (SciPy)*, 11–15. http://conference.scipy.org/proceedings/SciPy2008/paper_2/
- Hankes, K., & Zhang, S. (n.d.). A Waning Storm: Once the World's Most Popular White Nationalist Website, Stormfront is Running out of Steam. Southern Poverty Law Center. Retrieved April 28, 2021, from https://www.splcenter.org/hatewatch/2017/02/22/waning-storm-onceworld%E2%80%99s-most-popular-white-nationalist-website-stormfront-running-out
- Hartzell, S. L. (2018). Alt-White: Conceptualizing the "Alt-Right" as a Rhetorical Bridge between White Nationalism and Mainstream Public Discourse. 20.
- Hartzell, S. L. (2020). Whiteness feels good here: Interrogating white nationalist rhetoric on Stormfront. 21.
- Healy, J. (2021, January 11). These Are the 5 People Who Died in the Capitol Riot. *The New York Times*. https://www.nytimes.com/2021/01/11/us/who-died-in-capitol-building-attack.html
- Hoback, C. (2021). Q: Into the Storm. HBO. https://www.imdb.com/title/tt14215442/
- Isaac, M., & Browning, K. (2020, November 11). Fact-Checked on Facebook and Twitter, Conservatives Switch Their Apps. *The New York Times*.
- https://www.nytimes.com/2020/11/11/technology/parler-rumble-newsmax.html Jackson, S. (2019). "Nullification through armed civil disobedience": A case study of strategic framing in the patriot/militia movement. *Dynamics of Asymmetric Conflict*, *12*(1), 90–109.
 - https://doi.org/10.1080/17467586.2018.1563904
- Jasper, J. M. (2017). The doors that culture opened: Parallels between social movement studies and social psychology. *Group Processes & Intergroup Relations*, 20(3), 285–302. https://doi.org/10.1177/1368430216686405
- Jasper, M., Klein, O., & Dijkstra, G. (2021). Challenges and opportunities of social media research: Using Twitter and Facebook to investigate far right discourses. In S. D. Ashe, J. Busher, G. Macklin, & A. Winter (Eds.), Researching the Far Right: Theory, Method and Practice (pp. 147–163). Routledge.
- Key, V. O. (1966). The Responsible Electorate. Harvard University Press.
- Kim, T. K. (2005). Hate Website Stormfront Sees Rapid Growth of Neo-Nazi Community. Southern Poverty Law Center. https://www.splcenter.org/fighting-hate/intelligencereport/2005/hate-website-stormfront-sees-rapid-growth-neo-nazi-community
- Klandermans, B., & Mayer, N. (2006). Right-wing Extremism as a Social Movement. In *Extreme right activists in Europe: Through the magnifying glass*. Routledge. https://ceuedu.sharepoint.com/sites/itservices/SitePages/vpn.aspx

- Knüpfer, C., Hoffmann, M., & Voskresenskii, V. (2020). Hijacking "MeToo": Transnational dynamics and networked frame contestation on the far right in the case of the '120 decibels' campaign. *Information, Communication & Society*, 1–19. https://doi.org/10.1080/1369118X.2020.1822904
- Kruglanski, A. W., Gelfand, M. J., Bélanger, J. J., Sheveland, A., Hetiarachchi, M., & Gunaratna, R. (2014). The Psychology of Radicalization and Deradicalization: How Significance Quest Impacts Violent Extremism. *Political Psychology*, 35(S1), 69–93. https://doi.org/10.1111/pops.12163
- López, I. H. (2015). Dog Whistle Politics: How Coded Racial Appeals Have Reinvented Racism and Wrecked the Middle Class. Oxford University Press.
- Mamié, R., Ribeiro, M. H., & West, R. (2021). Are Anti-Feminist Communities Gateways to the Far Right? Evidence from Reddit and YouTube. *ArXiv:2102.12837 [Cs]*. http://arxiv.org/abs/2102.12837
- Marantz, A. (2019). Antisocial: Online extremists, techno-utopians, and the hijacking of the American conversation. Penguin Random House.
- https://www.penguinrandomhouse.com/books/562022/antisocial-by-andrew-marantz/ McIlroy-Young, R., & Anderson, A. (2019). From Welcome New Gabbers to the Pittsburgh Synagogue Shooting: The Evolution of Gab. *ArXiv:1912.11278 [Cs]*.

http://arxiv.org/abs/1912.11278

- McNeil-Willson, R. (2020). Framing in times of crisis: Responses to COVID-19 amongst Far Right movements and organisations. International Centre for Counter-Terrorism. https://www.jstor.org/stable/resrep25256
- McNeil-Willson, R., Gerrand, V., Scrinzi, F., & Triandafyllidou, A. (2019). Polarisation, violent extremism and resilience in Europe today: An analytical framework. https://ceuedu.sharepoint.com/sites/itservices/SitePages/vpn.aspx
- Meraz, S., & Papacharissi, Z. (2013). Networked Gatekeeping and Networked Framing on #Egypt. *The International Journal of Press/Politics*, *18*(2), 138–166. https://doi.org/10.1177/1940161212474472
- Miller-Idriss, C. (2020). Hate in the Homeland: The New Global Far Right. Princeton University Press.
- Newman, M. E. J. (2003). The structure and function of complex networks. *SLAM Review*, 45(2), 167–256. https://doi.org/10.1137/S003614450342480
- Oliver, P., & Johnston, H. (2006). What a Good Idea! Ideologies and Frames in Social Movement Research. *Mobilization: An International Quarterly*, *5*(1), 37–54. https://doi.org/10.17813/maiq.5.1.g54k222086346251

Panizo-LLedot, Á., Torregrosa, J., Bello-Orgaz, G., Thorburn, J., & Camacho, D. (2019). Describing Alt-Right communities and their discourse on Twitter during the 2018 US mid-term elections. ArXiv:1910.03431 [Cs]. http://arxiv.org/abs/1910.03431

Peita, L. R. (2021). Big Data and the Resurgence of the Far Right Within the United States. In S.
D. Ashe, J. Busher, G. Macklin, & A. Winter (Eds.), *Researching the Far Right: Theory, Method and Practice* (pp. 164–177). Routledge.

Phillips, J., & Yi, J. (2018). Charlottesville Paradox: The 'Liberalizing' Alt-Right, 'Authoritarian' Left, and Politics of Dialogue. *Society*, 55(3), 221–228. https://doi.org/10.1007/s12115-018-0243-0

- Radicioni, T., Squartini, T., Pavan, E., & Saracco, F. (2021). Networked partisanship and framing: A socio-semantic network analysis of the Italian debate on migration. *ArXiv:2103.04653 [Physics]*. http://arxiv.org/abs/2103.04653
- Rauchfleisch, A., & Kaiser, J. (2020). The German Far-right on YouTube: An Analysis of User Overlap and User Comments. *Journal of Broadcasting & Electronic Media*, 64(3), 373–396. https://doi.org/10.1080/08838151.2020.1799690
- Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A. F., & Meira, W. (2019). Auditing Radicalization Pathways on YouTube. *ArXiv:1908.08313 [Cs]*. http://arxiv.org/abs/1908.08313
- Rodríguez, D. (2017). White Supremacy. In *The Wiley-Blackwell Encyclopedia of Social Theory* (pp. 1–3). American Cancer Society. https://doi.org/10.1002/9781118430873.est0407
- Rogers, R. (2020). Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. *European Journal of Communication*, 35(3), 213–229. https://doi.org/10.1177/0267323120922066
- Sanchez, J. C. (2018). Trump, the KKK, and the Versatility of White Supremacy Rhetoric. 13.

Saresma, T. (2020). Circulating the Origin Myth of Western Civilization – The Racial Imagery of the 'Men of the North' as an Imaginary Heritage in White Supremacist Blogs. In D. E. Morse, Z. O. Réti, & M. Takács (Eds.), *The (Web)Sites of Memory: Cultural Heritage in the Digital Age* (pp. 68–81). Debrecen University Press. https://web.archive.org/web/20200118234638/https://s3.amazonaws.com/academia.e du.documents/61195901/Loci VIII20191112-53236-1u8vi04.pdf?response-content-

- disposition=inline%3B%20filename%3DLoci_Memoriae_Hungaricae_VIII.pdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-
- Credential=AKIAIWOWYYGZ2Y53UL3A%2F20200118%2Fus-east-
- 1%2Fs3%2Faws4_request&X-Amz-Date=20200118T234524Z&X-Amz-

Expires=3600&X-Amz-SignedHeaders=host&X-Amz-

Signature=aa6873fdb9948643eca7848adb3b3fabb42dc8af8d8df5f2e404b641db336d70

- Snow, D., Benford, R., McCammon, H., Hewitt, L., & Fitzgerald, S. (2014). The Emergence, Development, and Future of the Framing Perspective: 25+ Years Since "Frame Alignment." *Mobilization: An International Quarterly*, 19(1), 23–46. https://doi.org/10.17813/maiq.19.1.x74278226830m691
- Southern Poverty Law Center. (2020a). *Profile: Stormfront*. Southern Poverty Law Center. https://www.splcenter.org/fighting-hate/extremist-files/group/stormfront
- Southern Poverty Law Center. (2020b, April 23). 25 years later, Oklahoma City bombing still inspires antigovernment extremists. https://www.youtube.com/watch?v=YBHCLWZMJ_A&t=20s

Spencer. (2016). Euro-Skepticism Skepticism archived at Wayback Machine.

^{https://web.archive.org/web/20170119013901/http://disqusads.com/ads}iframe/prebid/?category=news&stories_allowed=1&service=dynamic&safetylevel=30& display_allowed=1&video_allowed=0&aol=0&thumbnails_allowed=1&provider=prebid &experiment=prebid_lift&variant=with_prebid&forum_pk=2675125&rubicon=0&ope nx=1&links_allowed=1&position=top&shortname=radixjournal&display_bidding_allow ed=1&forum_shortname=radixjournal&display_only=0&t=1484789939&anchorColor= %23990000&colorScheme=light&sourceUrl=https%3A%2F%2Fweb.archive.org%2Fwe b%2F20170312063232%2Fhttp%3A%2F%2Fwww.radixjournal.com%2Fblog%2F2016 %2F5%2F25%2Fmake-europe-great-again&typeface=sans-

serif&canonicalUrl=http%3A%2F%2Fwww.radixjournal.com%2Fblog%2F2016%2F5% 2F25%2Fmake-europe-great-again&disqus_version=d85161d

- Stewart, L. G., Arif, A., Nied, A. C., Spiro, E. S., & Starbird, K. (2017). Drawing the Lines of Contention: Networked Frame Contests Within #BlackLivesMatter Discourse. *Proceedings* of the ACM on Human-Computer Interaction, 1(CSCW), 1–23. https://doi.org/10.1145/3134920
- Stormfront Home Page. (2021). *Stormfront—White Nationalist Community*. https://www.stormfront.org/forum/
- Than, N., Rodriguez, M. Y., Yoong, D., & Windel, F. (2020). Welcome to Gab Alt Right Discourses. *ArXiv:2007.09685 [Cs]*. http://arxiv.org/abs/2007.09685
- The Independent. (2020). *Search: Christchurch attack*. The Independent. https://www.independent.co.uk/topic/christchurch-attack
- Thompson, K. C. (2001). WATCHING THE STORMFRONT: White Nationalists and the Building of Community in Cyberspace. *Social Analysis: The International Journal of Social and Cultural Practice*, 45(1), 32–52.
- Thorburn, J., Torregrosa, J., & Panizo, Á. (2018). Measuring Extremism: Validating an Alt-Right Twitter Accounts Dataset. Intelligent Data Engineering and Automated Learning – IDEAL 2018, 19th International Conference, Madrid, Spain, November 21–23, 2018, Proceedings, Part II. https://www.springerprofessional.de/en/measuring-extremism-validating-an-alt-righttwitter-accounts-dat/16260480
- Tufekci, Z. (2014). Social Movements and Governments in the Digital Age: Evaluating a Complex Landscape. *Journal of International Affairs*, 68(1), 1–18.
- Uddin, S., Choudhury, N., Farhad, S. M., & Rahman, M. T. (2017). The optimal window size for analysing longitudinal networks. *Scientific Reports*, 7(1), 13389. https://doi.org/10.1038/s41598-017-13640-5
- van der Vegt, I., Mozes, M., Gill, P., & Kleinberg, B. (2020). Online influence, offline violence: Language Use on YouTube surrounding the "Unite the Right" rally. *ArXiv:1908.11599 [Cs]*. http://arxiv.org/abs/1908.11599
- Wahlström, M., Törnberg, A., & Ekbrand, H. (2020a). Dynamics of violent and dehumanizing rhetoric in far-right social media. New Media & Society, 1461444820952795. https://doi.org/10.1177/1461444820952795
- Wahlström, M., Törnberg, A., & Ekbrand, H. (2020b). Dynamics of violent and dehumanizing rhetoric in far-right social media. New Media & Society, 146144482095279. https://doi.org/10.1177/1461444820952795
- WayBackMachine Internet Archive. (2008, December 28). Stormfront White Nationalist Community—Discussion Board for Activists.

https://web.archive.org/web/20081228233552/http://www.stormfront.org/forum/

- Wetts, R., & Willer, R. (2019). Who Is Called by the Dog Whistle? Experimental Evidence That Racial Resentment and Political Ideology Condition Responses to Racially Encoded Messages. Socius, 5, 2378023119866268. https://doi.org/10.1177/2378023119866268
- Wright, S. A. (2009). Strategic Framing of Racial-Nationalism in North America and Europe: An Analysis of a Burgeoning Transnational Network. *Terrorism and Political Violence*, 21(2), 189–210. https://doi.org/10.1080/09546550802544565

Appendices

Appendix 1. Replication code for data scraping and cleaning (Python)

Please find the notebook at this link:

https://drive.google.com/file/d/1NNvlK1wrdVUyxgc8Z8SrKEMAywxE2dF7/view?usp=sharing

Please note that this is an html file. Download the file and open it locally. Opening it in Google Drive might not display the file correctly.

Appendix 2. Replication code for NLP (R)

Please find the notebook at this link:

https://drive.google.com/file/d/1Xd8CsYZVDBsD_47DryVVyRVqBH8Xr3nz/view?usp=sha ring

Please note that this is an html file. Download the file and open it locally. Opening it in Google Drive might not display the file correctly.
Appendix 3. Replication code for Social Network Analysis (Python)

Please find the notebook at this link:

https://drive.google.com/file/d/1x7ayXjN2nBA 77m zdB9PHmIvQcYDpJB/view?usp=shari

<u>ng</u>

Please note that this is an html file. Download the file and open it locally. Opening it in Google

Drive might not display the file correctly.