Journal of Right-Wing Studies, Vol. 2, No. 1 (2024)
http://jrws.berkeley.edu

ARTICLE

The Gab Project

The Methodological, Epistemological, and Legal Challenges of Studying the Platformized Far Right

Tim de Winkel
University of Applied Sciences Utrecht, Erasmus University

Ludo Gorzeman
Utrecht University

Sofie de Wilde de Ligny
Utrecht University

Thomas ten Heuvel
ING

Melissa Blekkenhorst
Ministry of the Interior and Kingdom Relations

Sander Prins
Utrecht University

Mirko Tobias Schäfer
Utrecht University

Abstract: *In this article we describe our five-year research project on the notorious radical free speech service and fringe platform Gab. During these years we scraped an entire platform, prepared it into a dataset for analysis, and opened it up to a broader community of students and researchers. Each of these projects provides us not just with a small slice of platformized far-right culture but also with a larger sphere of a fringe platform. However, the overarching goal of the Gab project was to contribute to a methodology for the study of the contemporary platformized far right. The atypical nature of the project posed many methodological, epistemological, and legal challenges. It therefore kicked off an institutional learning process about the possibilities,*

legal boundaries, and best practices for research compliant with General Data Protection Regulation (GDPR). In this article we argue that the study of the platformized far right should have a thorough understanding of the medium on which the object is present, as well as the methods with which the object is captured. What is more, scholars that use digital tools and data methods for capture and analysis of web platforms must become literate in operating them. Consequently, data-driven research on the far right is naturally interdisciplinary and therefore cooperative and adherent to the principles of open science.

Keywords: fringe platforms, platformized public sphere, digital humanities, alt-right, critical data studies, Gab

Increasingly, radical free speech social media platforms have come under scrutiny for the societal impact of the far-right and alt-right publics they host.[1] Services such as 4chan, 8chan, Parler, and Kiwifarms have been implicated in acts of offline hate, harassment, or worse.[2] This is also true for radical free speech social media platform Gab. At the Tree of Life synagogue in Pittsburgh on October 27, 2018, during the morning Shabbat service, a domestic white supremacist terrorist opened fire and killed eleven people. The police found that the perpetrator had a Gab profile, which he used to disseminate his neo-Nazi views and where he announced his attack through a now infamous post condemning the Hebrew Immigrant Aid Society (HIAS): "HIAS likes to bring invaders in that kill our people. I can't sit by and watch my people get slaughtered. Screw your optics, I'm going in" (Hutchinson et al. 2018; Pagliery and Toropin 2018; Bagavathi et al. 2019; de Winkel 2023, 12).[3] The FBI raided the homes of Gab users who were in contact with the Pittsburgh killer and found heavy weaponry at one of them (Sommer 2018; Weill 2019). Due to the tragic attack, the possible dangers of radical free speech

1 We use "far right" as an umbrella term to describe the right that is "anti-system" or hostile to liberal democracy. "Alt-right" is an abbreviation of "alternative right." The term is a denominator for the loose affiliation of different online far-right groups, spawn from internet troll culture, with a common style, common tactics, and a common political identity. Most "mainstream" social media are also used to incite offline violence and hate; however, "radical free speech" platforms practice a governance style of inaction, condoning hate speech and political radicalism as part of their identity (de Winkel 2023, 77, 85–87).

2 In 2014, 4chan circulated celebrity nude pictures that were harvested in an iCloud hack (Patrick 2022). 8chan has been involved in plenty of scandals, including episodes involving death threats, doxing, political violence and terrorism, and child pornography. The message board gained notoriety through the infamous Gamergate harassment campaign (McLaughlin 2019; Rieger et al. 2021). Parler was used for the dissemination of the "Stop the Steal" conspiracy that led to the January 6, 2021, attack on the United States Capitol (Otala et al. 2021). Kiwifarms has been implicated in the stalking, doxing, and threat campaigns mounted against its critics and trans people, ostensibly with the aim of moving their victims to commit suicide (Breland 2023; Hern 2022).

3 As explained in "Fringe Platforms" (de Winkel 2023, 12, 63), HIAS provides humanitarian aid and assistance to refugees, which triggers in white supremacists the hoary antisemitic trope that Jewish citizens are diluting the whiteness of a nation through immigration.

platforms such as Gab took hold in public consciousness. Cries for regulation ensued, and moderation of the larger Web was amped up. Journalists reframed their reporting about Gab; once merely a platform for the deplatformed, Gab was now regarded as an echo chamber of the alt-right. Two contrasting narratives emerged about the role of Gab in the online public sphere—either it had a facilitating role as an ecosystem for discourse, or it played a more active role as a radical(izing) technology. These narratives echo the questions raised by political scientists and media scholars on how to study the platformized far right.

With the transition to a platformized public sphere (de Winkel 2023), the rise of a body of technologically and methodologically savvy humanities scholarship (e.g., the digital humanities, critical data studies, platform studies) has become ever more valuable. The possibilities offered by these new fields to scholarship on the far right were quickly embraced, not least because in the last ten years its object of study has been a moving target. The rise of the alt-right signified the far right's successful establishment on the contemporary Web. Initially inhabiting mainstream social media, these alt-right publics were compelled, in the wake of several waves of deplatforming, to search for their own services, eventually regrouping as the "Alt-Tech movement" (de Winkel 2023, 75–76). This development is best described as the infrastructural turn of the alt-right (Donovan, Lewis, and Friedberg 2019). The platformized public sphere thus presents scholars of the far right with new publics, new cultures, a new vernacular, and new technology—that is, with a new object of analysis "shaped and remediated by pre-existent imaginaries,"[4] both the same as and different from previous incarnations of the far right. The technological newness of this datafied object in a platformized environment forces scholars to apply new methods and to view the phenomenon through new lenses.

In this article we describe our efforts, results, and experiences performing data-driven research on the platformized far right as humanities scholars. Our case study is the aforementioned radical platform Gab. For this reason, the many research projects conducted on the Gab dataset over the past five years are encapsulated under the moniker "the Gab project." We will tell the story of the Gab project in the following order. In the first section we introduce and elaborate on the Gab platform. In the second section we describe the onset of the Gab project, the data capture, and the many subprojects performed on our dataset (this section includes all our empirical work and results). In the third section we describe the legal and epistemological challenges of data-driven humanities research through the narrative of our own trajectory. And finally, in the fourth section, we present our conclusions on Gab and propose several methodological guidelines for the study of the platformized far right. Through this article we hope

4 This is the description by the KNIR's Dr. Bonaria Urban of populist and nationalist collective identity. See "We Are the People: Transnational Imaginaries of (Anti-)Fascism and Populism," Koninklijk Nederlands Instituut Rome (KNIR), accessed June 3, 2024, https://www.knir.it/nl/cursus/we-are-the-people.

not only to do justice to our empirical work but also to disclose the methodological, epistemological, and legal journey entailed by this data-driven humanities research, a type of research that, at the time it was conducted, was unconventional and therefore uncomfortable given the structures available for faculty researchers.

The Platform Gab

Gab is a social networking site modeled after Twitter,[5] but parts of its design and affordances resemble those of the social news website and forum Reddit, known as the "front page of the Internet." Gab as a social medium thus aspires to be a networking site and news aggregator. The platform's core features, similar to those of most of its social media counterparts, include the ability to create user profiles, follow other users, participate in groups, and post or send text messages (Jasser et al. 2021). Additionally, Gab offers a news hub (Gab Trends), a blog (Gab News), a YouTube-like video platform (Gab TV), an online marketplace, its own pay service (GabPay), paid premium accounts (Gab Pro), a web browser and browser extension (Dissenter—a fork of Brave), and an online shop (Dissenter Shop) (de Wilde de Ligny 2022). These features exemplify Alt-Tech's striving to create its own web infrastructures, separate from those of Big Tech.

Launched in 2016, Gab soon became one of the most notorious social media platforms due to the abundance of hate speech and political radicalism present on the service, as well as the platform's public quarrel with Silicon Valley elites (de Winkel 2023, 18, 61–62). After Gab was implicated in the Pittsburgh synagogue shooting in 2018 it had to suspend operation for several days because many of its partner services cut ties. A seemingly independent platform service was thus pushed offline. This process, called "deplatformization," entails a systemic effort by Big Tech companies to push radical and controversial platforms—and their communities—to the fringes of the platform ecosystem by denying them access to basic infrastructural services needed to function online (Van Dijck, de Winkel, and Schäfer 2021, 2). Gab resurfaced a week later and eventually switched its software infrastructure to run on a forked version of Mastodon in 2019 (Van Dijck, de Winkel, and Schäfer 2021, 2024).

Like many alternative free speech platforms, Gab moderates what its users post to a far lesser extent than mainstream platforms, leading to the increased presence of far-right publics and content (Zannettou et al. 2018) Even so, the service rejects being characterized as an alt-right Twitter and denies that it has any political affiliation. Rather, Gab initially framed itself as a platform for the deplatformed and subsequently as a Christian nationalist platform.[6] However, academics (Zannettou et al. 2018; Zhou et al.

---

5 Since 2023, Twitter has been known officially as X, but we will refer to the platform as Twitter in what follows since all analyses predate the name change.

6 The self-proclaimed Christian nationalism of Gab is a form of badly concealed Christian fundamentalism and antisemitism, as shown by the following three quotes from Makuch (2019): 1) "We don't

2019; Munn 2019; Zeng and Schäfer 2021), journalists (Makuch 2019; Dougherty and Hayden 2019), and intelligence agencies (Nationaal Coördinator Terrorismebestrijding en Veiligheid 2018; Department of Homeland Security 2019, 10) alike describe Gab as a far-right platform. Research shows that Gab hosts far-right celebrities and publics, including explicit Nazis and white supremacists, and that it has a high percentage of hate speech (Zannettou et al. 2018; Zhou et al. 2019; Makuch 2019). Additionally, the communications of Gab CEO and founder Andrew Torba have involved plenty of white supremacist and antisemitic rhetoric, as well as amplification of right-wing conspiracies (Dickson and Wilson 2018). Moreover, Gab almost exclusively partners with far-right and conspiracy channels, figures, and services. The disparity between its political extremism and expressed self-image is likely a matter of deception. Typically for the alt-right, Gab engages in the ludic and tactical practice of signaling its own far-right ideology through dog whistles,[7] which grants it a veil of plausible deniability regarding incendiary claims while reaching its primary audience and attracting attention from journalists.[8]

Despite Gab's obvious political radicalism, a focus solely on its toxicity and far-right politics would yield a limited view of this sort of space. Work that engages with the platformized far right as part of the larger dynamics in the public sphere (Donovan 2019; Donovan, Lewis, and Friedberg 2019; Figenschou and Ihlebæk 2019; Van Dijck, de Winkel, and Schäfer 2021, 2024) reveals aspects of far-right platforms the usual lens does not. Analyses of communities, content, (hate) speech, and influencers engage with the far right at the level of the platform, while work that analyzes the far right at the level of a platform ecology instead emphasizes infrastructures, political economy, power structures, and ecosystems. For these reasons, de Winkel (2023) chose to analyze Gab as an antagonistic or contentious platform, conceptualizing it as a "fringe platform." In the context of technology platforms, "fringe" characterizes technologies on the far reaches

---

want people who are atheists. We don't want people who are Jewish. We don't want people who are, you know, nonbelievers, agnostic, whatever. This is an explicitly Christian movement because this is an explicitly Christian country." 2) "Our generation of Christians is not buying dispensational Zionist lies, we do not have a pessimistic eschatology." 3) "They're trying to subvert Christian nationalism. Turn off Ben Shapiro. This is not a Judeo-Christian movement. Those two terms are actually contradictory."

7 A dog whistle is a style of multivocal communication (Albertson 2015) where the full message is only heard by those for whom it is meant, with the rest of the audience remaining oblivious to the undertones. Alt-right dog whistles often conceal antisemitic or otherwise racist remarks or references to fascist theory.

8 A good example of this practice is Gab's former logo, a frog called "Gabby." This image dog-whistles the alt-right figure of Pepe the Frog, which has been declared a hate symbol by the Anti-Defamation League. However, every time Gab executives are confronted with the *alleged* meaning of the logo, they come up with a different meaning for the symbol. One particularly outrageous explanation claims that Gabby is a reference to the plagues in the story of Moses. Such "trolling" attracts attention from journalists, who will feel the urge to find out the symbol's true meaning, while simultaneously creating confusion among mainstream audiences and a feeling of superiority among the alt-right publics who get the joke.

of the Web, ideologically, infrastructurally, and in terms of power. Fringe platforms are defined as

> alternative platform services that were established as an explicit critique of the ideological premises and practices of mainstream platform services [and] that attempt to cause a shift in the norms of the platform ecology they contest by offering an ideologically different technology. (de Winkel 2023, 35)

Many fringe platforms are radical free speech spaces that facilitate right-wing extremism, much like Gab does, but the terminology "fringe platforms" is not reserved for any one specific political affiliation. What distinguishes fringe platforms is their explicit contestation vis-à-vis their mainstream counterparts. Fringe platforms are founded to challenge the norms of mainstream social media platforms by offering a space with different norms (de Winkel 2023, 18). As we will see in the next section, research that addresses these far-right services in the context of the broader platform ecology allows for interpretations and lenses that extend beyond the borders of their discursive spheres.

The Gab Project

In November 2018, Data School, a research and education platform of Utrecht University,[9] started a research project on Gab.ai. Under the supervision of Dr. Mirko Tobias Schäfer, a team led by Tim de Winkel initiated the project "Data-Mining Hate-Speech, or Analyzing Speech, Images and Interactions on Gab.ai, the 'Twitter of the Alt Right,'" henceforth called "the Gab project." We were awarded a grant in the focus area of Applied Data Science,[10] the main portion of which was used to hire an external expert for the necessary data capture, involving the scraping of the entire Gab platform's 17 million posts and 770,000 profiles.

After the deliverables of the grant were met, we decided to prolong the project, seeking 1) to research the far-right social media platform Gab through several different projects, and 2) to understand what it takes, as data-savvy humanities scholars, to research far-right media services within the contemporary platform ecology. We opened up our data, infrastructure, and expertise to staff and students who wanted to contribute to the Gab project and who possessed their own questions and research projects. Ultimately, the project ran for almost five years, and it will be finalized with the publication of this article. The Gab project has contributed to a dissertation (de Winkel 2023), several

9 Data School was formerly known as Utrecht Data School (UDS).

10 Specifically, we worked in the special interest group (SIG) Text Mining led by Dr. Pim Huijnen and Jan de Boer. See the Utrecht University Applied Data Science website at https://www.uu.nl/en/research/applied-data-science/research/research-grants.

masters theses and internships,[11] a journalistic article in a Dutch weekly (Boeschoten and Van den Ven 2019), and several presentations at academic conferences (Blekkenhorst et al. 2019; Salazar and de Winkel 2019; de Winkel et al. 2019; Gorzeman 2019). In this section we explain our process of data capture, all the subprojects performed on the dataset, which includes the bulk of our empirical work and data analyses, and some theoretical contributions. Many of the subprojects, including the analyses and results, have not yet been published as written texts.

Scraping the Platform

The scraping process met with both institutional and technical challenges. Despite those hurdles our primary goals were reuse, accessibility, and ease of usage. We wrote our toolkit in Python due to its popularity within the digital humanities, thus making it easier for colleagues within the field to understand, deploy, and adapt our tools. For similar reasons, in particular the ease of setup, we opted for MongoDB as the database to store the scraping results. Our dataset—too large to process on a desktop computer, too small to justify supercomputers—required renting computing infrastructure from a Dutch commercial provider. The initial plan was to use an existing scraper for the Gab platform; however, due to the interactive nature of the Gab website, regular web scraping techniques would not suffice.[12] Instead, we resorted to reverse-engineering the—undocumented—backend application programming interface (API) used by the Gab website to retrieve its data. By the time we discovered which relevant functions were accessible, what data they provided, and, perhaps most importantly, what rate limiting was implemented on retrieving said data, we had already written such substantial amounts of code for reverse-engineering and testing purposes that it was easier for us to reuse this code to build a scraper from scratch ourselves. The gabber toolkit was born.

The scraper was designed to work as follows. It would be fed a point of entry—a profile—and crawl to all connections this node had made to other nodes—users, groups, posts—making those nodes the new points of entry. The scraping process is completed once the entire network is scraped and no new nodes are found. Since the structure of a network allows for a "circular" data crawl, the scraper was programmed to compare each new data entry to the already existing dataset in order to prevent "doubles." If a data entry already existed, it was not written into the dataset and was therefore not used—again—as a new point of entry. After roughly two weeks of scraping, performing many iterations, we found ourselves in possession of a database filled with every post

11 See "Theses" in section 3.

12 Simply stated, existing public scrapers would collect all that was visible on the page, but, as we know, social media only show a fraction of their posts. After a certain number of posts—for example, ten—a user has to click "load more" to show another ten "Gabs." A scraper based on the backend API of the platform itself does not have this problem.

and profile of every non-isolated user and group publicly available on Gab. The question then became how to make sense out of these millions of database records.

After some harsh lessons in the importance of proper database indexing, several tools were written to analyze and enrich the dataset as well as to extract specific data. In particular, community detection algorithms were run to identify and export clusters based on following, repost, comment, and quote relationships. HateSonar,[13] a multiclass classifier based on crowdsourced data (Davidson et al. 2017), was used to detect hate speech and add metadata with classification and confidence to all posts; LDA (Latent Dirichlet Allocation) topic modeling was applied to all identified clusters; and hashtags were extracted and counted both globally and per cluster. These operations provided an overview of the different communities within Gab and laid the foundation for further research. Additionally, we performed activity counts and hyperlink analysis and experimented with network analysis in order to map communities. So that our scraping and analysis of Gab would be of educational value, we published the toolkit,[14] ensuring that the basic documentation and an abundant number of inline comments allowed for ease of understanding. Finalizing the development project, we presented the toolkit before the SIG Text Mining and focus area Applied Data Science.

Migrating Publics

With the amping up of regulation and moderating of mainstream social media, most notably Twitter, worries emerged that an exodus of Twitter and other deplatformed users would gather on Gab, and that this new platform would form a pipeline to right-wing extremism. Blekkenhorst et al. (2019) performed exploratory research into users arriving at Gab after being banned from Twitter in order to find some basis for accurately accessing these migrating publics and the Twitter-to-Gab pipeline. The first step of this research process was identifying Gab users who had been banned from Twitter. Attempts to link Gab users to banned Twitter accounts proved unsuccessful. With username diverging between platforms, and it being unclear whether and when banned accounts still publicly appeared on Twitter, there was no way to be certain, based on publicly available data, whether particular Twitter and Gab accounts corresponded to the same "real world" person. We ended up using the Gabbers' own admissions as indications of who had been banned from Twitter. Querying the database of all Gab posts and profiles for the string "twitter"—later supplemented by pejorative derivatives such as "Twatter" or "Shitter"—and subsequently performing a close reading of random samples of the results revealed that many Gab users who had apparently been banned

13 See HateSonar's open-source software at https://github.com/Hironsan/HateSonar (last released on July 1, 2020).

14 The Gabber toolkit published by the Center for Digital Humanities is available at https://github.com/CentreForDigitalHumanities/gabber (published on June 12, 2019).

from Twitter indeed addressed their bans in one of their first Gab posts. Another option was that Gab users mentioned a Twitter ban in their profile, seemingly as a badge of honor. Through this method, we were able to compose queries that identified at least the self-proclaimed banned Twitter users on Gab, of which we found 29,412.

Once these self-proclaimed banned Twitter users were identified, questions arose: Did these people actually stay on Gab? And if so, did they find footholds in unambiguously right-wing extremist communities? By analyzing the activity and lifespan of the accounts of the deplatformed publics, we showed that while the number of people who turn to Gab after being banned on Twitter might be significant, those who actually stay on Gab represent a far smaller group. Out of the 29,412 accounts, 11,042 showed no further public activity on Gab after the first week of their arrival on the platform, and only 3,452 were still active when the scrape was performed. Further community detection analysis and toxicity analyses revealed that of the ex-Twitter users who remained on the Gab platform, over 10 percent found a foothold in clearly right-wing extremist communities.

In every community detection analysis that was performed—be it based on repost, follow, quote, or comment relationships—one sizeable cluster was identified that had notably more hate speech than the others. Using the results of the topic modeling as well as random sampling of its posts, this cluster was identified as being neo-Nazi in nature. Of the 3,452 self-proclaimed banned Twitter users who actually stayed on Gab, 398 ended up as part of this neo-Nazi cluster. Of the 29,412 users who turned to Gab after presumably having been banned from Twitter, 37 percent were no longer active on Gab after the first week, and only 12 percent actually remained active users on Gab.[15] Despite this attrition, over 1 percent of these users ended up in a right-wing extremist community. These results do not indicate a strong waterbed effect or a robust Twitter-to-Gab pipeline, contrary to what critics of deplatforming often caution against.[16]

Being Nonradical on a Radical Platform

Analyses of the Gab project generally show elevated levels of hate speech and political posting on the fringe platform as compared to its mainstream counterparts (Zannettou et al. 2018; Zhou et al. 2019; Munn 2019; Zeng and Schäfer 2021). Salazar and de Winkel (2019) reengaged with the image of Gab as a radical space by researching the opposite presupposition. Whereas previous academic contributions considered Gab to be a homogenous platform and explored the connections between its users and alt-right

15 Again, Gab consisted of 770,000 profiles at the time of the scrape, meaning that under 4 percent of Gab users at that time were self-proclaimed deplatformed publics.

16 The "waterbed effect" refers to the idea that deplatformed content or users do not actually disappear from the web but just move to different spaces. It implies that the act of deplatforming is as useful as pushing down a lump in a waterbed.

political activity, they focused on the fragmented publics and nonradical activity. They asked, "Can you be nonradical on a radical platform?" By turning the question around they attempted to show, and thus account for, the manner in which a question to a large degree determines the answer.

To see whether Gab could also be a platform for non-right-wing publics, nonextremist content, or even nonpolitical discourse, Salazar and de Winkel focused on Gab "groups." Users on Gab have the opportunity to create or join groups, and to communicate with other members in a members-only part of the service. These occasional communities can revolve around any topic, ranging from discussing shared hobbies and interests to being instructive about the forum or bringing together people from a specific community, country, or region. By investigating whether these groups on Gab provide spaces to discuss a variety of non-right-wing topics, Salazar and de Winkel question the perceived homogeneity of far-right spaces on fringe platforms, and they show that these groups provide a rationale for research on communities or partial publics.

The Gab scrape provided 1,631 groups, of which 1,310 were discarded because they had 100 or fewer posts. The 321 groups left—with a combined 810,570 posts—were categorized qualitatively into six categories, based on the group title and a close reading of content samples. The largest category, named *topical*, included 155 different groups that each focused on one specific topic, interest, or hobby—such as pets, comics, cannabis, or computer technology. The second largest category consisted of 87 *explicitly political* groups, meaning that it was clear from these groups' names or descriptions that they were meant to talk about a given political issue or about politics itself. The third largest category, which was called *identitarian*, contained the 35 identity-driven groups—for example "Manly Men of Gab," "Traditional Lifestyle," and "Christianity."[17] A fourth category was made up of the 24 groups that revolve around the *nationality or locality* of the users. The fifth category consisted of the 17 groups that center on forum *communication*, such as "Introduce Yourself," "Breaking News," or "Gab on Android." (A sixth category consisted of only three groups categorized as *miscellaneous*.) The three smallest categories of groups—*forum communication*, *nationality or locality*, and *miscellaneous*—are rather typical for groups on social media platforms. But the three largest categories—*topical*, *explicitly political*, and *identitarian*—might be a good indicator of whether nonradical and nonpolitical speech is possible on Gab. Especially salient is the high number of topical communities because this significant footprint

17 The differentiation between political and identitarian groups is somewhat contrived, especially in the far-right context. However, they wanted to allow for the hypothesis that you could be nonradical and/or nonpolitical on a politically radical platform; moreover, there is a difference in political explicitness between a "Trump2020" group and a "traditional lifestyle" group. For that reason, we used the definition of "identitarian" in the *Collins English Dictionary online*, which gives "concerned with promoting the interests of one's own cultural group," as a categorization. See https://www.collinsdictionary.com/us/dictionary/english/identitarian (accessed June 1, 2024).

suggests that almost half of the forum is organized around people's nonpolitical interests. And while the identitarian groups often carry a strong political connotation, since identity is—especially on Gab—political by nature, the topical groups seem rather neutral, sometimes even innocent. After the categorization, they performed a series of computational analyses on this corpus, including determining the number of active users, visualizing which groups shared active members, and calculating the distribution of hate speech. Findings on user participation and hate speech confirmed earlier findings of elevated toxicity as well as greater activity in the identitarian, political, and national groups, confirming the notion that social media platforms, and especially Gab, are used for political discourse (Salazar and de Winkel 2019; de Winkel et al. 2019). A network analysis (figure 1) reveals that the platform was broadly divided into two major spheres, namely the topical groups and the explicitly political groups. The division between the political—and many of the identitarian—groups and the topical groups suggests that those who comment in the political groups often comment in other political groups, while users in topical groups generally tend to be more active in other topical groups than they are in political groups. So, if we assume that the topical groups are indeed relatively nonradical and nonpolitical, we can also assume that the active users in such groups have a considerably less radical and political platform experience on Gab in general.

The network shows ample connections between the political groups and the nonpolitical groups, indicating that Gabbers in fact meet in shared spaces instead of inhabiting segmented trenches. One could suppose that Gab is a heterogenous platform in terms of discourse and usage; nonetheless, it is hard to believe that members of nonpolitical groups are drawn to Gab for the opportunities it provides—for example—to share their photography. Gab is too small, too inactive, too controversial, and too political for such goals. In all likelihood, users are drawn to Gab because they are looking for an alternative social network that suits their political affiliation, and which will allow them to escape the censorship regulations found on mainstream social media platforms. So, while partial publics are present on this fringe platform, they do not "prove" that Gab functions as an alternative sphere for discourse, instead of as a far-right radical(izing) platform.

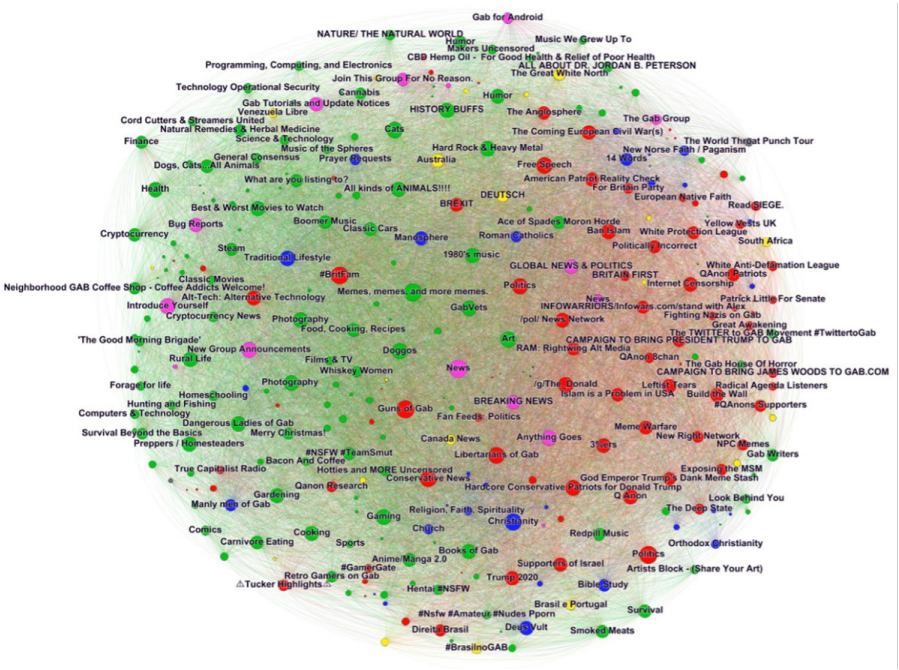

Figure 1. The network graph represents the landscape of the groups on Gab. Groups that are close have a high probability of sharing a relatively high number of active users. All 321 groups that made the final selection are represented by a node in the graph, and the edges that connect the nodes of the network represent the activity of shared users. We used the force-directed layout algorithm called "force atlas 2" of the network analysis software Gephi (Bastian, Heymann, and Jacomy 2009), and so the logics of the network are therefore based on gravitational forces, where the nodes reject each other and the edges can be likened to bands pulling the nodes toward each other. The more connections two nodes make, the closer they are pulled toward each other (Jacomy et al. 2014). Source: authors' construction.

Truth Claims on Gab

With its stated focus on freedom of speech, both Gab's technical architecture and its regulatory framework are aimed at the uninhibited dissemination of user content. However, even a brief glimpse at its content reveals Gab to be a gathering place for racism, antisemitism, and other forms of hate speech, together with an array of conspiracies: in all, it is a disturbing example of the troublesome relationship toward truth that has come to be popularly known as post-truth. Gab's view of itself as "defenders of free speech, individual liberty, and truth" was articulated in a 2017 blog post:

> In the free market of ideas, the best ideas will always win. When those ideas start to rise and challenge the establishment, they have no choice but to silence and censor. They have no choice but to purge any ideology that does not conform to their own echo chamber

> bubble world. They do not and cannot relate to the average middle class family that can barely put food on the table. **They care only about three things: money, power, and control.** Enough is enough. The time is now for patriots and free thinkers inside and outside of Silicon Valley to **organize, communicate in a safe way, and start building.**[18]

Gab appeals to the "marketplace of ideas," a central concept within US free speech law, especially since 1969, when it formed the theoretical foundation for the US Supreme Court to overturn the conviction of a Ku Klux Klan leader on charges of advocating violence.[19] In contrast, Zannettou et al. (2018, 1) characterize Gab's focus on free speech as "merely a shield behind which its alt-right users hide." Accordingly, it would be prudent to question whether Gab can in fact be seen as a "marketplace of ideas" and whether Gab's users are actually committed to such free exchange, or whether the emphasis on free speech should be seen through a different framework.

Gorzeman's (2019) empirical and philosophical inquiry into Gab engages with the plausibility of the radical platform as a "marketplace of ideas" by indexing differing and conflicting ideas and opinions. To make this assessment, all hashtags used in posts on Gab were extracted, counted, indexed, and presented. A frequency list of all Gab hashtags was subsequently used to search for hashtags both favoring and opposing the politics of former US president Donald Trump, who was—and still is—a major topic within the Gab community. The most common hashtag found on Gab, with 682,750 occurrences, was #maga, clearly indicating support for Trump. By stark contrast, the most frequent hashtag opposing Trump, #nevertrump (referring to opposition to Trump, mostly among Republicans and other conservatives), comes in at 658th place, with only 2,424 occurrences. The hashtag #hillary ranks 30th, with a total of 35,286 occurrences, but the close reading of samples shows it to be used almost exclusively to criticize Hillary Clinton rather than to support opposition to Trump. In a marketplace of ideas we would expect to see competition between differing ideas. While differing and conflicting ideas exist on Gab, on one of the platform's most central topics they are extremely out of balance, so much so that one cannot plausibly imagine Gab as a marketplace of ideas.

To make sense of whether and how freedom of speech and the construction of truth are entangled on Gab, a corpus of hashtags that can be understood as "truth claims"—such as the hashtag #truth—was built, resulting in a collection of signifiers that all fulfill

---

18 The post, from August 10, 2017, has since been deleted from *Medium* but is archived at https://web.archive.org/web/20180203222908/https://medium.com/@getongab/announcing-the-alt-tech-alliance-18bebe89c60a (accessed June 1, 2024, original emphasis).

19 See Brandenburg v. Ohio, 395 U.S. 444 (1969), Justia US Supreme Court Center, accessed June 1, 2024, https://supreme.justia.com/cases/federal/us/395/444/.

some rhetorical function linked to the truth claims made or referred to in the posts. By organizing these hashtags in a graph, where each hashtag is represented as a node and each co-occurrence is represented by an edge, and after filtering out all the nodes with a relatively low number of edges, we can see the most dominant co-occurrences (figure 2). Of particular interest are #freespeech, #speakfreely, and #censorship. Seeing these hashtags so prominently intertwined with #truth would seem to suggest that Gab users commit to an entanglement of truth with freedom of speech.

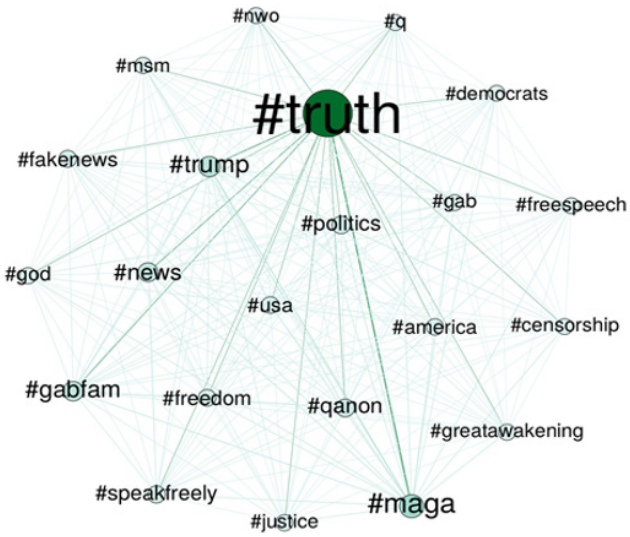

Figure 2. A co-occurrence graph of the selected corpus of hashtags; line weight reflects the number of co-occurrences between hashtags. Source: authors' construction.

Gab, which describes itself as a free speech platform for "defenders of truth," asserts, in effect, that it is a platform for the parrhesiastes, a term embraced by Gab project researchers to name one who practices "boldness or freedom of speech."[20] With almost no regulation of content and prevalent use of pseudonyms disguising posters' identities, Gab does seem to provide the ideal conditions for its users to speak freely and frankly. But precisely by providing such a safe haven, it appears to have also undermined the very possibility of a meaningful confrontation in speaking truth to power, that is,

20 The definition of "parrhesia" is from *Merriam-Webster Dictionary online*, https://www.merriam-webster.com/dictionary/parrhesia (accessed June 1, 2024). In parrhesia, truth, free speech, and speaking freely are closely entangled with one another. The fruitful discussion that followed the publication of Foucault's lectures on parrhesia, and in particular the works of Seitz (2016) and Simpson (2012), have given rise to a nuanced view of the particular sort of "truth" that is spoken in parrhesia, and its relation to power and subjectivity.

the very possibility of parrhesia. Foucault (2001) argues that "[s]omeone . . . merits consideration as a parrhesiastes only if there is a risk or danger for him in telling the truth." But on a platform like Gab, where free speech is nearly absolutely guaranteed, and moreover where the vast majority of users post under pseudonyms, it becomes hard to imagine any real risk or danger in speaking freely and frankly, except perhaps for certain extreme cases that merit criminal prosecution. Furthermore, Gab is an online platform, and such technologically mediated communication generally results in less inhibited speech than in face-to-face settings. Euripides characterizes a "bad" parrhesiastes as, in Foucault's words, "someone who has a tongue, but not a door," that is, someone with poor judgment who does not know when not to speak (Foucault 1983). We argue that the technological nature of Gab presents users with a tongue while holding open the door: it creates an environment devoid of many inhibitions that could otherwise aid in better judgment about when and when not to speak. In short, the racist—but parrhesiastic—rhetoric on Gab serves as encouragement, support, and opportunity for others to "freely" make racist truth claims, thus contributing to the emergence of a new regime of truth. The focus on free speech on Gab is thus more than a shield that its alt-right users can hide behind. Parrhesia is used as a rhetorical device that opens up space for political subjugation. When we focus on free speech not as a commitment to freedom of speech as a human right but rather as a promotion of parrhesiastic activity, it becomes clear how free speech fits within the political project of the alt-right—namely as a technique for political subjugation. The condition that Gab creates as a platform is the ability to safely break with nomos—the laws, norms, or conventions governing human conduct—without the need to accept any risk that would require one to have an existential stake in the matter, providing the ideal breeding ground for the technique of rhetorical parrhesia to flourish.

Theses

Three theses have been finalized as part of the Gab project. The first thesis is "Moderating Online Extremism on Fringe and Mainstream Platforms: An Analysis of Governance by Gab and Twitter" by Melissa Blekkenhorst (2019). In her research, Blekkenhorst compared the moderation on Gab and on its mainstream counterpart Twitter. By analyzing the guidelines, affordances, and enforcement efforts of both social media platforms she concluded that, though both platforms struggle with the responsibility of enabling free speech while attempting to prevent extremism, illegality, and disruption, Gab possesses a fundamentally different ideology and has divergent goals when compared to Twitter. Blekkenhorst calls for a deeper investigation into fringe platforms like Gab, specifically into their role within online public discourse and their motivations, which seem clandestine but are nonetheless very important to their moderation practices.

The second thesis, Sofie de Wilde de Ligny's (2022) "An Analysis of How Fringe Platform Gab.com Relates to the Process of Platformization," explores the

development of Gab's infrastructure and links the rise of fringe platforms to the process of platformization, or "the rise of the platform as the dominant infrastructural and economic model of the social web" (Helmond 2015, 1). Although platformization has been intensively studied and theorized in recent years (Van Dijck, de Winkel, and Schäfer 2021; Helmond 2015), de Wilde de Ligny stresses that the theoretical framework of platformization has been applied only to Big Tech or mainstream services. The inclusion of fringe platforms in the theoretical purview of platformization allows for new angles and a broader scope on the Web. Through a platform infrastructure analysis, she traces and maps Gab's technical infrastructure and finds a notable absence of platform mechanisms and API-business partners. Instead, the fringe platform increasingly uses—and offers—services from its own tech sphere. De Wilde de Ligny concludes that however much Gab may wish to separate itself from the Big Tech companies—by using and offering alternative services—the mainstream and fringe will still be part of the same online dynamic in the platform ecosystem.

The Gab project's third thesis, Thomas ten Heuvel's (2019) "Gab.ai: Het Platform waar Hate Speech een Recht is" (Gab.ai: The platform where hate speech is a right), uses text-mining methods to reveal the ideological positions and targets of speakers on Gab. By assigning a toxicity score to more than two million posts, Ten Heuvel filtered and analyzed posts containing hate speech, showing that hate speech on Gab is frequently aimed at Jews, Muslims, and people of color.[21] Furthermore, Ten Heuvel shows that the posters' ideological positions are highly white nationalist, conservative, and racist in nature. The analysis also reveals Gab users to be strong advocates for free speech who strongly oppose censorship and political correctness. At the same time, paradoxically, they attack and seek to silence people with other ideological beliefs. Free speech therefore seems to be used as a justification for spreading racist and antisemitic views. Ten Heuvel concludes that although Gab styles itself as a platform for free speech and individual freedom, it is in fact an echo chamber where hate speech is a right.

Researching Platforms in the 2020s

While every subproject was an investigation into Gab as an alt-right or fringe platform, the Gab project's overarching goal with the computational methods employed to create and research our dataset was to develop a methodological framework for researching the platformized far right through data practices. As stated in the introduction, when the far right was platformized its manner of communication and the dissemination of its views—along with its tactics, vernacular, and ultimately its culture—were

21 Ten Heuvel used the Google API called "Perspective" to determine toxicity, defined as a "rude, disrespectful, or unreasonable comment that is likely to make people leave a discussion." See "Attributes and Language," Perspective, https://support.perspectiveapi.com/s/about-the-api-attributes-and-languages?language=en_US (accessed June 26, 2024).

transformed. To ask how one can research a far-right platform is, at least partly, to ask how to research a platformized community or movement. Because our datafied object warranted the use of data methodology, we had to incorporate insights and frameworks from the fields of the digital humanities, critical data studies, and platform studies. In this section we discuss data-driven humanities research conducted on the platformized far right. The structure of this section is aligned with the two types of challenges we encountered: institutional challenges, which were mostly legal and will be discussed first, and epistemological challenges that come with conducting data-driven humanities research.

Legal and Institutional Challenges

The first institutional challenge was a lack of resources and infrastructure. On the platform side, there were no application interfaces that would allow us to easily access data, and on the university side, there was a lack of expertise in collecting data through other techniques. As stated, the data collection was eventually made possible through a modest grant made by Utrecht University, which allowed us to hire an expert who could reverse-engineer the undocumented backend API of the Gab website. At that time no one on our faculty possessed the necessary skills to perform this task, and only due to the Data School's extensive professional network were we able to employ someone with requisite expertise at such an affordable price. A university research group lacking half a decade of data-driven research with various outside partners would probably not have been able to get the Gab project off the ground. Given the volatility of the platform ecosystem, the project at hand also presented a welcome opportunity to learn more about accessing platforms that do not provide well-documented APIs (Perriam, Birkbak, and Freeman 2020; Tromble 2021), since we assume that platform researchers will encounter such challenges more often. We believe it is epistemologically problematic to largely focus on easily accessible platforms.

Ethical and legal issues presented another challenge (see Krotov, Johnson, and Silva 2020). Novel data repositories and the means to access them can challenge research integrity and require rigorous ethical consideration (Van Schie, Westra, Schäfer 2017). Web scraping became a controversial issue, with platform providers often challenging researchers and questioning their justifications for data collection (e.g., Luscombe, Dick, and Walby 2022). Compiling and capturing a dataset of a web platform are routine activities in social media research (e.g., Marres and Weltevrede 2013), but we were adamantly committed to using this project as an exemplary case for doing ethical data research and for complying with academic integrity standards (Algra et al. 2018; Franzke et al. 2020) and GDPR (General Data Protection Regulation) regulations. We consulted a privacy lawyer for developing a GDPR-compliant research process.[22]

22 This lawyer later developed a "research justification, methodology, and integrity" manual for this

With an eye to contract law, the scraping was in line with Gab's own terms of use, as the platform did not prohibit automated data collection at the time of scraping.[23] We carried out a data protection impact assessment and developed a data management plan prior to the data collection. Both endeavors were necessary for compliance with our university's guidelines. Our data collection meets the terms of the GDPR's exceptions for research purposes.[24] Of course, researchers must take appropriate precautions to prevent breaches of privacy. Drawing from our data protection impact assessment, we concluded that many usernames were pseudonyms that did not represent the users' real identities. In addition, we decided to anonymize the dataset, as the general dynamics of Gab as a communication platform were more relevant to our project than individual users.

While commercial entities are required to inform users about data collection, the GDPR allows researchers to make their findings public through general publications about their research activities or by publishing their research in scholarly journals and disseminating their work at conferences. Data collection thus can take place without informing the data subject prior to the data scrape. This is permitted under the terms of the GDPR when the gathering of such information would appear otherwise infeasible or if data and research quality would be negatively affected. Of course, there was no way to inform hundreds of thousands of—largely anonymous—account holders on a far-right platform in advance of a data scrape. Legally, the data collection was covered because the research interest was legitimate and because Gab's terms of use did not prohibit data collection. To ensure GDPR compliance and data protection, we conducted ethical deliberations about the research project and its method and process, applied data minimization through the anonymization of the dataset, restricted access to the dataset to researchers working on the project, and stored the data in a secure environment using the research "workspace" compartments available via the research data management service Yoda, developed at Utrecht University.[25] We wrote up a detailed research plan and closely documented the data collection process. In addition, a public seminar was held to inform interested parties about the project, its research design, and the data collection; the seminar complemented its public informational role

---

kind of digital humanities research and web scraping, at least in part as a result of the Gab project. It has now become a set of official guidelines for data collection at Utrecht University. See Gerritsen (2021).

23 Gab's terms of service are archived at https://archive.ph/PZEqE (last updated August 18, 2016).

24 As per Article 43 of the Dutch law "Uitvoeringswet Algemene verordening gegevensbescherming," Overheid.nl Wettenbank, accessed June 1, 2024, https://wetten.overheid.nl/jci1.3:c:BWBR0040940&hoofdstuk=4&artikel=43&z=2018-05-25&g=2018-05-25.

25 See Ton Smeele and Lazlo Westerhof, "Using iRods to Manage, Share and Publish Research Data: Yoda," Utrecht University, n.d., accessed June 1, 2024, https://irods.org/uploads/2018/Smeele-UtrechtUni-Yoda-paper.pdf.

by educating our colleagues regarding practices of research design, data management, and GDPR compliance for research projects requiring web scraping.

Throughout the process, we kept in contact with the privacy officer as well as the privacy advisers and data managers at the university. It turned out that a lack of jurisprudence, the fear of significant penalties, and an unfamiliarity with social media research were at the root of many difficult but valuable conversations between our research team and the university's privacy and data team. The minimization of liability sometimes seemed at odds with guaranteeing the fundamental right to research. Our project therefore initiated an institutional learning process about the possibilities, legal boundaries, and best practices for GDPR-compliant research. It launched, albeit slowly, the development of a dedicated ethics board for social media data research.

Epistemological and Ethical Challenges

Studying cultural phenomena through data analysis affects the methodology employed (Van Es and Schäfer 2017). We need the lessons provided by the field of critical data studies to address the methodological and—to an even greater extent—the epistemological challenges posed by such a transformation. First, capturing all the data points of a social media service in a dataset means that you have altered your object of analysis and must now reverse-engineer the dataset back into the online public and/or discourse. All our choices regarding data capture, data storage, data analysis, and data presentation—from our programming language to our visualization software—are axiomatically part of, and inseparable from, our research object. A thorough and critical review of the array of data-driven methods we use within the humanities must become part and parcel of the research process (Van Es, Schäfer, and Wieringa 2021). We will refrain from declaring a new methodological tradition or coining new terminology, but the involvement of the entire process of data methodology (capture, storage, selection, analysis, visualization) and the limited use of ready-made and/or commercial tools and packages seems to deploy the framework that critical data studies aims to teach. Second, although we scraped an entire forum, by no means do we now possess a god's-eye perspective. Haraway (1988) has long reminded us that the god-trick fallacy of seeing everything as if from nowhere—whereby the researcher or analyst imagines not being situated in the world but serving as a convergence point of objective observation—is especially potent in technology-enhanced vision and representations of big data knowledge (Leurs 2017) such as data visualizations (Kennedy et al. 2016). Therefore, the challenge is to fully understand within our methodology that we do not in fact see the entirety of the platform, we see only our dataset. A distant reading perspective is better described as a bird's-eye view: this metaphor encompasses the notion that a bird sees a great deal due to a very specific position in the world, even if this same bird sees very little of what is visible from another perspective, namely the one we have on the ground and up close.

Finally, we must not confuse the map with the territory. A user is not the same as a person, and likewise a datafied social media public is not equivalent to an offline community or culture. An online public is both socioculturally and technologically shaped and is cocreated by the medium on which it has assembled (de Winkel 2023, 40). The affordances and algorithms, and all the political and economic incentives they embody, are just as much part of the activity and community as social actors such as users, publics, and influencers. These considerations are more fundamental yet less obvious than the challenges posed by the white noise of bots or trolls or any other disruptive, misleading, or automated activity. A dataset like ours becomes meaningful when the distance separating the object from the digital presence of that object, and from the representation of that presence in the dataset, are thoroughly understood. As Caplan (2016, 6) concludes,

> [W]e should be warned: although data is neither map nor territory, it can foster their confusion. It is a seductive mode of representation that can easily trap an intellectual milieu terrified by representation, providing a method for running away from its history and its own activity in the present.

So, the far right does not become known to us in this dataset, but the interpretation of the data analyses can take place within the context of the medium and when the interpretations are applicable to the datafied object. We are able to analyze traits of the platformized far right. For example, the assertion that 1 percent of members of the deplatformed publics of Twitter joins a neo-Nazi cluster on Gab is meaningful within the framework of online migration and the waterbed effect. It does not specify for us whether the Twitter publics were radicalized by their deplatforming or, alternatively, were neo-Nazis to begin with and then found a community. The usage of new methodologies such as data analyses thus leaves the researcher vulnerable to epistemological misunderstanding. Employing such a methodology requires an in-depth understanding of what you are studying and a familiarity with the field of critical data studies. Additionally, since the far right is increasingly often also a platformized object of study, we might incorporate the lessons provided by the field of platform studies—most notably, the understanding that a platform is a multidimensional ecosystem, determined not solely by user activity and technology but also by economic and governance models (Van Dijck 2013). To study a social media platform is also to study the data economy (Gerlitz and Helmond 2013), partnerships (Van der Vlist 2022), and infrastructures (Plantin et al. 2018). De Winkel (2023), developing this framework, argues that far-right fringe platforms such as Gab should be studied as part of the larger platformized public sphere and through the power dynamic between a dominant center and the fringes—taking to heart Federico Finchelstein's (2019, 2021) suggestion that fascist ideologies are best studied from the margins. An example of such research on the far right would be the subproject of migrating publics, since it applies a

cross-platform lens and acknowledges the dispersed and translocal nature of the online far right.

Finally, there are research ethics and considerations specifically warranted by study of the far right. We decided to largely suspend judgment with regard to their radicalism, without ignoring the toxicity and danger that fascism and adjacent ideologies present for free and open societies. The posture of suspended judgment is due to our desire to maintain our curiosity and open-mindedness toward the dataset and our research object, but such generosity is never extended to the actual politics of the far right.[26] Additionally, researching the online far right brings with it the possibility of virality. An important goal of far-right content producers is to have their content circulate as long and widely as possible. The online far right needs mainstream actors to disseminate their radicalism to a broader audience. This point is made most strikingly by Whitney Phillips (2018), whose report "The Oxygen of Amplification" implores center-left media, including journalists and editors, to stop giving oxygen to hate speech: even as they debunk or critically cover alt-right channels and discourses, their reporting on such noxious content nonetheless signal-boosts the far right. Journalists and scholars are often baited into investigating and explaining their content, which is packaged in such a way that the general public takes note of their tropes and presence without fully understanding their underlying dynamics. Far-right publics, for their part, often immediately recognize the meaning behind the far-right content and are emboldened by seeing a far-right presence in the mainstream. We have withheld spectacular details and sources of far-right propaganda, such as manifestos, in order to avoid being useful idiots for their cause. Every research subcomponent of the Gab project was aimed at finding the structure instead of particular cases, and this is especially true of the two subprojects on Gab's publics (migrating publics) and discourse (truth-telling). Research that focuses on what happens *on* a far-right platform, rather than what happens *with* a far-right platform, has a greater risk of highlighting the spectacular and the horrific, thereby perpetuating attention to such repellent content and platforming that which ought not to be amplified.

Conclusion and Discussion

Over a period of five years, we ran a research project analyzing the far-right service Gab. Our Gab project included scraping this entire social media platform into a dataset and the employment of a spectrum of data methods. The analyses disclosed Gab to be a far-right social media platform deeply rooted in alt-right culture. Data analyses show Gab to have high percentages of hate speech and political speech, low levels of user participation, and very limited moderation, yielding a platform full of

26 De Winkel (2023, 14) is very clear on this point in his dissertation: "[W]e are dealing with violent and illiberal extremism here, and . . . I condemn and oppose the far-right, openly and vehemently."

noise and toxicity. Although text and network analyses had shown that nonradical and nonpolitical discourse and activity are certainly possible and are in fact present on Gab, we refuted that their presence is indicative of any flourishing public sphere or diversity of opinion. Subsequently, we confirmed this refutation of diversity through a hashtag analysis. Furthermore, we argue that the platform creates conditions where it is safe to engage in racism, antisemitism, and sexism. While Gab disguises itself as ostensibly a platform for free speech and individual freedom, it is an echo chamber where hate speech is encouraged. This "rhetorical parrhesia"—speaking freely without consequence—functions as a device of political subjugation and thus is aligned with the political project of the alt-right. Engaging with the notion of Gab as a radicalizing platform, we investigated the migrating publics that were deplatformed on Twitter and are now present on Gab. However, these analyses cannot unambiguously confirm that deplatforming users off of Twitter might cause them to be radicalized on Gab. The majority of those from deplatformed publics who land on Gab do not stay there for long—presumably for however long they are banned—and only a very small percentage joins an overtly neo-Nazi community.

However, the Gab project was not interested solely in researching a single alt-right platform. The overarching goal of all the subprojects and the plethora of analyses was to contribute to a methodology for the study of the contemporary online far right. As scholars working at the intersection of media and the far right, we recognized both the challenges and the necessity of doing this type of data-driven research. Taking our atypical project as a case study, we went through all the structures of our institution to experience and document how to perform data analyses as humanities scholars. Our first contribution to data-driven far-right scholarship is our appeal to incorporate the fields of critical data studies and platform studies when analyzing the platformized far right. Researching the far right in the new decade is going to involve—to some extent—researching the platformized far right, alt-right, and Alt-Tech. Since the research object of scholars of the far right is now partly datafied, the field has to take notice of the possibilities of data methods, or at least be related to research that does. This is not an appeal to solely perform data analyses, but scholars that do use digital tools and data methods for the capture and analysis of web platforms must become literate in operating them. Other fields that become relevant with the study of a datafied object—specifically social media data—are data ethics and law. Humanities scholars who apply data science methods must also develop an understanding of their legal responsibilities and the practices of compliant research. In addition, their research institutions must be able to provide the necessary legal support to defend their fundamental right to research and to advise, effectively and competently, on proper conduct. Institutions can either build structures that bolster the legal and methodological foundations of this type of research or, indeed, open up unspecified money grants for anomaly research. However, the presence of financing for data science does not mean legal bodies, facilitating services, or departmental research agendas are necessarily prepared for this research. We advise

universities to allocate time and resources to build tool repositories and create flexible access to computational resources and legal support.

In addition, we want to state the importance of open and interdisciplinary science for this type of study. The Gab project would not have been possible without these principles. There is a lot of valuable output besides the publication of journal articles, and we have made extensive efforts to guarantee the accessibility, usability, and reusability of our work. The caveat is that currently, and until further notice, the dataset is reserved for internal reuse only in order to guarantee good practice.[27] The dataset has been of great value and has already contributed greatly to knowledge transfer, and the same holds true for the toolbox. It is of great importance that all programming should come with loads of comments and documentation; tools, scrapers, and the dataset itself should be explained. Writing the code is only 30 percent of the work, with the rest devoted to explaining your code and facilitating use. Digital humanities research and data-driven scholarship are, intrinsically, forms of interdisciplinary research and therefore warrant cooperation. Contrary to popular belief, data science usually increases the workload of those engaged in it. Additionally, it is unrealistic and unadvisable to demand that everybody learn to code. Cooperation and interdisciplinary scholarship, however, are well equipped to overcome these challenges.

References

Albertson, Bethany. 2015. "Dog-Whistle Politics: Multivocal Communication and Religious Appeals." *Political Behavior* 37: 3–26. https://doi.org/10.1007/s11109-013-9265-x.

Algra, Keimpe, Lex Bouter, Antoine Hol, and Jan van Kreveld. 2018. "Nederlandse Gedragscode Wetenschappelijke Integriteit." *KNAW; NFU; NWO; TO2-federatie; Vereniging Hogescholen; VSNU*. https://doi.org/10.17026/dans-2cj-nvwu.

Bagavathi, Arunkumar, Pedram Bashiri, Shannon Reid, Matthew Phillips, and Siddharth Krishnan. 2019. "Examining Untempered Social Media: Analyzing Cascades of Polarized Conversations." In *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* 2019, 625–32. http://dx.doi.org/10.1145/3341161.3343695.

Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy. 2009. "Gephi: An Open Source Software for Exploring and Manipulating Networks." In *Proceedings of the International AAAI Conference on Web and Social Media* 3: 361–62. https://doi.org/10.1609/icwsm.v3i1.13937.

27 Internal reuse does not mean only researchers from the Data School can use the data; rather, it means that we keep this dataset on our infrastructures. For access or request of transparency, please contact dataschool@uu.nl.

Blekkenhorst, Melissa. 2019. "Moderating Online Extremism on Fringe and Mainstream Platforms." Master's thesis, Utrecht University. https://studenttheses.uu.nl/handle/20.500.12932/33879.

Blekkenhorst, Melissa, Ludo Gorzeman, Mauricio Salazar, and Tim de Winkel. 2019. "Studying the Fringe of the Platform Society: Gab.ai." Presentation, 3rd International Data Power Conference, University of Bremen, September 12, 2019.

Boeschoten, Thomas, and Coen van den Ven. 2019. "Je Gaat Me Toch Niet Als Genuanceerd Wegzetten Hè?" *Groene Amsterdammer* 35, August 28, 2019. https://www.groene.nl/artikel/je-gaat-me-toch-niet-als-genuanceerd-wegzetten-he.

Breland, Ali. 2023. "The Website That Wants You to Kill Yourself—and Won't Die." *Mother Jones*, March/April 2023. https://www.motherjones.com/politics/2023/02/kiwi-farms-die-drop-cloudflare-chandler-trolls/.

Caplan, Lindsay. 2016. "Method without Methodology: Data and the Digital Humanities." *E-Flux* 72, April 2016. http://worker01.e-flux.com/pdf/article_9006656.pdf.

Davidson, Thomas, Dana Warmsley, Michael Macy, and Ingmar Weber. 2017. "Automated Hate Speech Detection and the Problem of Offensive Language." In *Proceedings of the international AAAI conference on web and social media* 11 (1): 512–15. https://doi.org/10.1609/icwsm.v11i1.14955.

Department of Homeland Security. 2019. *DHS Strategic Framework for Countering Terrorism and Targeted Violence*. Washington, D.C., September 20, 2019. https://www.dhs.gov/publication/dhs-strategic-framework-countering-terrorism-and-targeted-violence.

Dickson, Caitlin, and Christopher Wilson. 2018. "Who Is Gab Founder Andrew Torba?" *HuffPost*, October 30, 2018. https://www.huffpost.com/entry/gab-founder-andrew-torba_n_5bd89d70e4b0dee6eececbc5.

Dijck, José van. 2013. *The Culture of Connectivity: A Critical History of Social Media*. New York: Oxford University Press.

Dijck, José van, Tim de Winkel, and Mirko Tobias Schäfer. 2021. "Deplatformization and the Governance of the Platform Ecosystem." *New Media & Society* 0 (0): 1–17. https://doi.org/10.1177/14614448211045662.

———. 2024. "Deplatforming and Deplatformization as Governance Strategies." In *Handbook of Media and Communication Governance*, edited by Manuel Puppis, Robin Mansell, and Hilde van den Bulck, 503–16. Cheltenham: Edward Elgar.

Donovan, Joan. 2019. "Navigating the Tech Stack: When, Where and How Should We Moderate Content." *Centre for International Governance Innovation*, October 28, 2019. https://www.cigionline.org/articles/navigating-tech-stack-when-where-and-how-should-we-moderate-content/.

Donovan, Joan, Becca Lewis, and Brian Friedberg. 2019. "Parallel Ports: Sociotechnical Change from the Alt-Right to Alt-Tech." In *Post-Digital Cultures of the Far Right: Online Actions and Offline Consequences in Europe and the US*, edited by Maik Fielitz and Nick Thurston, 49–65. Bielefeld: transcript Verlag.

Dougherty, John, and Michael Edison Hayden. 2019. "How Gab Has Raised Millions Thanks to This Crowdfunding Company." *Southern Poverty Law Center*, January 24, 2019. https://www.splcenter.org/hatewatch/2019/01/24/how-gab-has-raised-millions-thanks-crowdfunding-company.

Es, Karin van, and Mirko Tobias Schäfer. 2017. "Introduction: New Brave World." In *The Datafied Society: Studying Culture through Data*, 13–22. Amsterdam: Amsterdam University Press.

Es, Karin van, Mirko Tobias Schäfer, Maranke Wieringa. 2021. "Tool Criticism and the Computational Turn: A 'Methodological Moment' in Media and Communication Studies." *Medien & Kommunikationswissenschaft* 69 (1): 46–64.

Figenschou, Tine Ustad, and Karoline Andrea Ihlebæk. 2019. "Challenging Journalistic Authority: Media Criticism in Far-Right Alternative Media." *Journalism Studies* 20 (9): 1221–37.

Finchelstein, Federico. 2019. *From Fascism to Populism in History*. Oakland: University of California Press.

———. 2021. "On Why the History from the Margins Is Central to Fascism and Populism." Class lecture at KNIR, Rome, December 16, 2021.

Foucault, Michel. 1983. "Parrhesia in the Tragedies of Euripides" (part 2). In Discourse and Truth: The Problematization of Parrhesia. Six lectures given by Michel Foucault at the University of California at Berkeley, October–November 1983. MΦ (Michel Foucault), "Explore Parrhesia," accessed June 14, 2024. https://foucault.info/parrhesia/foucault.DT2.parrhesiaEuripides.en/.

———. 2001. *Fearless Speech*, edited by Joseph Pearson. Los Angeles: Semiotext(e).

Franzke, Aline, Anja Bechmann, Michael Zimmer, Charles Ess, and the Association of Internet Researchers. 2020. *Internet Research: Ethical Guidelines* 3.0. https://aoir.org/reports/ethics3.pdf.

Gerlitz, Carolin, and Anne Helmond. 2013. "The Like Economy: Social Buttons and the Data-Intensive Web." *New Media & Society* 15 (8): 1348–65. https://doi.org/10.1177/1461444812472322.

Gerritsen, Joost. 2021. *Tactvol contactloos onderzoek—Handleiding voor onderzoekers die de AVG respecteren* (Tactful non-contact research—Handbook for researchers who respect the GDPR). Utrecht University. https://dspace.library.uu.nl/handle/1874/406645.

Gorzeman, Ludo. 2019. "#SpeakingTruth: Understanding False Epistemologies on an Alt-Right Fringe Platform." Project presented at the International Conference on "Technology and Society," KU Leuven, Belgium. https://nomennesc.io/presentatie-leuven/.

Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14 (3): 575–99. https://doi.org/10.4324/9780203427415-40.

Helmond, Anne. 2015. "The Platformization of the Web: Making Web Data Platform Ready." *Social Media + Society* 1 (2): 1–11. https://doi.org/10.1177/2056305115603080.

Hern, Alex. 2022. "TechScape: How Kiwi Farms, the Worst Place on the Web, Was Shut Down." *Guardian*, September 7, 2022. https://www.theguardian.com/technology/2022/sep/07/techscape-kiwi-farms-cloudflare.

Heuvel, Thomas ten. 2019. "GAB.AI: Het Platform waar Hatespeech een Recht is." Bachelor's thesis, Utrecht University. https://studenttheses.uu.nl/handle/20.500.12932/33790.

Hutchinson, Bill, Mike Levine, Janet Weinstein, and Matt Seyler. 2018. "'Screw the Optics, I'm Going In': Alleged Synagogue Shooter Posts on Social Media Moments before Massacre." *ABC News*, October 28, 2018. https://abcnews.go.com/US/pittsburgh-synagogue-alleged-mass-shooter-told-swat-officers/story?id=58803485.

Jacomy, Mathieu, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian. 2014. "ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software." *PloS One* 9 (6): e98679. https://doi.org/10.1371/journal.pone.0098679.

Jasser, Greta, Jordan McSwiney, Ed Pertwee, and Savvas Zannettou. 2021. "'Welcome to #GabFam': Far-Right Virtual Community on Gab." *New Media & Society* 25 (7): 1728–45. https://doi.org/10.1177/14614448211024546.

Kennedy, Helen, Rosemary Lucy Hill, Giorgia Aiello, and William Allen. 2016. "The Work That Visualisation Conventions Do." *Information, Communication & Society* 19 (6): 715–35. https://doi.org/10.1080/1369118x.2016.1153126.

Krotov, Vlad, Leigh Johnson, and Leiser Silva. 2020. "Tutorial: Legality and Ethics of Web Scraping." Faculty and Staff Research and Creative Activity, Murray State University. https://digitalcommons.murraystate.edu/faculty/86/.

Leurs, Koen. 2017. "Feminist Data Studies: Using Digital Methods for Ethical, Reflexive and Situated Socio-Cultural Research." *Feminist Review* 115 (1): 130–54. https://doi.org/10.1057/s41305-017-0043-1.

Luscombe, Alex, Kevin Dick, and Kevin Walby. 2022. "Algorithmic Thinking in the Public Interest: Navigating Technical, Legal, and Ethical Hurdles to Web Scraping in the Social Sciences." *Quality & Quantity* 56 (3): 1023–44.

Makuch, Ben. 2019. "Mastodon Was Designed to Be a Nazi-Free Twitter—Now It's the Exact Opposite." *Vice*, July 11, 2019. https://www.vice.com/en_us/article/mb8y3x/the-nazi-free-alternative-to-twitter-is-now-home-to-the-biggest-far-right-social-network.

Marres, Noortje, and Esther Weltevrede. 2013. "Scraping the Social? Issues in Live Social Research." *Journal of Cultural Economy* 6 (3): 313–35.

McLaughlin, Timothy. 2019. "The Weird, Dark History of 8chan." *Wired*, August 6, 2019. https://www.wired.com/story/the-weird-dark-history-8chan/.

Munn, Luke. 2019. "Alt-Right Pipeline: Individual Journeys to Extremism Online." *First Monday* 24 (6). https://doi.org/10.5210/fm.v24i6.10108.

Nationaal Coördinator Terrorismebestrijding en Veiligheid. 2018. *NCTV Report. De Golfbewegingen van Rechtsextremistisch Geweld in West-Europa*. Den Haag. https://

www.nctv.nl/documenten/rapporten/2018/11/05/de-golfbewegingen-van-rechts-extremistisch-geweld-in-west-europa.

Otala, Jacqueline, Gillian Kurtic, Isabella Grasso, Yu Liu, Jeanna Matthews, and Golshan Madraki. 2021. "Political Polarization and Platform Migration: A Study of Parler and Twitter Usage by United States of America Congress Members." In *Proceedings of the Web Conference* 2021, 224–31. https://doi.org/10.1145/3442442.3452305.

Pagliery, Jose, and Konstantin Toropin. 2018. "Social Network Gab, a Home for Anti-Semitic Speech, Produced Some of Its Own." *CNN*, October 30, 2018. https://www.cnn.com/2018/10/30/tech/gab-anti-semitic-speech-invs/index.html.

Patrick, Stephanie. 2022. "More than a Sex Crime: A Feminist Political Economy of the 2014 ICloud Hack." *Media, Culture & Society* 44 (1): 39–55. https://doi.org/10.1177/01634437211022713.

Perriam, Jessamy, Andreas Birkbak, and Andy Freeman. 2020. "Digital Methods in a Post-API Environment." *International Journal of Social Research Methodology* 23 (3): 277–90. https://doi.org/10.1080/13645579.2019.1682840.

Phillips, Whitney. 2018. "The Oxygen of Amplification." *Data & Society*, May 22, 2018. https://datasociety.net/library/oxygen-of-amplification/.

Plantin, Jean-Christophe, Carl Lagoze, Paul Edwards, and Christian Sandvig. 2018. "Infrastructure Studies Meet Platform Studies in the Age of Google and Facebook." *New Media & Society* 20 (1): 293–310. https://doi.org/10.1177/1461444816661553.

Rieger, Diana, Anna Sophie Kümpel, Maximilian Wich, Toni Kiening, and Georg Groh. 2021. "Assessing the Extent and Types of Hate Speech in Fringe Communities: A Case Study of Alt-Right Communities on 8chan, 4chan, and Reddit." *Social Media + Society* 7 (4). https://doi.org/10.1177/20563051211052.

Salazar, Mauricio, and Tim de Winkel. 2019. "Performing Diversity on Toxic Soil? Researching Topical Groups on the Radical Free Speech Platform Gab." Poster presentation at the AoIR Flashpoint Symposium "Below the Radar: Private Groups, Locked Platforms and Ephemeral Contents," University of Urbino, June 24, 2019.

Schie, Gerwin van, Irene Westra, and Mirko Tobias Schäfer. 2017. "Get Your Hands Dirty: Emerging Data Practices as Challenge for Research Integrity." In *The Datafied Society: Studying Culture through Data*, edited by Mirko Tobias Schäfer and Karin van Es, 183–200. Amsterdam: Amsterdam University Press.

Seitz, Sergej. 2016. "Truth beyond Consensus: Parrhesia, Dissent, and Subjectivation." *Epekeina: International Journal of Ontology, History and Critics* 7 (1–2): 1–13. https://doi.org/10.7408/epkn.

Simpson, Zacharia. 2012. "The Truths We Tell Ourselves: Foucault on Parrhesia." *Foucault Studies*, no. 13: 99–115. https://doi.org/10.22439/fs.v0i13.3461.

Sommer, Will. 2018. "Pittsburgh Synagogue Shooter Spewed His Hate on Gab, the Alt-Right's Favorite Social Network." *The Daily Beast*, October 27, 2018. https://www.thedailybeast.com/pittsburgh-synagogue-shooter-spewed-his-hate-on-gab-the-alt-rights-favorite-social-network.

Tromble, Rebekah. 2021. "Where Have All the Data Gone? A Critical Reflection on Academic Digital Research in the Post-API Age." *Social Media + Society* 7 (1). https://doi.org/10.1177/20563051219889.

Vlist, Fernando van der. 2022. "The Platform as Ecosystem: Configurations and Dynamics of Governance and Power." PhD dissertation, Utrecht University. https://dspace.library.uu.nl/handle/1874/422213.

Weill, Kelly. 2019. "Gab Is in Full Meltdown, and Its Founder Blames the 'Deep State.'" *The Daily Beast*, January 30, 2019. https://www.thedailybeast.com/gab-is-in-full-meltdown-and-founder-andrew-torba-blames-the-deep-state.

Wilde de Ligny, Sofie de. 2022. "An Analysis of How Fringe Platform Gab.com Relates to the Process of Platformization." Master's thesis, Utrecht University. https://studenttheses.uu.nl/handle/20.500.12932/41875.

Winkel, Tim de. 2023. "Fringe Platforms: An Analysis of Contesting Alternatives to the Mainstream Social Media Platforms in a Platformized Public Sphere." PhD dissertation, Utrecht University.

Winkel, Tim de, Mauricio Salazar, Ludo Gorzeman, and Melissa Blekkenhorst. 2019. "Hosting Radically Fuelled Periphery Networks: Investigating the Fringe Platform Gab.ai." Project presented at the 15th International Conference of Sociocybernetics, "'Dark Ages 2.0': Social Media and Their Impact," University of Urbino, Italy.

Zannettou, Savvas, Barry Bradlyn, Emiliano De Cristofaro, Haewoon Kwak, Michael Sirivianos, Gianluca Stringhini, and Jeremy Blackburn. 2018. "What Is Gab? A Bastion of Free Speech or an Alt-Right Echo Chamber?" In *Proceedings of the Companion of the Web Conference* 2018, 1007–14. https://doi.org/10.1145/3184558.3191531.

Zeng, Jing, and Mike Schäfer. 2021. "Conceptualizing 'Dark Platforms': Covid-19-Related Conspiracy Theories on 8kun and Gab." *Digital Journalism* 9 (9): 1321–43. https://doi.org/10.1080/21670811.2021.1938165.

Zhou, Yuchen, Mark Dredze, David Broniatowski, and William Adler. 2019. "Elites and Foreign Actors among the Alt-Right: The Gab Social Media Platform." *First Monday* 24 (9). https://doi.org/10.5210/fm.v24i9.10062.