

Variations on a Theme? Comparing 4chan, 8kun, and Other chans' Far-Right “/pol” Boards

by Stephane J. Baele, Lewys Brace, and Travis G. Coan

Abstract

Online forums such as 4chan and 8chan have grown in notoriety following a number of high-profile attacks conducted in 2019 by right-wing extremists who used their “/pol” boards (dedicated to “politically incorrect” discussions). Despite growing academic interest in these online spaces, little is still known about them; in particular, their similarities and differences remain to be teased out, and their respective roles in fostering a certain far-right subculture need to be specified. This article therefore directly compares the content and discussion pace of six different /pol boards of “chan” forums, including some that exist solely on the dark web. We find that while these boards constitute together a particular subculture, differences in terms of both rate of traffic and content demonstrate the fragmentation of this subculture. Specifically, we show that the different /pol boards can be grouped into a three-tiered architecture based upon both at once how popular they are and how extreme their content is.

Keywords: far-right; Internet; chan boards; 4chan; 8kun; language; computational methods.

Introduction, Research Question, and Hypothesis

Brenton Tarrant’s attack on two mosques in Christchurch, New Zealand, on 15 March 2019 was not only shocking in its own right, but also drew the public’s attention to 8chan, an anonymous “image-board” (online forum where users share images alongside text) where far-right discussions were carried out and extremist content shared. Indeed, shortly before his shooting, Tarrant announced on the forum “I will carry out an attack against the invaders”, and posted a manifesto as well as links to a Facebook livestream video of his attack. Among the first replies came encouragements; for instance, a picture of Hitler was uploaded with the comment “Good luck shitposter. Rolling for many dead chinks and niggers”. In the months following this attack, two shooters (John Earnest and Patrick Crusius) followed the exact same modus operandi, announcing their imminent attacks and posting their manifestos on the forum.

While 8chan’s shutting down in August 2019, a few days after Crusius’ attack, came as a relief to many, the closure of the platform seemed to trigger a migration of its users to other comparable websites. It turned out that 8chan was merely one node in a much broader, ever-changing and expanding constellation of almost identical image-boards. For the purpose of the present article, we refer to this constellation as the “chans”: indeed almost all iterations of these sites have the same visual layout, internal architecture, and a variation of the term “chan” in their name. A specific site name also includes either a number indicating its chronological distance from the two original chans (2channel [founded in 1999, renamed 5channel in 2017] and FutabaChannel [founded in 2001 with the confusing extension 2chan.net]), or a word denoting a specific focus (the “nein” in NeinChan was for example supposed to signify that the site was both the ninth iteration and had a focus on particularly fascist content). While initially limited, the chans constellation seemed to expand in the aftermath of the closure of 8chan: the likes of 8kun, 9chan, 16chan, ShitChan, EndChan, or NeinChan appeared (and sometimes shut down). All of them, including the long-standing 4chan, host “/pol” thematic boards dedicated to “politically incorrect” (a euphemism for racist and anti-Semitic) conversations. Iterations in other, non-English languages have also appeared. One such board, the /pol board of the German-language kohlchan, featured Tobias Rathjen’s video and manifesto very soon after his attack on the two shisha bars in Hanau, Germany, in February 2020, and already disseminated Stephan Balliet’s Halle shooting video (initially published on Twitch) a couple of months before.[1] While the most popular chans are accessible on the open Internet, others exist on the dark web as “onion” sites.

In spite of its offensive content and links to offline violence, this “semi-dark web” [2] far-right online constellation

remains largely unexplored. Specialized news outlets like Bellingcat have shed light on their main aspects,[3] but in-depth scientific work is almost inexistent [4] and, consequently, key questions remain unanswered. The present article examines one of the most pressing research questions: to what extent are the /pol boards of the chans similar in terms of content? In other words, do these boards host the same type of far-right content, constituting a coherent whole, or does each occupy a specific niche in a more fragmented ecosystem?

On the one hand, all these boards look interchangeable: they all have the same structure, publish highly offensive extremist far-right content (both text and visual memes), have users who write using the same characteristic lingo, and act as gateways to a range of different types of extreme-right wing platforms (archived extremist content, far-right organizations' websites, etc.). They also seem to experience regular user migrations from one to the other; 8chan, for example, grew in popularity during the “#gamergate” controversy when some 4chan users looked for an even less scrutinized forum. In other words, they all seem to contribute to the fabric of a single online “subculture”, understood here as a “culturally bounded (but not closed) network of people who come to share the meaning of specific ideas, material objects, and practices through interaction”. [5] Key to the establishment of subcultures through such interactions is the construction of a typical identity for both the ingroup and the perceived outgroup(s),[6] which occurs in various ways such as the use of new or unusual group labels to name them [7] or the glorification/vilification of individuals crystallizing the ingroup/outgroup values.[8] Under this light, the /pol chan boards would constitute yet another of the many particular subcultures populating the much broader far-right online social environment.[9]

Yet, on the other hand, a close reading of the /pol boards leaves the impression that each possesses its own identity; in other words, each /pol iteration appears to have its own subculture. For instance, whereas some appear to host favorable opinions of Brenton Tarrant, who became a cult “saint” figure, others regularly criticize the Christchurch shooter because he targeted Muslims instead of Jews. Overall, while all these /pol boards might be grouped under the very broad common umbrella of right-wing extremism, differences in ideological content and perhaps interaction practices do seem to exist.

To reconcile this apparent contradiction, we put forward—and empirically test—the hypothesis that while the /pol boards indeed partake in a common far-right online subculture, the developments recalled above have led to a fragmentation of this subculture into a series of distinct yet overlapping “sub-subcultures”. This hypothesis rests on recent work documenting “longitudinal fragmentation processes”, that is, the tendency of online political movements to fragment into self-segregating “homophilic” enclaves.[10] More precisely, we hypothesize that the mechanism driving the fragmentation of the /pol subculture is the radicalization of its members into increasingly extreme niches through forces akin to those described by McCauley and Moskalenko as “fission” (extremist organizations organically splitting up in increasingly radical groups) and “condensation” (extremist groups emerging from less radical ones following exogenous factors).[11] Our empirical analysis sought to test this hypothesis by comparing the /pol boards' content in such a way that the extremity of their views can be approximated.

Addressing this question incidentally brings about two important additional benefits. First, it allows us to assess whether the various /pol boards are equally popular, that is, feature comparable levels of discussion activity. From an outsider's perspective, it is hard to establish which of these boards are most attended and which ones are, conversely, unpopular; a problem reinforced by their contributors' full anonymity. The data collected here to compare the content of the different boards allows us to provide insight into this question, offering essential information about the overall structure of the chans' constellation. As we demonstrate below, with what we believe is the first study to include chan forums located in the dark net only, the traffic/popularity structure of the chans does in fact somewhat correspond with their ideological positioning. Second, by shedding light on an under-researched part of the wider far-right online ecosystem, our effort contributes to the growing literature on right-wing extremism on the Internet. This paper builds on recent work tracing the historical development of this ecosystem,[12] unpacking its cultural dimension and significance,[13] clarifying its relationship to radicalization,[14] and explaining processes of far-right identity construction online.[15] More specifically, it further demonstrates the depth of this ecosystem and investigates its dynamics of ideological fractionation/consolidation, prolonging important recent research on these aspects.[16]

Data and Methods

To answer our research question, we scraped the linguistic content of six chan /pol boards and analyzed it with a three-step methodology that utilizes computational text-analysis tools.[17]

Data. We developed custom software to extract and process all Original Posts (OPs) made to *4chan/pol*, *8kun/pol*, *16chan/pol*, *InfinityChan/bestpol* (also known as *9chan*), *Endchan/pol*, and *Neinchan/pol*, over the three-month period covering April, May, and June 2020.[18] Simultaneous collection during this timeframe allowed for the development of comparable corpora across the various boards (i.e., making it less likely that an external event could shape the discussion in board A collected at a time X, and not in board B collected at a time Y). Covering that 3-month timeframe—rather than a shorter period of time or a similar one at another moment—had the benefit of avoiding the content of the forums to be overtly skewed by an important event (such as the U.S. presidential elections, the initial weeks of the coronavirus outbreak, the killing of George Floyd, etc.). Innovative techniques were, in particular, developed to extract the content from *InfinityChan/bestpol* and *Neinchan/pol*, which are located on the Dark Net. These six boards were selected for two main reasons: first because they were recently identified as the most important ones to include by a range of government units in charge of online extremism,[19] and second because these were the most active of the several /pol boards our team was monitoring and scraping at the time of data collection (as **Figure 1** will make clear, other boards did not demonstrate a sufficient level of user engagement to motivate inclusion). As a result, our work does not claim to investigate a representative sample of the /pol ecosystem, but rather to focus the analysis on the most relevant of them.

Both **Table 1** and **Figure 1** below provide key descriptive statistics on the overall level of posting activity for the six /pol boards. Based on the number of OPs made per day, *4chan/pol* remains by far the most popular board, with OP traffic being almost twice as big as that of all other boards combined. *8kun/pol*, which is the direct successor to the now defunct *8chan/pol*, comes second, with all the other boards characterized by low traffic. The structure of the /pol boards constellation is therefore a 3-tiered hierarchy with one popular board, one less popular board, and many poorly attended boards.

The graph demonstrates a decrease in activity for *4chan/pol* from the beginning of June; however, a more in-depth investigation into the typical level of *4chan/pol* traffic, with data from further back in time, reveals that the high posting activity observed in April and May corresponds to an unusual spike in pace of discussion due to the COVID-19 lockdown. As such, the observed decline represents the level of discussion returning to pre-COVID levels. We also notice that *8kun/pol*'s popularity seems to have slowly increased over the period under scrutiny; given that forum's extreme content (cf. below), this is a worrying trend.

Table 1: Descriptive statistics of the six corpora over the period 1 April 2020 – 30 June 2020.

	<i>4chan/pol</i>	<i>8kun/pol</i>	<i>16chan/pol</i>	<i>Endchan/pol</i>	<i>Infinitychan/bestpol</i> (Dark Net)	<i>Neinchan/pol</i> (Dark Net)
Total number of Original Posts (OPs)	8393	978	123	51	334	63
Average number of OPs per day	92.23	10.74	1.35	0.56	3.67	0.69
Total number of words	378,051	102,640	19,158	14,599	47,598	7,985
Average of average word length per OP	45.04	104.94	155.75	286.25	142.50	126.74

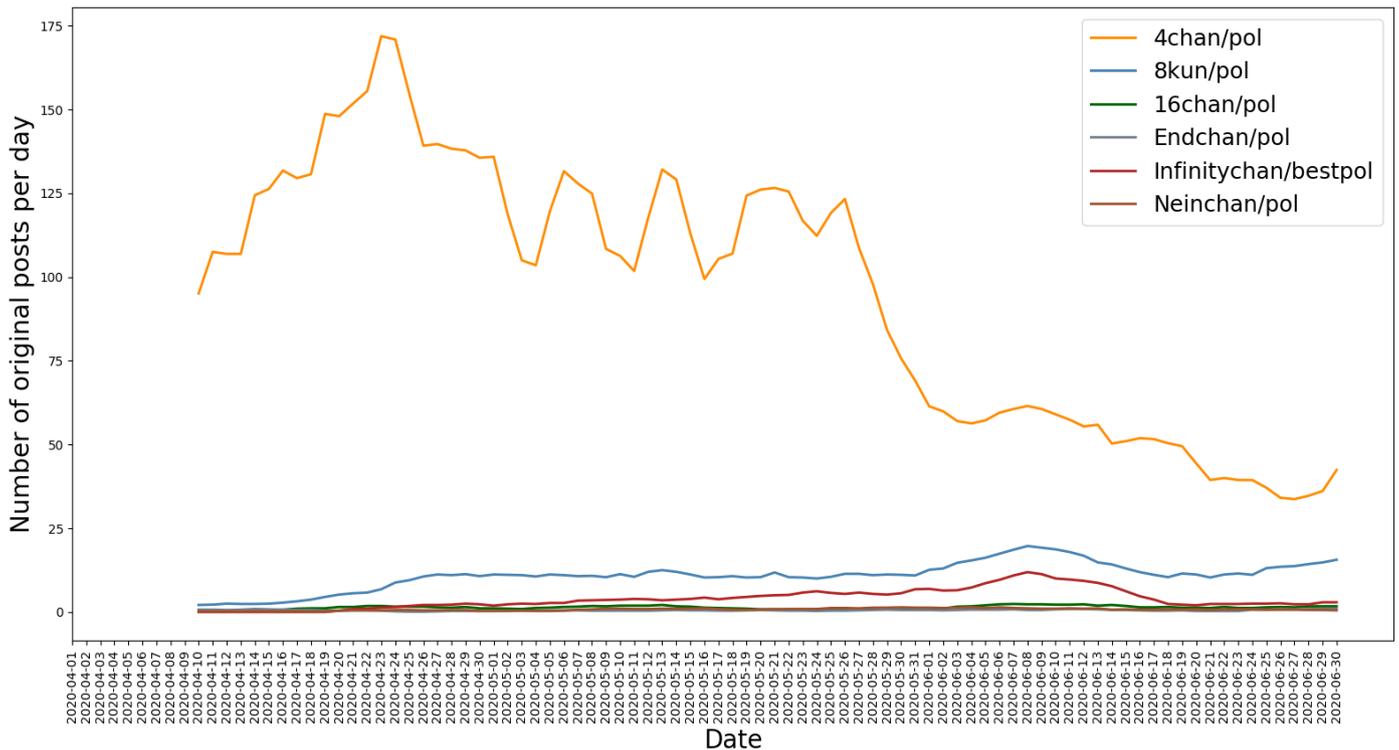


Figure 1: Posting activity in the six /pol boards over the period 1 April 2020 – 30 June 2020. The number of posts presented here is shown as a 10-day moving average in order to make the posting behavior on the less popular chans amenable to study. This is due to the fact that some data is missing for some of the days in early-to-mid April for 4chan due to web driver issues. However, the decline in 4chan traffic from the end of May is due to lower traffic and not caused by these missing data.

Method

To evaluate the similarity between these boards’ content, we proceed in three steps. First, we identified, for each board, its 20 most frequent nouns and bigrams (sequences of two words that frequently co-occur in the corpus, such as “white house”). Comparing the relative frequency of the boards’ prominent nouns and bigrams provides a first indication of the main preoccupations of the posters and may already highlight differences in emphases.

Second, we conducted a correspondence analysis of the boards’ most frequent terms, which visually represents how specific these terms are to the various boards. Correspondence analyses, which are built from contingency tables (wherein entities are categorized into two or more different sets of categories), are a standard technique to visualize semantic differences between long texts and corpora [20]—indeed it has been used in endeavors as diverse as evaluating the evolution of Dickens’ style,[21] settle the disputed authorship of Gospels,[22] or, more pertinently for us, compare the content of Islamic State’s *Dabiq* and *Rumiyah* magazines.[23] We plotted the 200 most frequent nouns for the six boards in a contingency table (a 2- dimensional table with each row corresponding to a frequent word and the frequency counts for each /pol board appearing in the columns), and plotted on that basis a two-dimensional scatter diagram to visualize the results. As Kroonenberg and Greenacre explain, such “graphical representations of the row and column variables [...] highlight the salient aspects of their relationship”, in our case the deviations between the lexical fields of the various boards.[24]

Third, we generated four co-occurrence networks; one for *4chan/pol*, one for *8kun/pol*, one for *16chan/pol*, and one for *Endchan/pol* (the rationale for this selection is explained below). Co-occurrence networks, also known as semantic networks, are generated from co-occurrence matrix tables. They visually represent both at once the most frequent words of the text as nodes (whose sizes vary according to their relative frequencies) and the probability that these terms co-occur in the same sentence or paragraph in the text (represented as more or less thick and distant links connecting the nodes). This method allows for a more granular representation of each board’s lexical field, exposing the texts’ “interconcept organization” (or “associative network”, that is, the way

the concepts composing the texts take their meaning from their regular association with other ones).[25] These advantages make co-occurrence networks useful tools to identify the major frames, narratives, and discourses of political texts,[26] including violent extremists’ prose.[27] These networks, as well as the correspondence analysis for step 2, were generated with the KH Coder software.[28]

Findings and Discussion

Step 1: Most frequent terms and bigrams. Table 2 below displays the most frequent nouns for each of the six boards. A number of features are immediately obvious. First, the most popular nouns for all the boards concern the phrase “white people”, with “jew” or labels depicting black people (“black”, “nigger”) also being frequently used, supportive of other work that has discussed the in-group and out-group categories of the far-right online environment [29]. This is indicative of ingroup and outgroup labels being common to the various boards, and hence evidence a common subculture based on the racial distinction between white people and others.

Second, we see that compared to *4chan/pol*, *8kun/pol* appears to have more of a racial component to its discussions: the word “black” is third most frequent (compared to 11th), and the terms “jews”, “nigger”, and “race” appear in the table (none of them count among *4chan/pol*’s 20 most frequent words). Third, the table seems to suggest that *NeinChan* hosts the most extreme content, which is perhaps not surprising given it exists solely on the dark web. A cursory look at this particular board demonstrates that users tend to discuss matters directly related to violent attacks, such as comparing the manifestos of various shooters (with Brenton Tarrant even having his own board aside from the /pol one). These results indicate that not all these boards are equally extreme—rather, they suggest that the more established and longstanding *4chan/pol* hosts less extreme discussions, which stands in line with our fragmentation hypothesis.

Table 2: Most frequent nouns on the six /pol boards over the period 1 April 2020 – 30 June 2020 (absolute count and ratio for the whole corpus).

	<i>4chan/pol</i>	<i>8kun/pol</i>	<i>16chan/pol</i>	<i>Endchan/pol</i>	<i>Infinitychan/bestpol</i>	<i>Neinchan/pol</i>
1	“people”: Frequency: 1580 Ratio: 0.0079	“people”: Frequency: 431 Ratio: 0.0076	“white”: Frequency: 120 Ratio: 0.0117	“people”: Frequency: 64 Ratio: 0.0097	“people”: Frequency: 168 Ratio: 0.0065	“white”: Frequency: 56 Ratio: 0.0130
2	“white”: Frequency: 1059 Ratio: 0.0053	“white”: Frequency: 389 Ratio: 0.0069	“people”: Frequency: 68 Ratio: 0.0066	“white”: Frequency: 50 Ratio: 0.0075	“white”: Frequency: 147 Ratio: 0.0057	“thread”: Frequency: 54 Ratio: 0.01262
3	“pol”: Frequency: 677 Ratio: 0.0033	“black”: Frequency: 189 Ratio: 0.0033	“group”: Frequency: 68 Ratio: 0.0066	“jew”: Frequency: 38 Ratio: 0.0057	“us”: Frequency: 111 Ratio: 0.0043	“tarrant”: Frequency: 56 Ratio: 0.0107
4	“trump”: Frequency: 593 Ratio: 0.0029	“time”: Frequency: 188 Ratio: 0.0033	“state”: Frequency: 37 Ratio: 0.0036	“race”: Frequency: 34 Ratio: 0.0051	“time”: Frequency: 96 Ratio: 0.0037	“meme”: Frequency: 41 Ratio: 0.0095
5	“world”: Frequency: 592 Ratio: 0.0029	“post”: Frequency: 171 Ratio: 0.0030	“time”: Frequency: 34 Ratio: 0.0033	“jewish”: Frequency: 27 Ratio: 0.0041	“thread”: Frequency: 84 Ratio: 0.0032	“brenton”: Frequency: 38 Ratio: 0.00888
6	“view”: Frequency: 575 Ratio: 0.0028	“jew”: Frequency: 171 Ratio: 0.0030	“jew”: Frequency: 34 Ratio: 0.0033	“control”: Frequency: 27 Ratio: 0.0041	“world”: Frequency: 65 Ratio: 0.0025	“jew”: Frequency: 56 Ratio: 0.0130

7	“time”: Frequency: 556 Ratio: 0.0027	“state”: Frequency: 152 Ratio: 0.0027	“year”: Frequency: 31 Ratio: 0.0030	“life”: Frequency: 26 Ratio: 0.0039	“niggers”: Frequency: 61 Ratio: 0.0023	“world”: Frequency: 20 Ratio: 0.0046
8	“text”: Frequency: 545 Ratio: 0.0027	“view”: Frequency: 149 Ratio: 0.0026	“world”: Frequency: 27 Ratio: 0.0026	“woman”: Frequency: 26 Ratio: 0.0039	“black”: Frequency: 60 Ratio: 0.0023	“people”: Frequency: 17 Ratio: 0.0039
9	“click”: Frequency: 524 Ratio: 0.0026	“nigger”: Frequency: 149 Ratio: 0.0026	“life”: Frequency: 27 Ratio: 0.0026	“state”: Frequency: 23 Ratio: 0.0034	“power”: Frequency: 57 Ratio: 0.0022	“man”: Frequency: 17 Ratio: 0.0039
10	“country”: Frequency: 575 Ratio: 0.0028	“text”: Frequency: 146 Ratio: 0.0025	“thread”: Frequency: 26 Ratio: 0.0025	“system”: Frequency: 23 Ratio: 0.0034	“post”: Frequency: 56 Ratio: 0.0021	“year”: Frequency: 13 Ratio: 0.0030
11	“black”: Frequency: 505 Ratio: 0.0025	“race”: Frequency: 145 Ratio: 0.0025	“leftist”: Frequency: 26 Ratio: 0.0025	“thing”: Frequency: 17 Ratio: 0.0024	“pol”: Frequency: 52 Ratio: 0.0020	“thing”: Frequency: 13 Ratio: 0.0030
12	“shit”: Frequency: 495 Ratio: 0.0024	“police”: Frequency: 143 Ratio: 0.0025	“war”: Frequency: 24 Ratio: 0.0023	“virus”: Frequency: 16 Ratio: 0.0024	“whites”: Frequency: 51 Ratio: 0.0020	“site”: Frequency: 12 Ratio: 0.0028
13	“going”: Frequency: 489 Ratio: 0.0024	“thread”: Frequency: 139 Ratio: 0.0024	“thing”: Frequency: 24 Ratio: 0.0023	“german”: Frequency: 16 Ratio: 0.0024	“war”: Frequency: 49 Ratio: 0.0019	“faggot”: Frequency: 12 Ratio: 0.0028
14	“year”: Frequency: 454 Ratio: 0.0022	“click”: Frequency: 135 Ratio: 0.0024	“violence”: Frequency: 24 Ratio: 0.0023	“post”: Frequency: 15 Ratio: 0.0022	“police”: Frequency: 49 Ratio: 0.0019	“race”: Frequency: 12 Ratio: 0.0028
15	“day”: Frequency: 445 Ratio: 0.0022	“world”: Frequency: 128 Ratio: 0.0022	“men”: Frequency: 23 Ratio: 0.0022	“place”: Frequency: 14 Ratio: 0.0021	“way”: Frequency: 48 Ratio: 0.0018	“anons”: Frequency: 9 Ratio: 0.0021
16	“life”: Frequency: 431 Ratio: 0.0021	“video”: Frequency: 127 Ratio: 0.0022	“trump”: Frequency: 23 Ratio: 0.0022	“symptom”: Frequency: 14 Ratio: 0.0021	“going”: Frequency: 46 Ratio: 0.0018	“men”: Frequency: 9 Ratio: 0.0021
17	“china”: Frequency: 417 Ratio: 0.0020	“group”: Frequency: 119 Ratio: 0.0021	“control”: Frequency: 23 Ratio: 0.0022	“cause”: Frequency: 14 Ratio: 0.0021	“jewish”: Frequency: 46 Ratio: 0.0018	“manifesto”: Frequency: 9 Ratio: 0.0021
18	“woman”: Frequency: 387 Ratio: 0.0019	“shit”: Frequency: 119 Ratio: 0.0021	“month”: Frequency: 21 Ratio: 0.0020	“mind”: Frequency: 14 Ratio: 0.0021	“race”: Frequency: 45 Ratio: 0.0017	“pol”: Frequency: 9 Ratio: 0.0021
19	“thing”: Frequency: 380 Ratio: 0.0019	“trump”: Frequency: 118 Ratio: 0.0021	“going”: Frequency: 20 Ratio: 0.0019	“coronavirus”: Frequency: 13 Ratio: 0.0019	“system”: Frequency: 44 Ratio: 0.0017	“creator”: Frequency: 9 Ratio: 0.0021
20	“way”: Frequency: 379 Ratio: 0.0018	“thing”: Frequency: 116 Ratio: 0.0020	“medium”: Frequency: 20 Ratio: 0.0019	“number”: Frequency: 13 Ratio: 0.0019	“jews”: Frequency: 42 Ratio: 0.0016	“movement”: Frequency: 8 Ratio: 0.0018

Table 3 provides the most frequent bigrams and is generally consistent with the analysis provided in **Table 2**. Namely, while race is a prominent theme indicating a common subculture across all boards, the more esoteric chans have more extreme content demonstrating fragmentation. *4chan/pol* indeed appears to be more concerned with a broader range of political issues such as Donald Trump and COVID-19, the issue of race seems much more prominent on *8kun/pol* and other boards, and terms indicative of conspiratorial narratives (e.g. “jew” + “control”) and slurs (e.g. “cock” + “faggotity”, “naughty” + “cock”, “piss” + “earth”) appear to characterize the fringe ones. Using the concept of “whiteness” as an example, one can indeed see a gradual increase in radical views, starting with *4chan/pol* on the left mentioning “white people”, to the more racial focus of *8kun/pol*, to *16chan/pol* mentioning “white nationalist”, to “national socialist” on *InfinityChan/bestpol* board, finally ending with open discussions of Brenton Tarrant on *NeinChan/pol*. Overall, the first step of our analysis offers preliminary evidence in favor of our analysis.

Table 3: Most frequently occurring bigrams on the six /pol boards over the period 1 April 2020—30 June 2020.

	<i>4chan/pol</i>	<i>8kun/pol</i>	<i>16chan/pol</i>	<i>Endchan/pol</i>	<i>Infinitychan/bestpol</i>	<i>Neinchan/pol</i>
1	“covid”, “19”: 156	“white”, “people”: 62	“white”, “people”: 9	“white”, “people”: 7	“right”, “wing”: 20	“brenton”, “tarrant”: 38
2	“white”, “people”: 110	“united”, “state”: 44	“white”, “nationalist”: 9	“al”, “moallem”: 7	“dont”, “care”: 18	“tarrant”, “meme”: 32
3	“pic”, “related”: 86	“white”, “race”: 27	“social”, “medium”: 9	“neurological”, “symptom”: 6	“social”, “media”: 17	“meme”, “thread”: 32
4	“swine”, “flu”: 65	“black”, “life”: 24	“united”, “state”: 6	“jew”, “control”: 6	“national”, “socialist”: 17	“naughty”, “cock”: 8
5	“sequenced”, “live”: 65	“life”, “matter”: 24	“white”, “civilization”: 6	“span”, “n”: 6	“united”, “states”: 13	“white”, “man”: 8
6	“people”, “lockdown”: 64	“year”, “old”: 24	“non”, “white”: 6	“face”, “mask”: 6	“iov”, “cn”: 13	“white”, “race”: 6
7	“donald”, “j”: 62	“anti”, “white”: 22	“living”, “soul”: 6	“syrian”, “people”: 6	“anymore”, “dont”: 13	“basis”, “morality”: 5
8	“hasnt”, “done”: 62	“make”, “sure”: 21	“6”, “million”: 6	“united”, “state”: 5	“wh”, “iov”: 12	“non”, “white”: 5
9	“done”, “anything”: 61	“social”, “medium”: 19	“right”, “wing”: 6	“april”, “30”: 5	“anti”, “white”: 12	“u”, “army”: 4
10	“j”, “trump”: 61	“george”, “floyd”: 19	“anti”, “white”: 5	“minneapolis”, “police”: 5	“murdoch”, “murdoch”: 11	“race”, “war”: 4
11	“trump”, “potus”: 61	“national”, “guard”: 19	“fuck”, “right”: 5	“non”, “white”: 5	“shopping”, “cart”: 11	“dont”, “know”: 4
12	“trump”, “hasnt”: 61	“various”, “location”: 19	“jewish”, “american”: 5	“social”, “medium”: 5	“dont”, “want”: 11	“cock”, “faggotity”: 4
13	“potus”, “teamtrump”: 60	“survival”, “bread”: 18	“white”, “supremacist”: 5	“9”, “11”: 5	“white”, “people”: 10	“human”, “conscience”: 4
14	“teamtrump”, “trumpwarroom”: 60	“cvg”, “survival”: 18	“right”, “winger”: 5	“neurological”, “disorder”: 5	“even”, “though”: 10	“white”, “world”: 4
15	“president”, “donald”: 60	“yet”, “another”: 17	“group”, “people”: 4	“national”, “socialism”: 4	“national”, “socialists”: 10	“piss”, “earth”: 4
16	“black”, “people”: 59	“terrorist”, “organization”: 16	“deutsch”, “pol”: 4	“police”, “department”: 4	“white”, “man”: 9	“discussion”, “meme”: 3

One observation appears that strengthens the findings from step 1, further backing our hypothesis: words that represent a more openly xenophobic and confrontational type of far-right language—such as “nigger”, “jew”, “white”, “war”, “race”, or “power”—are located at the right side of the plot, where all the third-tiered /pol boards are also located, including the two Dark Web boards of *InfinityChan* and *NeinChan*. *4chan/pol*, albeit without a doubt a far-right forum, is situated at the other side of the graph, surrounded by less contentious words that denote the broader interest in political news (e.g. “president”, “money”, “government”, “coronavirus”, “country”). This corroborates our earlier suggestion that the content of the minor chans is more extreme than that of *4chan/pol*, with *8kun/pol* situated in between the two groups: in other words, it appears that the 3-tier structural hierarchy of the boards (in terms of their posting frequency) corresponds to increased ideological extremism, in line with our working hypothesis. This said, while *16chan/pol*, *InfinityChan/bestpol*, and *NeinChan/pol* occupy a comparable position, *EndChan/pol* is located very far from them in the graph, with very few of the most common words located near it; this means that its lexical field contains specificities that make it stand out from that of the other boards (as we foresaw in Table 2).

Step 3: Co-occurrence networks. We extend our previous analysis by generating and analyzing four co-occurrence networks (Figures 3, 4, 5, and 6). We estimate a separate network for *4chan/pol* and *8kun/pol*, given their respective status as first and second most visited boards, but only plotted two networks for the remaining smaller boards for the sake of clarity and pertinence. From these marginal chans, we selected *16chan/pol* because it stands at the center of the semantic cluster (in Figure 2) occupied by *NeinChan/pol* and *InfinityChan/bestpol*, and *Endchan/pol* because of its apparent oddity revealed by the first two steps of the analysis. While these graphs’ rich content offers multiple analytical insights, we can only focus here on what we believe are the two most crucial findings in terms of our research question and hypothesis.

First, our main observations from steps 1 and 2 are further confirmed: *4chan/pol* appears to display less extreme content than the other boards. Its most central discussion theme (green cluster) is the coronavirus situation, and while there are political discussions (blue cluster on Trump and political ideologies; small clusters on news events such as the Black Lives Matter protests) as well as a concern for race indicative of a common far-right subculture (e.g. black-white top cluster), brazenly racist lingo is absent and the most extreme right-wing concerns are not prominent. On the contrary, *8kun/pol*’s central cluster contains that language (“chimpout” [a derogatory term for the Black Lives Matter movement], “nigger”, “officer” co-occurring with “ass”) and features important far-right themes (see for example the right side of the purple cluster with “hitler”, “propaganda”, “struggle”, “control”, or “socialism”) and framing (see for example the co-occurrence of “black” with “cop”), together with indications that far-right conspiracy theories are talked about (“bill” – “gates”, “murdoch” – “channel”). Trump and U.S. politics feature less centrally. *16kun/pol*’s discussions are less clearly clustered, indicating that themes are more mixed up in discussions; moreover, xenophobia permeates the content of all these amalgamated debates. For example, the blue cluster on the left shows a concern for “white” “civilization”. That board appears to have a more anti-Semitic content than the others, with words such as “zionist”, “kosher”, “israel”, “kike” [a derogatory slang term for Jews], or “holocaust” spread across the network, which indicates a specific “sub-subculture”. In sum, *8kun/pol* and *16kun/pol* are both more extreme than *4chan/pol*, but while *8kun/pol* is marked by white supremacist concerns, *16kun/pol* presents a more anti-Semitic tone.

Second, the peculiarity of *EndChan* further appears. While the forum does contain far-right anti-Semitic discussions (see for example the green cluster on the left-hand side, with “trump” co-occurring with “goyim” and “jew” with the discussions about the “looting” and “burning” of the protests [following the death of George Floyd]), the forum seems to be an eclectic aggregation of different discussions about themes as diverse as ancient European culture (“greek”, “culture”, “history”, “art”, “light”, “paint”, etc.), the 5G conspiracy (“5g”, “phone”), and the coronavirus (blue cluster, bottom right). This heterogeneous assemblage of different themes, with recurring far-right concerns, is partly due to the low popularity of the forum, which hasn’t succeeded in constituting a solid user base and coherent “identity”.

Overall, this last step’s granular analysis delivers results that confirm those of the first two steps, and so largely confirms our working hypothesis: while all the /pol boards together contribute to a common online far-right subculture structured by the racist opposition between “white” people and their “enemies”, clear differences

exist that evidence the fragmentation of this particular ecosystem—through fission/condensation—of this subculture into a range of ever-more extremist niches, or “sub-subcultures”, each with its particular ideological nuances and linguistic practices.

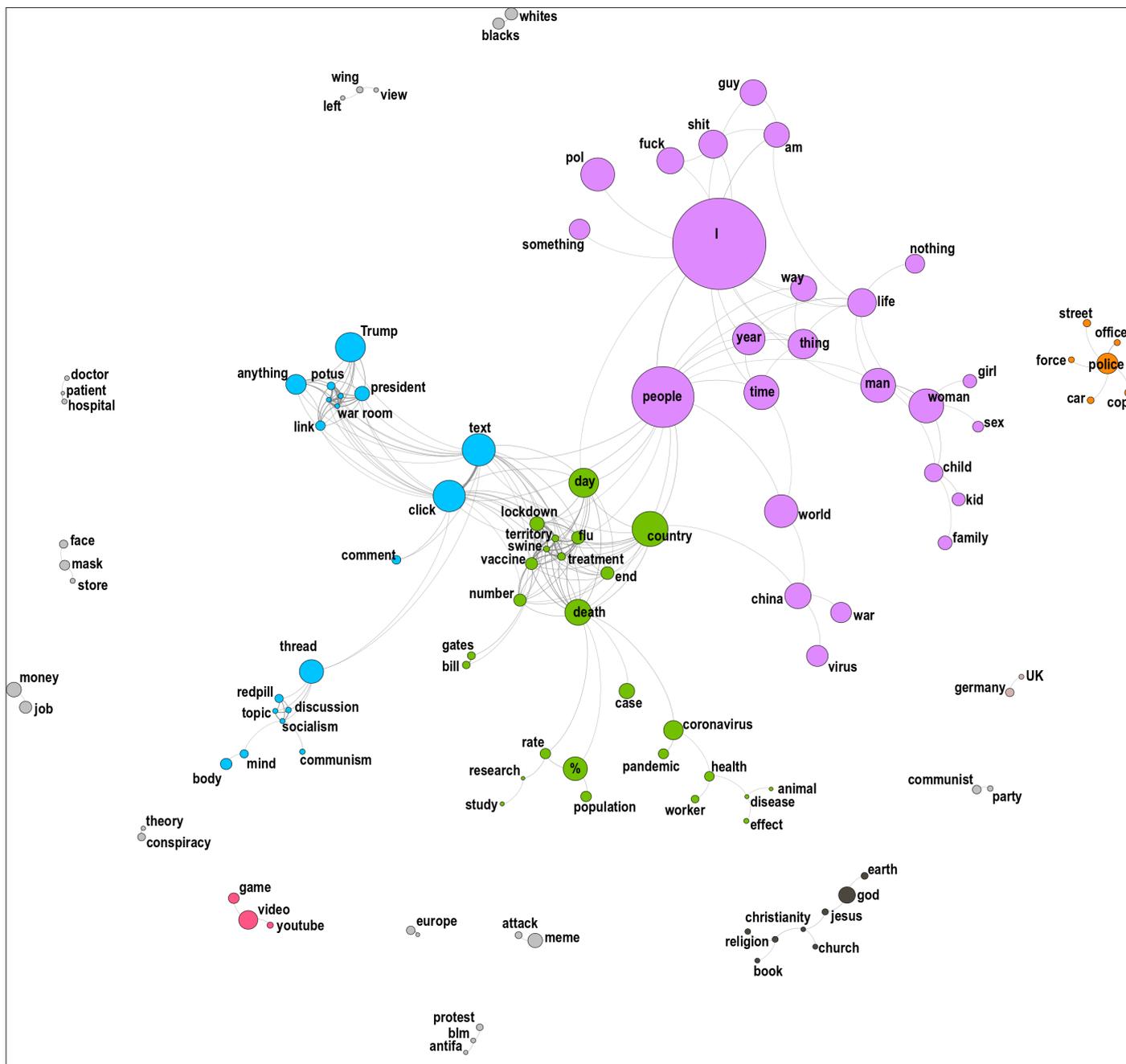


Figure 3: Co-occurrence network for 4chan/pol.

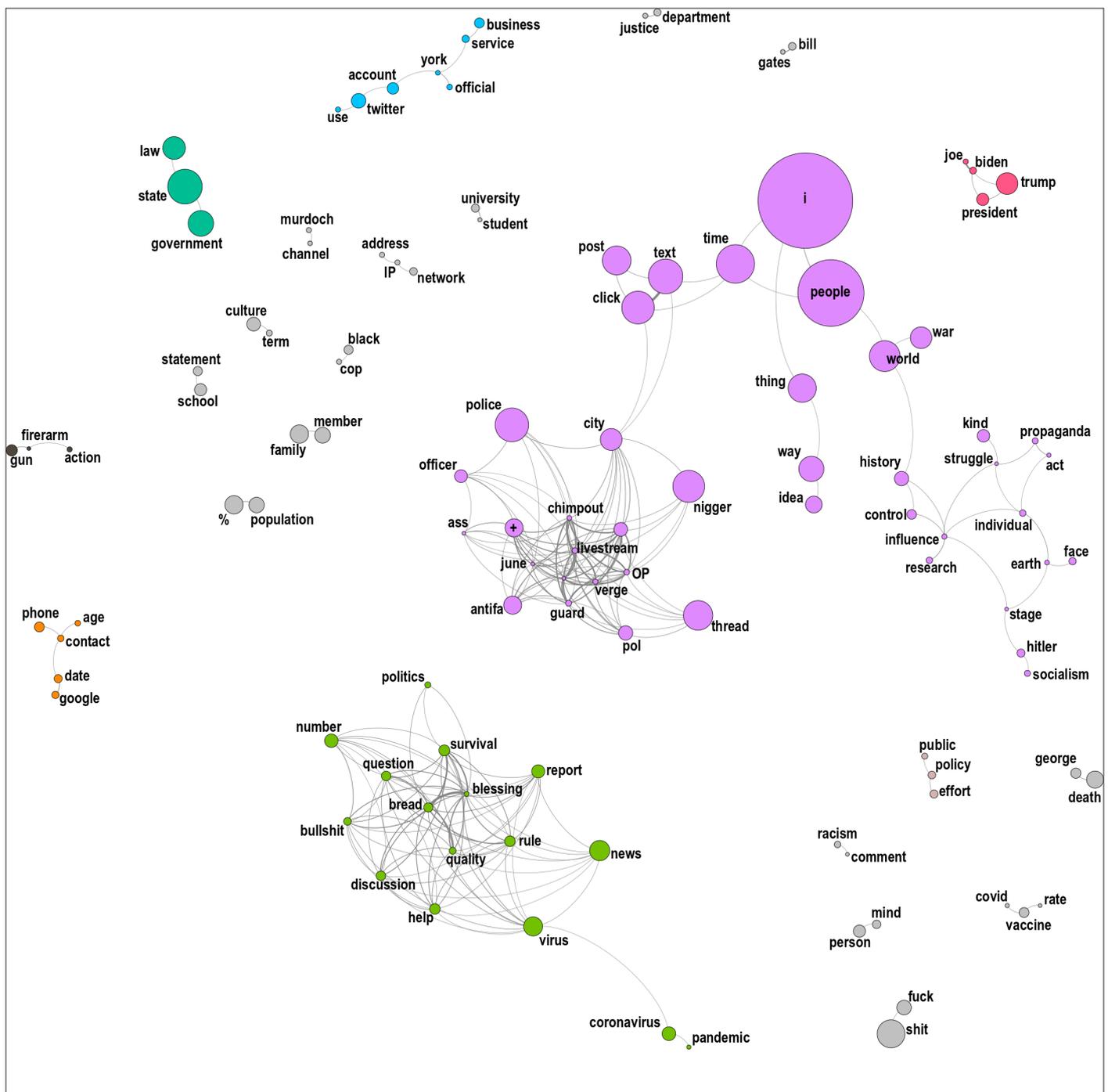


Figure 4: Co-occurrence network for 8kun/pol.

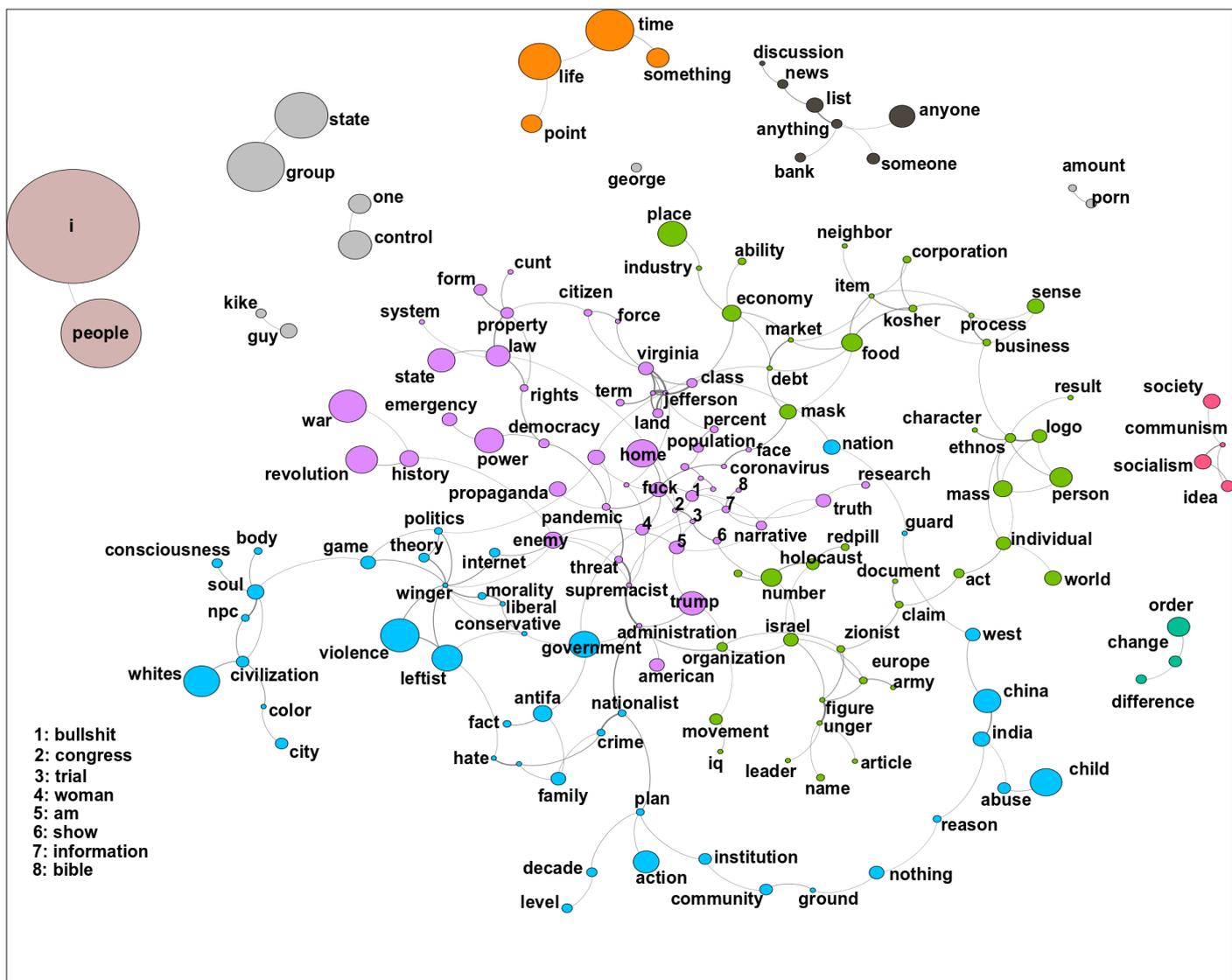


Figure 5: Co-occurrence network for 16chan/pol.

far the most popular /pol board at the top of the three-tiered hierarchy wherein the newer, more esoteric chans occupy a fringe position, hosting much fewer—but clearly more extreme—participants.

Whether that pattern is a common evolution of extremist subcultures on the Internet, or only a specific feature of the chans' ecosystem, is a crucial question with important policy implications that ought to be answered by further research. As “alt-tech” platforms such as Gab, Discord, Slug, or BitChute pull more and more users from increasingly strict companies like Facebook, Twitter, or YouTube, investigations of extremist community formation should evaluate whether the same dynamics also hold there.

Acknowledgement

This work was supported by Economic and Social Research Council (ESRC) via the Center for Research and Evidence on Security Threats (CREST) consortium under grant number ES/N009614/1.

About the Authors:

Stephane J. Baele, Ph.D., is Senior Lecturer in International Relations at the University of Exeter, U.K., where he is the deputy director of the Centre for Advanced International Studies (CAIS). His multidisciplinary research, which focuses on the linguistic and visual dimensions of extremist communications (as well as on IR theories), has appeared in *Political Psychology*, *Journal of Conflict Resolution*, *International Studies Review*, *Terrorism & Political Violence*, *Journal of Language & Social Psychology*, *International Theory*, and *Journal of Global Security Studies*, among others.

Lewys Brace, Ph.D., is Lecturer in Data Analysis at the University of Exeter, where he is part of the Q-Step Centre. His research focuses on online extremism and radicalization in regards to political violence, as well as the application of computational methods to social science research and policing. His research has previously appeared in *Terrorism & Political Violence* and *Artificial Life*.

Travis G. Coan, Ph.D., is a senior lecturer in Quantitative Politics at the University of Exeter and codirector of the Exeter Q-Step Centre. While Dr. Coan has published on a wide range of topics, his core research examines questions in the areas of environmental and political communication, with an interest in the study of polarization and extremism. His work has appeared in *International Studies Quarterly*, *Social Networks*, *Studies in Conflict and Terrorism*, and *Terrorism & Political Violence*, among others.

Notes

[1] Read for example *Le Monde's* articles dedicated to these shooters' online interventions; URL: https://www.lemonde.fr/pixels/article/2020/02/20/alle-magne-le-terroriste-presume-de-hanau-avait-publie-un-manifeste-conspirationniste-et-raciste_6030204_4408996.html, or https://www.lemonde.fr/pixels/article/2019/10/10/attentat-de-halle-un-manifeste-qui-pose-plus-de-questions-qu-il-n-apporte-de-reponses_6014973_4408996.html, among others.

[2] Maryam Zamani, Fereshteh Rabbani, Attila Horicsányi, Anna Zafeiris, and Tamas Vicsek, “Differences in Structure and Dynamics of Networks Retrieved from Dark and Public Web Forums,” *Physica A*, no. 525 (2019), pp. 326–336.

[3] See for example: Robert Evans, “Shitposting, inspirational terrorism, and the Christchurch Mosque Massacre,” *Bellingcat*, 15 March 2019; URL: <https://www.bellingcat.com/news/rest-of-world/2019/03/15/shitposting-inspirational-terrorism-and-the-christchurch-mosque-massacre/> [accessed 16/03/2019]; Robert Evans, “Ignore The Poway Synagogue Shooter's Manifesto: Pay Attention To 8chan's /pol/ Board,” *Bellingcat*, 28 April 2019; URL: <https://www.bellingcat.com/news/americas/2019/04/28/ignore-the-poway-synagogue-shooters-manifesto-pay-attention-to-8chans-pol-board/> [accessed 16/03/2019]; Robert Evans, “The El Paso Shooting and the Gamification of Terror,” *Bellingcat*, 4 August 2019; URL: <https://www.bellingcat.com/news/americas/2019/08/04/the-el-paso-shooting-and-the-gamification-of-terror/>.

[4] For exceptions, see for example D. Ludemann, “/pol/emics: Ambiguity, Scales, and Digital Discourse on 4chan,” *Discourse, Context & Media*, vol. 24 (2018), pp. 92–98; Florence Keen, Blyth Crawford, and Guillermo Suarez de-Tangil, “Memetic Irony And The Promotion Of Violence Within Chan Cultures”, CREST Report, 15 December 2020, URL: <https://crestresearch.ac.uk/resources/>

[memetic-irony-and-the-promotion-of-violence-within-chan-cultures/](#).

[5] J. Patrick Williams, and Heith Copes, “How Edge Are You? Constructing Authentic Identities and Subcultural Boundaries in a Straightedge Internet Forum,” *Symbolic Interaction* vol. 28 no. 1 (2005), p. 70.

[6] Idem, pp. 67–89. Also Susan Zickmund, “Approaching the Radical Other: The Discursive Culture of Cyberhate,” in David Bell and Barbara Kennedy, eds., *The Cybercultures Reader* (London: Routledge, 2000), pp. 137–253.

[7] Holt Thomas, Blevins Kristie, and Burkert Natasha, “Considering the Pedophile Subculture Online,” *Sexual Abuse*, vol. 22 no. 1 (2010), pp. 3–24.

[8] See for example how the online “incel” subculture rests on the pseudo-canonization of attackers such as Elliot Rodger or Alek Minassian: Stephane J. Baele, Lewys Brace, and Travis G. Coan, “From ‘Incel’ to ‘Saint’: Analyzing the violent worldview behind the 2018 Toronto attack,” *Terrorism & Political Violence*, online before print (2019); URL: <https://doi.org/10.1080/09546553.2019.1638256>; or Taisto Witt, “If I Cannot Have It, I Will Do Everything I Can to Destroy It. The Canonization of Elliot Rodger: ‘Incel’ Masculinities, Secular Sainthood, and Justifications of Ideological Violence,” *Social Identities*, online before print (2020); URL: <https://doi.org/10.1080/13504630.2020.1787132>.

[9] Such as those already described by John M. Cotter, “Sounds of Hate: White Power Rock and Roll and the Neo-Nazi Skinhead Subculture,” *Terrorism & Political Violence*, vol. 11 no. 2 (1999), pp. 111–140; or Zickmund, op. cit.

[10] Read Deen Freelon, Marc Lynch, and Sean Aday, “Online Fragmentation in Wartime: A Longitudinal Analysis of Tweets about Syria, 2011–2013,” *The ANNALS of the American Academy of Political and Social Science*, vol. 659 no. 1 (2015), pp. 166–179.

[11] Clark McCauley, and Sophia Moskalkenko, “Mechanisms of Political Radicalization: Pathways Toward Terrorism,” *Terrorism & Political Violence*, vol. 20 no. 3 (2008), pp. 415–433.

[12] See for example Maura Conway, Ryan Scrivens, and Logan Macnair, “Right-Wing Extremists’ Persistent Online Presence: History and Contemporary Trends,” *ICCT Policy Briefs* (2019); URL: <https://icct.nl/publication/right-wing-extremists-persistent-online-presence-history-and-contemporary-trends/>.

[13] Maik Fielitz and Nick Thurston, *Post-Digital Cultures of the Far Right. Online Actions and Offline Consequences in Europe and the US* (Bielefeld: Transcript Verlag, 2019).

[14] Tiana Gaudette, Ryan Scrivens, and Vivek Venkatesh, “The Role of the Internet in Facilitating Violent Extremism: Insights from Former Right-Wing Extremists,” *Terrorism & Political Violence*, online before print (2020); URL: <https://doi.org/10.1080/09546553.2020.1784147>.

[15] Consider, for example, the following studies of identity construction dynamics in the neo-Nazi forum “Stormfront”: Lorraine Bowman-Grieve, “Exploring ‘Stormfront’: A Virtual Community of the Radical Right,” *Studies in Conflict & Terrorism*, vol. 32 no. 11 (2009), pp. 989–1007; Willem De Koster, and Dick Houtman, “Stormfront is like a Second Home to Me’. On Virtual Community Formation by Right-wing Extremists,” *Information, Communication & Society*, vol. 11 no. 8 (2008), pp. 1155–1176.

[16] See for example Bennett Kleinberg, Isabelle van der Vegt, and Paul Gill, “The Temporal Evolution of a Far-right Forum,” *Journal of Computational Social Science*, online before print (2020); URL: <https://doi.org/10.1007/s42001-020-00064-x>; Jacob Davey and Julia Ebner, *The Fringe Insurgency. Connectivity, Convergence and Mainstreaming of the Extreme Right* (London: ISD, 2017); Manuela Caiani and Linda Parenti, *European and American Extreme Right Groups and the Internet* (New York: Routledge, 2016); Caterina Froio, “Race, Religion, or Culture? Framing Islam between Racism and Neo-Racism in the Online Network of the French Far Right,” *Perspectives on Politics* vol. 16 no. 3 (2018), pp. 696–709; or Caterina Froio, and Bharat Ganesh, “The Transnationalisation of Far Right Discourse on Twitter,” *European Societies* vol. 21 no. 4 (2019), pp. 513–539.

[17] Computational text-analysis methods are now a well-established method for the study of extremist content. See for example Shuki Cohen, Arie Kruglanski, Michele Gelfand, David Webber, and Rohan Gunaratna, “Al-Qaeda’s Propaganda Decoded: A Psycholinguistic System for Detecting Variations in Terrorism Ideology, Terrorism and Political Violence,” vol. 30 no. 1 (2018), pp. 142–171; Sylvia Jaki, Tom De Smedt, Maja Gwóźdz, Rudresh Panchal, Alexander Rossa, Guy De Pauw, “Online Hatred of Women in the Incels.me Forum: Linguistic Analysis and Automatic Detection,” *Journal of Language Aggression and Conflict*, vol. 7 no. 2 (2019), pp. 240–268. Baele, Brace, and Coan, op. cit. On far-right content, see for example Shuki Cohen, Thomas Holt, Seven Chermak, and Joshua Freilich, “Invisible Empire of Hate: Gender Differences in the Ku Klux Klan’s Online Justifications for Violence,” *Violence & Gender*, online before print (2020); URL: <https://doi.org/10.1089/vio.2017.0072>.

[18] The scraper that was developed by the authors to collect this data was developed in Python using a mixture of the *Requests* and *Selenium* packages and is available, along with user documentation, on the authors’ GitHub page.

[19] The source of this information can unfortunately not be disclosed.

[20] Pieter Kroonenberg, and Michael Greenacre, “Correspondence Analysis”; in: Samuel Kotz, Campbell Read, N. Balakrishnan, and Brani Vidakovic, (Eds.), *Encyclopedia of Statistical Sciences* (2006); URL: <https://doi.org/10.1002/0471667196.ess6018.pub2>. See also Phillip M. Yelland, “An Introduction to Correspondence Analysis,” *The Mathematica Journal*, vol. 12 (2010); URL: <https://content.wolfram.com/uploads/sites/19/2010/09/Yelland.pdf>.

- [21] Tomoji Tabata, "Investigating Stylistic Variation in Dickens through Correspondence Analysis of Word-Class Distribution," in Christian Mair, Charles F. Meyer, (Eds.), *Language and Computers. Studies in Digital Linguistics* (The Hague: Brill, 2002), pp. 165–182.
- [22] David Mealand (1997) "Measuring Genre Differences in Mark with Correspondence Analysis," *Literary & Linguistic Computing*, vol. 12 no. 4 (1997), pp. 227–245.
- [23] Stephane Baele, Katharine Boyd, and Travis Coan, "The Matrix of Islamic States' Propaganda: Magazines"; in: Stephane Baele, Katharine Boyd, and Travis Coan, (Eds.), *ISIS Propaganda: A Full-Spectrum Extremist Message* (New York: Oxford University Press, 2019), pp. 84–126.
- [24] Kroonenberg and Greenacre, op. cit.
- [25] See Ziva Kunda, *Social Cognition. Making Sense of People* (Cambridge, MA: MIT Press, 1999), pp. 42–51.
- [26] For example Junseop Shim, Chisung Park, and Mark Wilding "Identifying Policy Frames through Semantic Network Analysis: An Examination of Nuclear Energy Policy across Six Countries," *Policy Sciences*, vol. 48 (2015), pp. 51–83; Leo Kim, "Denotation and Connotation in Public Representation: Semantic Network Analysis of Hwang Supporters' Internet Dialogues," *Public Understanding of Science*, vol. 22, no. 3 (2013), pp. 335–350.
- [27] For example Stephane Baele, Katharine Boyd, and Travis Coan, "Extremist Prose as Networks," *CREST Security Review*, no. 5 (2017), pp. 20–21; Baele, Brace, and Coan, op. cit.
- [28] Higuchi Koichi, *KH Coder 3 Reference Manual* (2017); URL: <https://github.com/ko-ichi-h/khcoder>. All input information and files are available upon request from the authors.
- [29] For example Conway, Scrivens, and Macnair, op. cit.
- [30] Freelon, Lynch, and Aday, op. cit.
- [31] McCauley and Moskalenko, op. cit.