

ISD

Powering new
generations
against extremism

ichbinhier 

HATE AT THE PUSH OF A BUTTON

Right-wing troll factories and
the ecosystem of coordinated
hate campaigns online

Philip Kreißel

Julia Ebner

Alexander Urban

Jakob Guhl



Organisations:

The **Institute for Strategic Dialogue** (ISD) is an independent "Think and Do Tank" that works with leaders in policy, business, civil society and academia to seek transnational responses to the geostrategic, social and security challenges of our time. ISD's goal is to fight extremism worldwide and bridge inter-community divides.

The Facebook action group **#ichbinhier** was launched by Hannes Ley in December 2016 following the Swedish model of #jagärhär. The group's aim is to work against the increasingly widespread hostility found in social media comment threads and to bring about a better culture of dialogue. The group members do this by arranging unified counter arguments on the Facebook pages of major German media outlets. The ichbinhier e.V. association was founded at the end of July 2017 and is dedicated to supporting the action group, initiating educational projects and to media work for raising awareness.

Acknowledgements

The authors would like to express special thanks to Prof. Dr. med. Martin Emmer of the Institute of Media and Communication Studies, Freie Universität Berlin, Prof. Dr. med. Nejdli (L3S Research Center), Dr. med. Gerrit Eichner (Justus Liebig University Giessen) and Alex Krasodonski-Jones (Center for the Analysis of Social Media) for their critical feedback on our methodology and conclusions. We also wish to thank the teams at ISD, the Online Civil Courage Initiative (OCCI) and #ichbinhier.

English translation by Dominica Campman
metatetxt Ltd. | metatetxt.co.uk.

© ISD, 2018
London | Washington DC | Amman | Beirut | Toronto

This material is offered free of charge for personal and non-commercial use, provided the source is acknowledged. For commercial or any other use, prior written permission must be obtained from ISD.

In no case may this material be altered, sold or rented. ISD does not generally take positions on policy issues. The views expressed in this publication are those of the authors and do not necessarily reflect the views of the organisation.

Designed by forster.co.uk. Typeset by Lookandfeelstudio.

Key insights

The increase in coordinated hate: Analysis of over 1.6 million extreme right social media posts (Twitter and public Facebook pages) from February 2017 – February 2018 shows that even though explicitly racist, anti-Muslim and anti-Semitic posts have decreased since the NetzDG came into force in October 2017, coordinated right-wing extremist online hate campaigns have been more than three times as common since December (about 300,000 posts/month) as in the previous ten months (about 90,000 posts/month).

Vocal minority: Analysis of 700 posts, 16,830 comments, and 1.2 million likes shows that 5,500 accounts (5% of all hate speech accounts in January) account for 50% of the hate-related likes in the comment threads. About 1% of the accounts attract 25% of the likes.

Coordinated right-wing extremist accounts dominate: The discourse in many comment threads on Facebook is not a reflection of the entire political spectrum, rather it is dominated by sympathisers of extremist and anti-constitutional organisations. In many cases, these are targeted attempts to pressurise and intimidate politicians, journalists and activists.

Influential campaigns: We identified three major hate campaigns initiated and coordinated by right-wing extremists in the early months of 2018: #Kikagate, #Kandelistueberall and #120dB. The pro-AfD (right-wing Alternative for Germany party) election campaign with the hashtags #AfD, #TraudichDeutschland (Trust in yourself, Germany) and #Merkelmussweg (Merkel must go) in August-September 2017 was the right-wing extremist troll accounts' longest and most influential campaign.

Coordination in extreme right chats: These campaigns are spread by Reconquista Germanica and related right-wing extremist online networks from such Identitarian circles on social media using coordinated times and hashtags, which leads to them sometimes spending weeks in the top trends and thus setting the agenda for political discourse online.

Mainstreaming through politics and the media: The hashtags are also picked up by AfD accounts and Russian media such as RT and Sputnik and thereby get into the mainstream media. RT and Sputnik were among the first media outlets to report on the campaigns. Even mainstream media took up the Identitarian themes, but not their campaigns.

Table of contents

Organisations	2
Acknowledgements	2
Key insights	2
Contents	3
Glossary	5
I. Introduction (Summary)	7
Levels of analysis	8
Background ISD annual review: hate speech in Germany	8
II. Vocal minority: #ichbinhier analysis of the comments on Facebook	9
1. A Facebook group committed to decency	9
1.1 #ichbinhier action framework	9
1.2 Further qualitative criteria	10
2. Initial insights	10
III. Collection and evaluation of Facebook data	11
1. Making hateful speech measurable	11
1.1 Creating an information base	11
1.2 Hateful speech	11
1.3 Identification of hate speech accounts	11
1.4 A small number of users provide likes for hateful speech	12
2. Who is behind the accounts?	13
2.1 Preliminary considerations	13
2.2 The role of right-wing organisations: AfD & the Identitarian Movement	13
2.3 Coordination	14
2.4 Examples	15

Table of contents

IV. Background to hate speech coordination	16
1.1 Spreading hate: Coordination by organised troll factories	16
1.2 Campaigns	19
1.3 The role of bots	22
1.4 The role of the German media	23
1.5 The role of Russian media	23
V. Limitations	24
1.1 Classification of hate speech accounts	24
1.2 E-sympathiser	24
1.3 Differences in the approach of #ichbinhier and Reconquista Germanica	24
Conclusions	25
Recommendations: coordinated moral courage against coordinated hate	26
Further reading	27
Footnotes	28

Glossary

#120db:

#120db is a campaign by activists of the Identitarian Movement who want to draw attention to the alleged threat to European women from Muslim immigrants.

Der große Austausch – The great replacement:

One of the central concepts of the Identitarian Movement claims that the native European population is going to be replaced by non-European immigrants.

Discord:

Discord is an online chat app that was originally developed for video game fans, but now also includes several right-wing extremist channels.

E-sympathisers:

In this document, e-sympathisers are classified as accounts, not necessarily real people, who use social media to promote content that has been posted by parties or organisations such as the Identitarian Movement, using positive interactions (likes), thus showing their agreement with the shared or posted content.

Hate campaigns:

Hate campaigns are coordinated online actions with a high concentration of "hate speech" aimed at suppressing counter-speech, influencing opinions, drawing the resources of editors and addressing commentators and authors of videos, articles or photos directly, offensively and in a degrading way. In the context of this publication, the term "shitstorm" is equivalent to the term "hate campaign".

Hate comments:

Hate comments are comments below posts on social media that fall under our subsequent definition of hate speech.

Hate speech:

Hate speech is comparatively new as a term and so far, not conclusively and authoritatively defined. The Council of Europe's Committee of Ministers calls hate speech "all forms of expressions that spread, incite, promote or justify racial hatred, xenophobia, anti-Semitism or other forms of hatred

based on intolerance, including intolerance in the form of aggressive nationalism and ethnocentrism, discrimination and hostility toward minorities and people with a migrant background."¹

Hateful speech:

Hateful speech is the term used to refer to posts and comments that in their entirety contain significantly more hate speech than regular sets of comments and posts.

Identitarian Movement:

The Identitarian Movement (IM) is an extreme right movement that considers "preserving our ethnocultural identity" and "defence of what defines us" against multiculturalism, globalisation and "Islamisation" as its major goals. It is being monitored by the Federal Office for the Protection of the Constitution because of overlaps in its personnel with nationwide right-wing extremism groups.

#Infokrieg – Info Wars:

#Infokrieg is a extreme right German-speaking Discord group, which among other things attempted to influence the 2017 federal election by means of coordinated online campaigns. The group's name is a direct translation of "Info Wars," the site of conspiracy theorists Alex Jones and Paul Joseph Watson.

Kandel and "kandelistueberall" (Kandel is everywhere)

At the end of December 2017, an Afghan refugee stabbed his 15-year-old ex-girlfriend in Kandel (in the state of Rhineland-Palatinate). The case sparked discussions on social media and was interpreted by right-wing groups in particular as an example of a failed immigration policy.

KiKAgate:

The youth television channel Kinderkanal (KiKA) portrayed the relationship between a Syrian refugee (whose exact age was unclear) and a 15-year-old German girl in an episode of the "Schau in meine Welt" (Look into my world) series at the end of November 2017. The broadcast led to a shitstorm in early January 2018, mainly through right-wing groups on social media.

Mass media or mainstream media:

Mass media indirectly reaches recipients, who are neither specified nor defined, through different channels (visual, auditory, audio-visual). So-called mainstream media can be described as "more or less extensive media consensus on certain issues, or also as: a number of topics and opinions that dominate in the media landscape in a certain period of time and thus forming a 'mainstream' or 'main direction'."²

NetzDG:

The Network Enforcement Act (NetzDG) is a German law intended to combat hate speech on social media. It requires platform operators to remove "obviously unlawful" content within 24 hours after being alerted to it.

Right-wing extremism:

According to right-wing extremism expert Cas Mudde (2000), extreme right-wing groups and individuals tend to have at least three of the following five characteristics: nationalism, racism, xenophobia, antagonism to democracy and a yearning for a strong state.

Right-wing populism:

Right-wing extremism expert Cas Mudde (2004) calls populism "an ideology that assumes that society is divided into two homogeneous, antagonistic groups, the 'pure people' and the 'corrupt elite', and that argues that politics should be an expression of the *volonté générale* or the wishes of the general population."³

Reconquista:

The Identitarian Movement considers that the term "Reconquista" (the reconquest of the Iberian Peninsula from the Moors) represents a historical analogy to the current situation in which the Identitarians also face a Muslim 'invasion'.

Reconquista Germanica:

"Reconquista Germanica" is an extreme right Discord group whose primary goal is to influence online discourse and thereby apply political pressure. The group's members are organised based on strict hierarchies, clearly distributed thematic and geographical areas of responsibility and with military vocabulary based on the Wehrmacht (unified armed forces of Nazi Germany).

Remigration:

The "remigration" of non-European ethnic groups is one of the Identitarian Movement's central demands, which it sees as the necessary answer to the "great replacement".

Troll or troll account:

A troll or troll account refers to a user who deliberately interferes with communication through provocation and thus wants to provoke an emotional reaction with the aim of stoking conflict and influencing opinions. Trolling, troll campaigns or troll groups can thus be understood as targeted, coordinated activities of such troll accounts.

I. Introduction (Summary)

An increasingly irritable mood has been observed in our society for some years now, which finds particular expression in the comment section on social media. Insults and threats, the exclusion of and hostility towards entire groups of people and even the blatant advocacy of violence are the agenda of the day on Facebook, Twitter and the like. It is difficult to have factual, constructive conversations in an environment where pejorative, aggressive and sweeping comments dominate.

The aim of this joint study by ISD and ichbinhier e.V. is to better understand the origins and spread of online hate speech in order to develop more effective methods to identify or contain it. This should improve the culture of discussion on social media comment columns so that the silent social media users can feel empowered to participate in debates on social or political issues again.

"[counter-speech] I have tried (to speak up) in the past, but stopped again because at some point it was just a drain on my time and also on my own peace of mind [...]"

Facebook comment on SZ-online

Therefore, in a first step (Chapter II), the self-defined scope of activities of the #ichbinhier group is explained in more detail, where hate speech was initially sought out based on qualitative criteria. There are certain difficulties though in classifying comments as hate speech. Rather than simply relying on legal categories, the phenomenon is approached primarily in terms of linguistics (Chapter II, point 1.1). However, even if an authoritative definition were used, it would be arguable on a case-by-case basis whether or not a given comment falls within that definition. It is currently not possible to find hate speech, which is based on such a linguistic definition, solely by automated means. In addition, online activists constantly succeed in designing intimidation and hate campaigns against individual players and organisations to circumvent the new NetzDG provisions. For this reason, the authors explicitly do not pursue the goal of conclusively identifying hate speech in a more unambiguous, narrower sense. The term "hateful speech" will therefore be introduced later.

Unsurprisingly, it seems that certain topics generate a particularly high level of comments. Chapter II, points 1.2 and 1.3 describe the basis for finding hate speech and hateful speech using data analysis. In the course

of this data analysis, it has been shown that many comment columns are dominated by accounts that have a right-wing to extreme right-wing background (Chapter II, point 3).

It should not be forgotten that apart from extreme right groups such as Reconquista Germanica and #Infokrieg, there are also other extremist activists who misuse social media for their political ends. However, in their analyses over the past few months, ISD and the OCCI – as well as #ichbinhier – have repeatedly ascertained that hate speech from the extreme right-wing spectrum is a particularly virulent problem. These initiatives have observed that there is currently no such comparable level of coordinated hate campaigns in the German-speaking world in either left-wing extremist or Islamist circles. The coordination of such right-wing extremist campaigns will be considered in Chapter IV.

Islamist extremists in particular are now less present on the major platforms and are heavily dependent on encrypted messaging applications such as Telegram due to strengthened measures taken by social media companies last year. This also means that Islamist groups, which often posted overtly extremist content until early in 2017, are now limited to non-political religious posts (Fielitz et al. 2018).

Names and accounts of individual users have been hidden to protect their privacy. It has been ensured that the collected data is published in an aggregated form which does not allow for any conclusions on individual conduct.

The following three levels of analysis, described in more detail below, were examined for the present investigation of the spread of hate speech on Facebook and the intensity in participation of extreme right users of hate speech.

Levels of analysis

Level 1: Annual trend analysis

First, an annual analysis of general developments relating to hate speech online was conducted using the Crimson Hexagon digital analytics tool. Trends in the volume of coordinated hate posts and shared media articles have been noted in particular. A total of 1.6 million posts were evaluated during the period February 2017 to February 2018.

Level 2: Hate comment analysis

Three datasets were evaluated using R statistical programming language (R-Studio) in conjunction with the freely available R Facebook plugin and Facebook's own API functions (API version 2.9). Dataset I refers to #ichbinhier campaigns. Dataset II contains 700 comment columns that are thought to contain hateful speech. Dataset III contains a cross-section of 1,000 comment threads.

Level 3: Hate-spreading methods

The data analysis was supplemented with 6 months of ethnographic research within the encrypted chats of German right-wing extremists on the Discord platform, where the hate campaigns were coordinated. This qualitative long-term observation enabled us to analyse the strategies and techniques for propagating hate and to explain the effects shown at level 2.

Background ISD annual review: hate speech in Germany

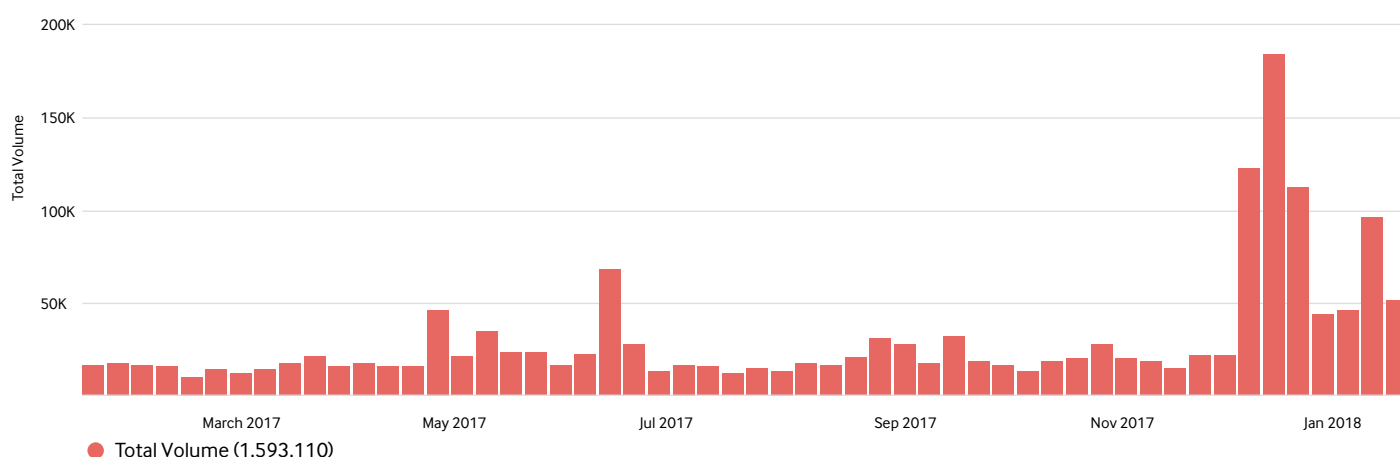
It was possible to investigate developments in the number of posts associated with right-wing extremist hashtags and keywords using the Crimson Hexagon digital analysis tool, which combines linguistic analysis capabilities with machine learning. In total, we analysed over 1.6 million posts on Twitter and public Facebook pages in the period from February 2017 to February 2018.

On the basis of our 6-month ethnographic research in right-wing extremist forums and chat groups, keywords and hashtags, which were defined and used in connection with coordinated hate campaigns in right-wing extremist Discord groups, were identified. The following recurring hashtags and keywords were used accordingly for the annual analysis:

#120dB, #Kandel, #kandeliseverywhere, refugee crime, #NetzDG, #Censorship, #120DEZIBEL, Merkelstan, remigration, close the borders, #thegreatexchange, half negro, mass immigration, fear of being overwhelmed by strangers, islamisation, criminal migrants, isolated case, Reconquista.

German original: #120dB, #Kandel, #kandelistueberall, Flüchtlingskriminalität, #NetzDG, #Zensur, #120DEZIBEL, Merkelstan, Remigration, Grenzendicht, #dergroßeaustausch, halbneger, Masseneinwanderung, Überfremdung, Islamisierung, Krimigranten, Einzelfall, Reconquista.

We trained the algorithm used to distinguish between positive and negative mentions of keywords and hashtags, and in which media they were mentioned. We noted a large increase of posts related to coordinated hate campaigns on social networks (Twitter and public Facebook pages). Between February 2017 and November 2017, there were an average of approximately 90,000 posts per month related to coordinated extreme right campaigns, however we identified an average of over 300,000 such posts per month between December 2017 and February 2018.



The most common hashtags used by Identitarians in January/February 2018 were:

#NetzDG	410,000+	<div></div>
#Kandel	190,000+	<div></div>
#AfD	140,000+	<div></div>
#Einzelfall	100,000+	<div></div>
#Zensur	100,000+	<div></div>
#Islamisierung	94,000+	<div></div>
#120db	64,000+	<div></div>
#Maas	49,000+	<div></div>
#Remigration	44,000+	<div></div>
#Merkel	40,000+	<div></div>



This observed increase in the total volume of posts related to the above hashtags and keywords forms the background for the in-depth analyses of such posts, which will be described in more detail in Chapter IV.

II. Vocal minority: #ichbinhier analysis of comments on Facebook

1. A Facebook group commits itself to decency

When Hannes Ley founded the #ichbinhier Facebook group in December 2016, he had the urgent need, in his own words, to speak about emotionally charged topics in a fact-based and nuanced way, while avoiding the generalisations that he often encountered when such topics were raised. A constructive exchange should once again take place on social media, free of verbally aggressive devaluation of dissenters or whole groups of the population. What had already broken the surface in the past could no longer be overlooked in the year of the Brexit referendum and Donald Trump's election as US President: the harsh, aggressive tone prevailing in the comment sections, and those who tried to contribute to a fact-based discussion being insulted and frightened off. Not only those who had previously opposed these (hate-)inciting comments on Facebook, but also those people who had withdrawn before that, having become alienated from the comment sections, readily accepted the opportunity to actively contribute to decency, social peace and democracy on social media. At the beginning of 2017, the number of members of the action group grew to several thousand within a few weeks. Since then (as of June 2018), more than 39,000 people have signed up to the #ichbinhier group.

1.1 #ichbinhier action framework

When it was set up, the moderator team determined that #ichbinhier would be active on the Facebook pages of high-reach media from across the political spectrum with more than 100,000 followers, but not on private pages, not in Facebook groups and not on the pages of political parties.

In exceptional cases, the scope of activities was extended to include the Facebook pages of public figures, NGOs, foundations or other initiatives, when it had to be assumed that they would be subject to organised and otherwise controlled campaigns (shitstorms). In the further course of this publication, coordinated campaigns (see Chapter IV.) are described as examples.

Essentially, the major media sites (e.g. Tagesschau, ZDF heute, BILD, SPIEGEL Online, Welt, FOCUS Online) were checked daily to see whether there were any sweeping, inciting and/or derogatory comments below the articles. If the moderators found such a comment section, a "campaign" was proclaimed within the Facebook group; each member could then take part free of any kind of guidelines in using their own comments to oppose hateful contents. The only rule: they had to maintain objectivity and respect for the other person.

In order to distinguish hate speech from partly justified criticism or black humour, the team of moderators first of all referred to the definition that the Duden (definitive German language dictionary) offers for the concept of agitation. Very similar definitions can also be found in other encyclopaedic works. According to this, agitation is "the totality of unobjective [unsachlich], hateful, slanderous, disparaging utterances or actions (...) that produce feelings of hatred, hostility, and emotion against someone or something."

The moderators also consulted the 2007 publications of the Committee of Ministers of the Council of Europe and Prof. Anatol Stefanowitsch of Freie Universität Berlin. On this basis, the moderators understand the concept of hate speech as the deliberate disparagement and humiliation of individuals, groups, communities or ethnic groups. Persons are marked and excluded as "different" on the basis of certain characteristics ascribed to them – ancestry, origin, skin colour, sex, sexual orientation, religion, physical limitations, but also party affiliation, ideology, profession, etc. Following on from these characteristics, misinformation or false statements are deliberately disseminated, stereotypes and prejudices are used and those affected are insulted, threatened, scorned and dehumanised because of their group affiliation: examples being terms such as "hordes of asylum seekers" or "cattle", used to describe refugees. In some extreme cases, this may include advocating or sending a rallying call for physical violence against those affected.

1.2 Further qualitative criteria

However, the moderators did not only consider the content and tone of the comments written according to the criteria outlined above. In addition, they followed the development of the comment sections under articles on the Facebook pages of the major media organisations. It was noticeable that the amount of comments correlated strongly with the topic of the article in question.

In particular, reports on migrant crime or terrorist attacks as well as articles relating to refugees or Muslims in Germany, but also (albeit much less frequently) articles with a socio-political topic were frequented significantly more by commentators. Several hundred comments among such contributions were made within an hour. The first comments on these topics generated two- or even three-digit like numbers after only a few minutes.

There appears to be a direct correlation – and this has been observed over the entire period since the group was founded up to January 2018 – between the issues of "immigration" and the number of comments or their like numbers. In terms of content, the authors of the comments we identified as problematic positioned themselves particularly strongly against immigration, Muslims and politicians.

It should be emphasised here that #ichbinhier becomes active when all these types of derogatory comments cluster. For example, the comment sections sometimes contain attacks on the integrity of AfD politicians, which #ichbinhier members then also counter. However, these comments rarely reach the extent or severity described above as they do in the other areas, and they are quickly countered even without #ichbinhier's intervention. #ichbinhier is therefore very rarely actively coordinating in this area.

2. Initial insights

With respect to the scope of activities outlined at the beginning, it was noticeable over the entire period up to January 2018 that, in the final analysis, #ichbinhier campaigns arose almost exclusively from media articles dealing with the above-mentioned topics. Although the moderators did not primarily focus on the topics in the selection, but rather towards the amount of comments and the proportion of hate speech in a comment thread. This allows the (initially qualitative) assumption that (actively) inflammatory comments on media articles of the above-mentioned scope of topics tend to originate from supporters of the right-wing populist or right-wing extremist spectrum. In any case, there was neither a high number of comments nor hate speech that could be discerned on the media sites, which had a large reach, that were observed by the team of moderators and which could have been attributed to the left spectrum.

This was the reason we attempted to automatically identify relevant media articles on the basis of topic patterns. The procedure chosen and the insights gained form the subject of the following sections.

III. Collection and evaluation of Facebook data

1. Making hate speech measurable

1.1 Creating an information base

The goal was to collect various data from media comment sections on Facebook and examine it in the context of hate speech. This meant answering the following questions:

How large is the group of people, or in this case accounts, that is/are spreading hate speech in the media article's comment columns? Can the accounts be classified politically? How do these accounts work together and is it done in a coordinated way?

Since the entirety of all hate posts is not clearly describable and accessible, a sample of three datasets has been prepared that is as representative as possible under the technical circumstances and constraints of this study. The investigations described here were carried out with R Studio (see Annex for technical details).

First, all top-level comments from the coordinated #ichbinhier campaigns in December 2017 were downloaded and analysed (dataset I).

#ichbinhier campaigns alone do not offer a general overview of the comment sections of media pages on Facebook in their totality. However, this overview is necessary to be able to observe the behaviour of the accounts in a long-term and comprehensive way. Therefore, starting from #ichbinhier campaigns in December, another 700 similar comment sections were analysed on 16 media outlets in January (dataset II). As it will be shown these are comment columns with an increased amount of hate speech.

In order to gain an overall view, to validate the methodology and to prove any coordination, a cross-sectional dataset was downloaded (dataset III), which contained the last 100 posts from 10 major media pages.

1.2 Hateful speech

As outlined above, the concept of hate speech is difficult to grasp. Since algorithmic or automated identification of actual hate speech was not possible, a simplified definition of hate speech, the term "hateful speech", was used here:

"Hateful speech" refers to a group of comments that has a significantly higher proportion of hate speech than general groups of comments on Facebook. The goal is therefore not to identify hate speech

(which may not have provided sufficient data due to the effects of the NetzDG and moderation activity on media pages), but to increase the probability of hate speech occurring in the data so it is possible to analyse the causal connections within the data. "Hateful speech" is a methodological construct. The conclusions of this work are not based on the rating of individual comments as hate speech and are therefore as objective as possible.

Firstly, this procedure ensures that a wide range of topics and forms of hate speech are included in the sample, minimising the risk of bias in certain topics. Secondly, there is a high likelihood of discovering all the current debates and issues that are dominant among the general public, as well as on Facebook, and which are characterised by an increased frequency of hate speech.

For #ichbinhier campaigns (dataset I), which indicate a significantly increased occurrence of hate speech, all comments which receive no likes from #ichbinhier members are considered hateful speech from a methodological point of view. As a reminder: "hateful speech" is not a normative label for a single comment. The basis here is the large and nonpartisan network of #ichbinhier. Regardless of the evaluation of individual actions, it can be assumed here that this data as a whole are more likely to contain hate speech than normal comments.

1.3 Identification of "hate speech accounts"

Based on and with the help of dataset I ("Campaigns from #ichbinhier in December 2017"), the larger dataset II was created in a multi-stage process:

Step 1: Since the purpose of this analysis is to analyse the like behaviour of users, accounts that interact positively with hateful speech comments, e.g. in the form of likes, were first identified on the basis of the above definition. From dataset I ("#ichbinhier campaigns December 2017"), the 1,000 most active non-ichbinhier accounts that supported hateful speech comments ("likely hate speech accounts") in at least four #ichbinhier campaigns during the December 2017 observation period were selected.

Step 2: In order to include media articles in the analysis that were not the subject of #ichbinhier campaigns as particularly incisive or prominent cases, media articles were selected from 16 media sites that aired intensive participation from these 1,000 accounts identified in December 2017. From this it can be concluded that these media articles are highly

likely to contain hateful speech comments. Even if individual accounts had been assigned incorrectly, the large number of simultaneously occurring accounts permits this conclusion.

Step 3: Comments with likes from #ichbinhier members were excluded from the analysis because it can be assumed that it is less likely that they contain hateful speech in their entirety. "Contamination" of the data by coordinated #ichbinhier campaigns is also excluded. To avoid miscategorising comments based on a low number of likes we did not consider comments with less than 10 likes either, which also kept to amount of data gathered manageable.

This resulted in a dataset of about 14,000 comments in total, each of which received more than 10 likes, written in the comment sections below 700 media articles, with around 600,000 likes in total ("dataset II").

The procedure for selecting the sample should ensure that the analytical base is as representative as possible. Firstly, the diversity and large number of members of the #ichbinhier network and, secondly, the following of comments/likes influenced by particularly active authors of hate speech are emphasised. As a result, the probability of relevant and important debates shaped by hate speech being overlooked can be considered low.

However, since the information base is not a systematic random sample and cannot be a random sample due to the technical limitations of Facebook, there is always some uncertainty about the findings, which is why minor differences between groups or types should be interpreted with caution. In addition, actual conditions may deviate from the findings presented here for unusual periods of time or isolated topics or forum areas.

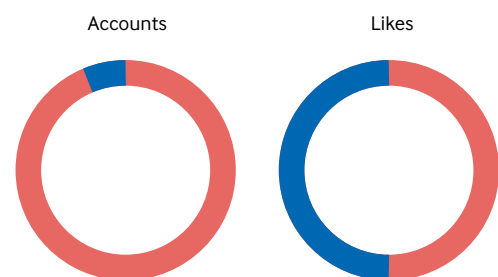
To verify that the extracted dataset actually contained more hate speech, two independent investigators manually compared blind samples with the definition of hate speech (see above). This confirmed that dataset II contained significantly ($p < 0001$) more hate speech than a random set of Facebook comments (for more details see Annex).

1.4 A small number of users provide likes for hateful speech

Dataset II allows the number of hate speech accounts in the comment columns on Facebook to be roughly quantified.

In addition to the 1,000 accounts ("probable hate speech accounts") used to identify those media articles among which a high percentage of hateful speech comments were written as defined above, nearly 5,500 more highly active accounts were identified in this dataset. These correspond to about 5% of the accounts involved in hateful speech with likes in dataset II. These accounts were responsible for about 50% of all likes for hateful speech comments in the 700 media articles examined. The most active 1% of users generated 25% of the likes for hateful speech comments. Comments that received likes from members of the #ichbinhier group were excluded from this estimate, as the action group is coordinated and therefore this data could falsify the outcome. So, the focus here is entirely on the potential hateful speech comments.

The 5% identified represent just 0.02% of all monthly Facebook users (for the month of September 2017: approximately 31 million Germans). A small group of highly active accounts dominate a large number of comment columns and interact positively with hateful speech.



2. Who is behind the accounts?

2.1 Preliminary considerations

Having demonstrated in the previous chapter that particularly active accounts dominate most of the hateful speech comment threads, the following questions then arise:

- a) Do hateful speech accounts constitute a representative cross-section of the population,
- or
- b) Do these accounts belong to a specific political camp?

First, a general definition of a "sympathiser of parties and organisations" must be found. In this context, it may be taken as a basis that the account has supported content posted by parties or organisations through positive interactions (likes) and thus indicated that they concur with the shared or posted content. The term "e-sympathiser" is used in the following to designate these accounts. The term "account" is important here. A real person can have several accounts, indeed in right-wing extremist forums it is explicitly encouraged to create several so-called "sockpuppets".

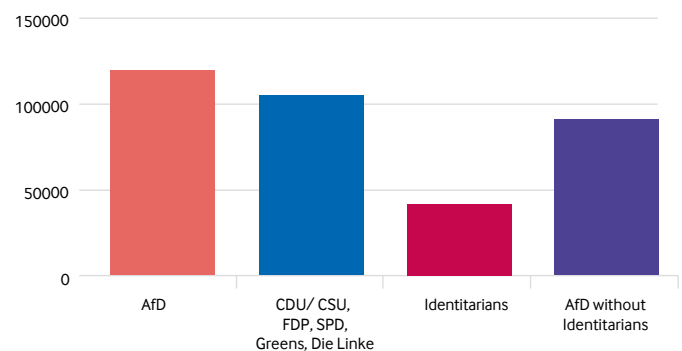
In order to map the widest possible spectrum, the likes for the most recent 200 posts were downloaded from the primary Facebook pages of the respective organisations via the Facebook interface. To be designated as an "e-sympathiser" of an organisation, the account must have liked at least one of these 200 posts. The focus of analysis was on all parties represented in the Bundestag. In each case, only the party's main page was evaluated. Pages of the regional parties (except the CSU party) or their personnel were not included to ensure maximum comparability. In addition, the activity of e-sympathisers of the Identitarian Movement (IM Germany) and Björn Höckes was investigated, as both play a major role in the coded Discord forums "Reconquista Germanica" and "Infokrieg".

The e-sympathisers of parties or organisations identified were compared with the 700 comment sections of the media articles that had already been analysed. As a result, one third of the likes for hateful speech in these 700 comment columns could be attributed to e-sympathisers of a party represented in the Bundestag.

2.2 The role of right-wing organisations: AfD and Identitarian Movement

It is especially striking that in these 700 comment columns among the media articles that were examined and contain hateful speech comments, there are about as many likes from e-sympathisers of the AfD as from e-sympathisers of all other parties put together.

Likes of e-sympathisers of the parties

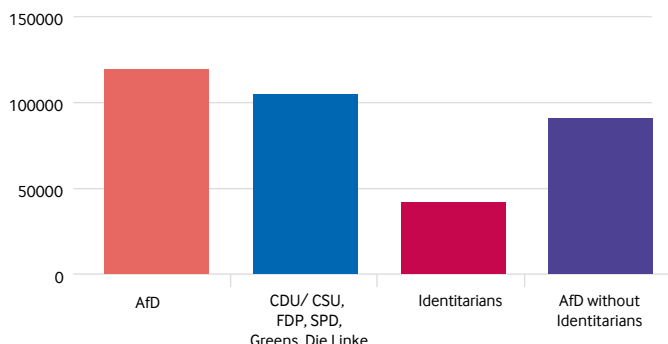


In addition, accounts that can be classified as e-sympathisers of the Identitarian Movement (IM), which is being monitored by the Federal Office for the Protection of the Constitution and classified as right-wing extremist are measurably more active than the accounts of all the other groups examined. This group is however significantly smaller in total than, for example, the e-sympathisers of the AfD. Thus, e-sympathisers of the IM distributed likes equivalent to a third of the likes of AfD e-sympathisers. Accounts of the e-sympathisers of the Identitarian Movement were on average clearly more active than AfD e-sympathisers and e-sympathisers of other parties represented in the Bundestag.

As an arithmetic mean, the IM was active among the 700 articles in dataset II with 40 accounts identified as e-sympathisers through positive interaction, while the AfD had 107 e-sympathiser accounts. If AfD e-sympathisers who are at the same time e-sympathisers of the IM (intersection AfD-IM) are discounted, this results in an average of 83 "pure" AfD e-sympathisers under one of the 700 articles. This again shows the extremely high activity of IM accounts compared to AfD.

In the media reports that led to #ichbinhier campaigns, the IM's involvement was even more intense:

Arithm. mean of accounts during #ichbinhier actions in Jan. 2018



E-sympathisers of the Identitarian Movement participated among such media articles with an average of 62 accounts during the observation period. There were at least 400 active accounts sympathetic to the Identitarian Movement during the campaign targeting a youth TV-channel ("KiKAgate") at the beginning of January 2018.

The figures given here are to be seen more as the lower end of the Identitarian Movement's engagement. It is probable that not all e-sympathisers could be identified as such, for example, if they were active with secondary accounts or just had not interacted with the organisation's main page in the observed period. It is noteworthy that the Identitarian Movement related accounts did not only appear occasionally. Rather, it became clear that there is a constantly intensive and coordinated appearance of accounts that had positively interacted with the Facebook page of the Identitarian Movement.

In addition, there was a strong correlation between the appearance of e-sympathisers of the Identitarian Movement and those of Björn Höcke, who were among the highly active profiles. No other politician is revered by Reconquista Germanica and other Discord servers in such a way (e.g. in militaristic presentation or in the form of memes).

Overall, these numbers are consistent with the composition of the extreme right "Reconquista Germanica" Discord server.

Figure 1: Recruiters from Reconquista Germanica consciously give AfD and IM members easy access



Ser Arthur Dayne heute um 22:01 Uhr

Eine grob umrissene Bewerbung wäre hilfreich und irgendwas, womit man dich verifizieren kann: Social Media Aktivitäten, Youtube, etwaige Mitgliedschaft bei IB oder AfD (auch Fördermitglied), Bilder von "gesinnungsfreundlichem Material"...



2.3 Coordination

The observation of highly active accounts of e-sympathisers of the Identitarian Movement is consistent with the "Handbuch für Medienguerillas" (Media Guerrillas' Manual) which Martin Sellner, co-director of the Identitarian Movement Austria, linked to his website as a handout and guide. It reads:

"Follow and like the accounts of all parties, especially the Greens, well-known feminists, government minions like Till Schweiger or Böhmermann and all propaganda government press, such as ARD, ZDF, Spiegel and the rest of the Fake News mixed clan. [...] All of the above is more effective if it is done not just by one but by a group of people. Make plans with your friends, and when the showdown comes, call them together to fight the lies and the poison together. If you do not have friends, create several accounts and run them in parallel. Build up an army of sockpuppets. Mankind is a herd animal. It is more willing to follow a group of people than one person."

One question that arises from this is the degree of coordination of accounts that appear in comment columns with hateful speech comments.

Against this background, the last 100 Facebook media articles from SPIEGEL ONLINE, VIP News, WELT, tagesschau, ZDF heute, Kronen Zeitung, BILD, HuffPost Germany, ZEIT ONLINE and the Frankfurter Allgemeine Zeitung were collected at the end of the period under review. All comments with more than 10 likes were then analysed. This resulted in a dataset (dataset III) of 6,000 comments with 300,000 likes, which come from 1,000 comment threads.

Each account's activity was calculated. The measure of activity was the number of comment sections in which this account had liked posts. Then it was determined how active the accounts are in these comment threads that like this comment. It turns out that some comments from much more active accounts were criticised more than others.

In reality, high-activity accounts seem to agree much more often than one would expect given normal probability. "Pure chance" can therefore be excluded with a high level of confidence (for more details see Annex).

Further analysis of the e-sympathisers in this dataset

shows that many likes of highly active accounts came mainly from sympathisers of the AfD and in particular the extreme right Identitarian Movement. IM and AfD sympathisers are significantly more active in dataset II than sympathisers of other parties and usually appear together.

The comments with the most active accounts are mostly pro-Trump and firmly against mainstream media, against the Greens, against Merkel, the EU and against refugees (see example Sample 2). These highly active accounts, a significant amount of which can be classified as e-sympathisers of right-wing extremist organisations, seem to spread certain narratives and content in the comment columns in such a way that they appear as "majority opinions" due to the number of their mentions and positive interactions.

Their consistency with the recommendations of the "Media Guerrillas Manual" suggests that that this effect is intentional. Extreme right-wing actors try to create the impression of representing the majority opinion through highly active and coordinated accounts.

2.4 Examples

Sample 1 of comments from **dataset II** (hateful speech) which were unanimously identified as smearing / inciting hatred by the examiners:

- "Right then, have fun in Algeria. Looking after goats, having 10 children, your morchid gets, as is tradition, an Algerian woman."
- "How long is this supposed to go on? This culture simply does not fit in here. And the whole thing probably happened because his pride was hurt. My condolences to the family. Can only guess how strong the hatred and the anger and at the same time the sadness is."
- "This kind of criminality by a senile SPD party member is criminal. It is also criminal to spread it about."
- "And where is the verdict for the politicians who have illegal border opening on their conscience. Should our children have to get by for 50 years or more before the carcass of these pigs gets convicted. Of course, if he's guilty, that's right but as one of the previous speakers said, why do you take so long?"
- "We feel so ripped off here in Germany. These eternal attempts to relativize separating Islamic immigration from the dramatic increase in crime are an insult to the intelligence of the people. Islam is spreading and Germany has changed as much in just 2 years as it did in 60 years."
- "most are illegal and not refugees so the thinking is wrong from the start and you cannot give underdeveloped people higher goals, these people come from a dark-ages society, how is that supposed to work? It is imperative to return them to their homeland and, above all, to cancel the development aid that leads to so many disputes in the affected countries"

As already mentioned, there are significantly more comments in dataset II that include hate speech or

smears.

Sample 2 of comments from **dataset III** (cross-section) with at least five times more active accounts than the average behind the likes (#ichbinhier deliberately hidden):

- "We should simply deport all criminal refugees and migrants ... help should be focussed on really threatened refugees".
- "Just ridiculous this number! Germany is and remains a banana republic! For the love of Germany"
- "And rights" is clear ... Just get beaten, then you're not a right-winger."
- Abolish GEZ fees (ed. note: for state TV). I am not watching or listening to this anymore. Very often very biased research, sentimentalism, completely excessive salaries for editors (Kleber, etc.) etc."
- "I also do not trust the public broadcasters on the refugee questions any more, they have lied to us and treated us like idiots too often"
- "New faces [politicians], same old story!!!"
- "Would be great! Then we all work for €6/hr again and our politicians and bankers and business leaders will be happy again!"
- "He [Macron] is losing more and more approval in France, at the moment about 10%. He probably should not plan that much"

This sample was taken from dataset III, with no other criterion than high activity (i.e. number of comment sections in which an account with likes was active) of the accounts behind the likes. This is not hate speech in the opinion of the examiners, but tend to conform with comments from users belonging to the right-wing populist spectrum.

Sample 3: For comparison, some randomly selected comments with users of very low activity from **dataset III**:

- "You can wash/peel vegetables and fruits! Just in case, I would only get sausage/cheese from the counter, served by the market's employees (1 job created + plastic avoided). Where there's a will there's a way!"
- "Not on the ascent, but on the descent... and not from 7,400 meters high... so many errors in one message. Embarrassing... It's a mystery why Urubko and Bielecki are not even mentioned..."
- "I'm not having any tests. My baby can be born just as it is and it will be loved. Done!"
- "It does not mean that every person who uploads edited images (or unedited ones) onto Facebook or Instagram can call themselves a photographer, though I often see great pictures by amateurs and very boring or even bad ones from the 'professionals'."
- Love is blind! Thank God there aren't the barriers that existed in my day. It was difficult to marry a partner from a different denomination then, let alone a different nationality! I'm glad that a lot has changed here!"

IV. Background of hate-speech coordination

1.1 Spreading hate: coordination by organised troll factories

Based on our 6-month ethnographic research in right-wing extremist forums and chat groups in Germany, we analysed the coordination of hate, intimidation and manipulation campaigns on the Internet.

The US elections showed for the first time the extent of the influence that coordinated troll factory campaigns can have on our political and social structures. Donald Trump's election victory, which the American alt-right celebrated as a result of its online trolling campaigns, motivated right-wing extremists around the world to try to influence the political online discourse in their country in favour of right-wing populist parties. These campaigns were typically planned in closed chat groups to spread hatred against minorities, the political establishment and the mainstream media. In 2017 alone, numerous coordinated attempts to influence the elections in France, the Netherlands, Germany and Austria by right-wing extremist groups online were observed.⁵

In the summer of 2017, German right-wing extremist groups joined forces for the first time with the aim of influencing the outcome of the Bundestag elections in favour of the AfD. Discussions were initially held on 4Chan and 8Chan, but quickly shifted to encrypted messaging applications such as Telegram and Discord. The #Infokrieg and Reconquista Germanica channels played a key role in the Discord encrypted gaming application. The Reconquista Germanica group already had almost 500 members the day after its foundation, about 5,000 on the day before the Bundestag elections and more than 7,000 in February 2018 before it was closed down for the first time by the operator. The group, which now calls itself "Revivica Germanica", has had well over 10,000 members since the end of April.

The authors have noted with great surprise that the Federal Government, in its response to a brief parliamentary enquiry by the DIE LINKE parliamentary group, has declared that it has not yet classified "Reconquista Germanica" as right-wing extremist. As can be seen from the screenshots in this study, nationalistic, racist, xenophobic and in some cases even National Socialist contents and symbols are so dominant that it is unfathomable to us how the Federal Government comes to this conclusion. The group name already contains the anti-Muslim message that Germany needs be forcibly "reconquered" from Islamic invaders, as the Iberian Peninsula once was.

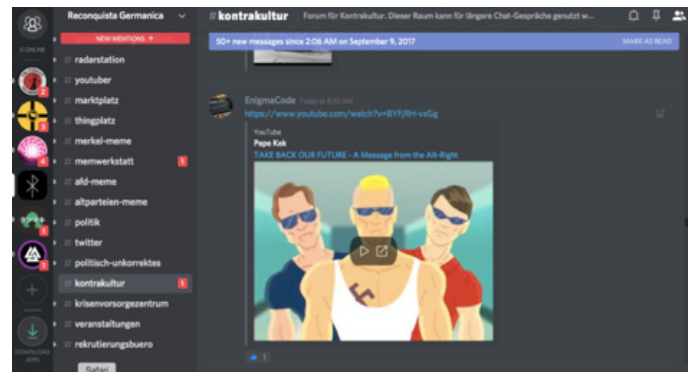


Figure 2: National Socialist symbolism in the Reconquista Germanica Discord channel

The organisational structures are strictly hierarchical with different strategic and regional levels. The self-proclaimed Commander-in-Chief Nikolai Alexander issues orders using military language with references to the Wehrmacht (see below). In addition to the Discord channels, the Twitter alternative Gab, which is mainly popular among right-wing extremists, and the Russian alternative to Facebook, VK.com, are used to coordinate and mobilise for trolling campaigns.

“The authors were very surprised to note that the German government has not yet classed "Reconquista Germanica" as right-wing extremist.”

terms	Use with Reconquista Germanica	Meaning
Case Blue	Rebuilding the server after it has been deleted	First part of the Wehrmacht summer conservation in 1942 during the German-Soviet War
Paladin	Second highest rank in RG	Nickname of Hermann Göring, "second man behind Hitler"
Army Group South	Regional Group in RG	Army Group in the Wehrmacht for the Polish campaign 1939
Probationary Battalion	Lowest rank of RG	Wehrmacht forces in World War II made up of criminals who were sent to the front for "probation"



Figure 3: Right-wing extremist troll groups: #Infokrieg (above) and Reconquista Germanica (below)

The so-called "Crisis Prevention Centre" on "Reconquista Germanica" shared violence and weapon-glorifying content as well as tips on how the users themselves can build their own stun gun.

The right-wing extremist trolls use different techniques and tactics to manipulate the social media algorithms, the coordinated dissemination of hatred and the targeted influencing of online discourse.

In so-called "sniper" missions, right-wing extremist activists try to provoke known personalities by trolling them into a verbal overreaction in order to

publicly expose them. On the other hand, "clear and hold" missions aim to infiltrate opposing hashtags (#refugeeswelcome) with their own content in order to force users who use such hashtags positively offline. Under "search and destroy" operations, a certain post, previously selected on Discord, is simply massively "spammed" with memes and comments.

Times and hashtags are synchronised by members of the extreme right-wing troll groups to attack "old-party" politicians, left-wing activists or mainstream media, "spam" the commentary area of videos with provocative messages or, in particular, distribute many "dislikes" in targeted "raids" on Facebook, Twitter and YouTube. The main themes are mass immigration, supposed "Islamisation", identity, freedom and tradition.

These operations are often extremely successful. For example, 15 days before the Bundestag elections, the extreme right-wing activists managed to place seven of their hashtags (including #TraudichDeutschland (go for it, Germany), #nichtmeinekanzlerin (not my leader), #merkelmussweg (Merkel must go), #reconquista) in the top 20 hashtags on Twitter in Germany.

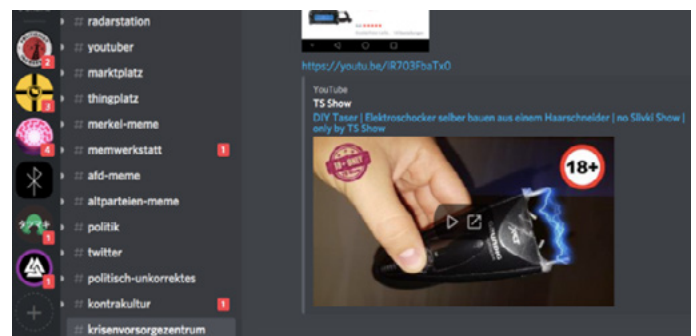
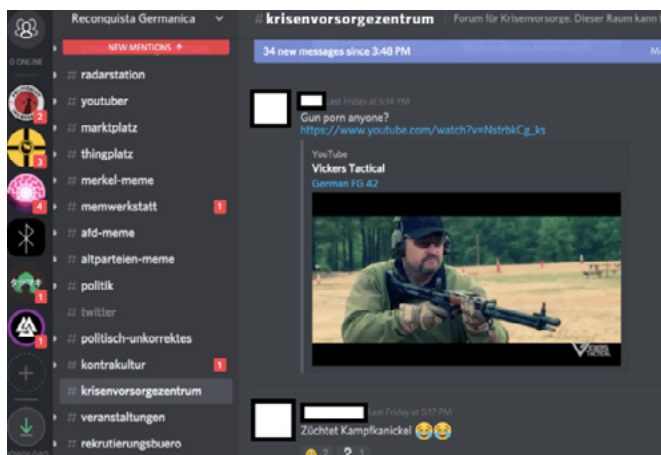


Figure 4: Instructions for building your own stun gun



Weapon-glorifying content in the "Crisis Prevention Centre" of "Reconquista Germanica"



Figure 5: The use of the #traudichdeutschland hashtag pushed by "Reconquista Germanica" increases significantly in the two weeks before the election.

#Infokrieg's administrators instruct their followers to accumulate several social media profiles (ideally outside their own filter bubble) and to post them simultaneously via several accounts with the "Ghostbuster" program in order to appear unsuspicious. In addition, they should tweet on

non-political issues (cars, football) to increase the number of their own followers. In addition, some of the extreme right trolls also get spoof accounts in order to troll their political opponents and create confusion. This also means that the investigation in Chapter III 2.2. likely underestimated the number of these accounts.

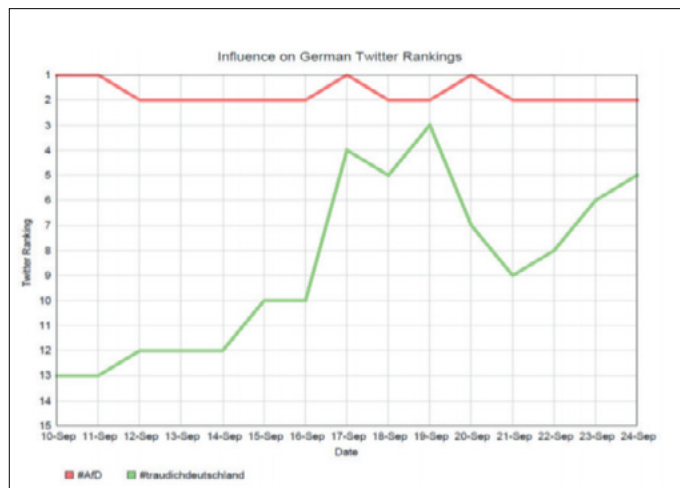


Figure 6: The use of the #traudichdeutschland hashtag pushed by "Reconquista Germanica" increases significantly in the two weeks before the election.

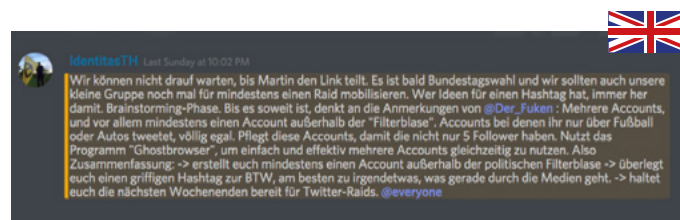


Figure 7: Instructions for using multiple accounts through the "Ghostbrowser" program and the importance of accounts outside the extreme right filter bubble (in which it is then tweeted via football or car themes to win followers, for example).

Provocative political memes are also developed and collected in the "Reconquista Germanica" and "#Infokrieg" Discord groups in meme databases, which are then disseminated on social media platforms such as Facebook and Twitter. In order to obtain politically sensitive material, "Reconquista Germanica" also distributes instructions on how to hide IP addresses, how to create automated social media accounts and how to prevent them from being taken down by the platform operators.



Figure 8: Chart of the "military" ranking within the "Reconquista Germanica" Discord channel.

1.2 Campaigns

Despite the formal AfD-resolution, which is intended to prevent cooperation between AfD members and extremist groups such as the Identitarian Movement, which is monitored by the Federal Office for the Protection of the Constitution, there are repeated reports that point to cooperation. Again and again, we came across AfD accounts that use the hashtags of these campaigns in their tweets in our analysis of campaigns organised by the extreme right troll factories. Due to their high number of followers on social media, these AfD-accounts function as an important vehicle for the right-wing Identitarian Movement to make their content accessible to a wider audience.

These include quite prominent members of the party, including several members of the Bundestag, who take up the contents of right-wing extremist Identitarians and spread them on social media. In addition to the Bundestag members Udo Hemmelgarn and Martin Reichardt, who are assigned to the right-wing and nationalist camp of the AfD party, the Leader Alice Weidel also used the #KiKAGate hashtag and asked whether KiKA was preparing "our girls for marriages with Islamic immigrants".

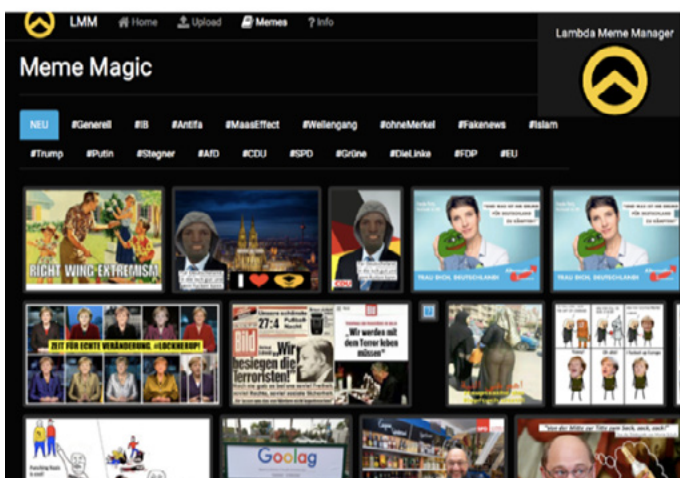


Figure 9: Online meme database where self-developed memes are shared and stored in the group to be used in appropriate campaigns



While it is sometimes possible that Identitarian Movement campaigns may be inadvertently amplified due to overlapping themes, in other cases deliberate intent appears to be the much more probable. The "Junge Alternative Hessen", for example, which is repeatedly accused of being close to the Identitarian Movement, the AfD-politician Martin Reichardt and the "Alternative NRW" all used the hashtag #120db, which clearly originated with the Identitarian Movement.⁹



Due to their high number of followers on social media, these AfD-accounts function as an important vehicle for the right-wing extremist Identitarian Movement to make its contents accessible to a wider audience.

In order to analyse a single campaign and to enable an even more detailed and direct comparison of the different activities and orientations, comment columns from dataset III on Holocaust Memorial Day were analysed.

This showed that the highly active accounts wanted to focus on anti-Semitism among refugees rather than commemorating the Holocaust. The same accounts also support an "end to the cult of guilt". In addition, statements by Holocaust survivor Hanni Levy, who had called for respectful treatment of refugees, are criticised in a targeted and coordinated manner.

The same could be observed in the "Reconquista Germanica" encrypted Discord forum.

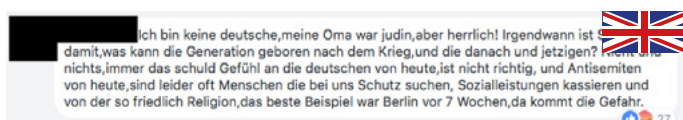


Figure 10: The likes for this comment came from accounts that were 60% more active than the average. At least 15% come from e-sympathisers of the IM. No statement about the authenticity of the commentator can be made, it is only about the likes.



Figure 11: This comment is on the WELT newspaper's Facebook page under a Sharepic with the inscription "We must never forget," which shows the trains standing in front of Auschwitz. It was liked by several e-sympathisers of the IM.

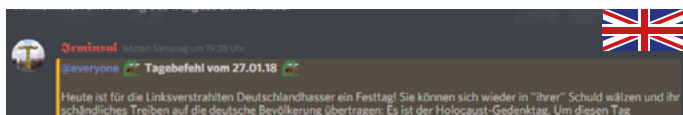


Figure 12: Screenshot from the "Reconquista Germanica" encrypted Discord forum



Figure 13: This ZDF post about the Holocaust survivor Hanni Levi features 104 e-sympathisers of the right-wing extremist Identitarian Movement in its "Like Activity" comment column. There were also 66 of the most active accounts that had participated in the #KiKagate shitstorm. All accounts interacted positively with comments criticising Hanni Levi.



Concerted right-wing extremist actions must be identified early so that civil society can take countermeasures.

In contrast to the highly active accounts, which want to transfer the common narratives described above and disseminated by "Reconquista Germanica" into the mainstream, there is much less interaction with comments that honour Holocaust Memorial Day as such.

It is noticeable, particularly in the analysis of the Holocaust Memorial Day, that individual accounts are not only significantly more active than others, but also that many very active accounts appear simultaneously, which in many cases also belong to the right-wing extremist Identitarian Movement. This leads to the conclusion that in many cases there was a coordinated appearance of these very active accounts.

In order to confirm that this observation does not describe an individual case but a joint strategic approach of the highly active accounts, the shitstorm against the ARD and ZDF children's channel ("KiKagate") was also examined in detail. Accounts of the Identitarian Movement were particularly active during this shitstorm, and a call to action also came from the right-wing extremist forum "Reconquista Germanica". The 1,000 most active accounts who were actively involved in hateful speech likes in the shitstorm against the children's channel were found in Holocaust Memorial Day comments with thousands of interactions.

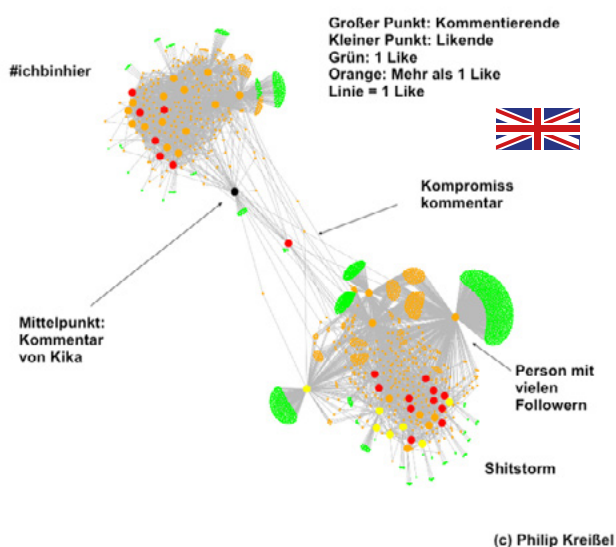


Figure 14: Visualisation of the KiK Shitstorm – the accounts involved in the Shitstorm raise each other's rating using likes.

1.3 The role of bots

On Twitter, there appears to have been an amplification of the planned hate campaigns by Russian-speaking bots, even though it seems that this had not been coordinated with the right-wing extremist online trolls. In the encrypted #Infokrieg group, for example, several users were taken aback that certain tweets had gone viral, and were surprised to find that many of the retweets were emanating from Russian-speaking bots.



Analysing bots is generally more difficult on Facebook: several authors have proposed criteria for identifying automated behaviour. These include: very high activity, all-day activity, etc. However, these criteria also apply to the coordinated behaviour of purely human-controlled coordinated actions. Therefore, unambiguous identification and differentiation of automated behaviour using the available data is not technically possible. In principle, however, the findings presented here represent a "proof of concept" because it can be shown that only a few accounts are needed to change the outlook of comment sections. It would be easy for non-state and state actors to

create fake accounts to fully control public debates in the comment sections on specific topics. However, it is not possible to prove the presence of state campaigns based on the available data.

1.4 The role of the German media

In the 1.6 million right-wing extremist posts analysed last year, links to the following media webpages were most frequently shared:

Most shared links to media webpages	Most discussed media articles
Philosophia-Perennis.com Refcrime.info Welt.de Tichyseinblick.de JungeFreiheit.de Bild.de Focus.de	1. Bild (newspaper): "This is Mia's killer posing in a tuxedo", 28 December 2018 2. Junge Freiheit (newspaper) "Censorship destroys Twitter", 31 May 2017 3. Welt (newspaper), "Previously convicted sex offender said to have raped 13-year-olds", 2 February 2018

Links to the following media websites were most frequently shared in connection with the #KiKagate, #Kandelistueberall and #120dB hate campaign hash tags:

Most shared links to media webpages	Most discussed media articles
Philosophia-Perennis Welt.de Bild.de Compact-online.de Focus.de pi-news.net Krone.at	1. Welt, "Whirlwind around refugee documentary in the KiKA children's channel", 9 January 2018 2. Bild, "Youth stabbed because of a girl", 18 January 2018 3. Bild (newspaper): "This is Mia's killer posing in a tuxedo", 28 December 2018

In contrast to the AfD accounts and the Russian media such as RT and Sputnik, German mainstream media such as BILD took up the themes of #Kandel and #KiKA, but did not refer to the Identitarian Movement campaigns.

1.5 The role of Russian media

Apart from AfD accounts, Russian media such as RT and Sputnik also play an important role as multipliers for the Identitarian Movement's hate campaigns. For example, both RT ("120 Dezibel – activists start campaign against "imported" violence against women in Europe") and Sputnik ("Women against 'Merkel's guests' – shift to the right or self-defence?") referenced the campaign of the Identitarian Movement #120db in their reports at the beginning of February.



V. Limitations

1.1 Classification of hate speech accounts

The classification of the 1,000 hate speech accounts (see Chapter 2.3) was based on their particularly frequent positive interaction with hateful speech. Even if individual interactions were wrongly classified as interactions with hate speech, the probability of an incorrect classification as a hate speech account is relatively low, because each of these accounts interacted positively with suspected hateful speech in at least four comment threads.

It should also be mentioned that articles with hateful speech might not have been recognised without the participation of the 1,000 hate speech accounts. This would have resulted in a selection bias. However, in the secondary selection period (January 2018), there was high participation from the identified hateful speech accounts in all media articles in which #ichbinhier became active, while without these 1,000 hateful speech accounts there was no significant occurrence of hateful speech.

Conversely, however, hundreds of comment sections with hateful speech and the participation of the hateful speech accounts were recognised in which #ichbinhier was not active, for example, because there was already counter-dialogue or there was too little interaction overall. This restriction was therefore not an obstacle to the identification of hateful speech, and at most led to underestimating the true extent of like numbers with a few accounts and the participation of extreme right accounts.

1.2 E-sympathiser

The definition of e-sympathisers may be too broad because a like does not allow a clear assignment of an account to a party. Nevertheless, the immense difference between the representation of e-sympathisers of the AfD vis-à-vis all other parties is so significant that this cannot be solely due to the methodology. The Identitarian Movement is more likely to be underestimated, as it often works with fake accounts that cannot be clearly assigned to the IM.

1.3 Differences in the approach of #ichbinhier and Reconquista Germanica

The #ichbinhier Facebook group also takes coordinated action in the comment sections to create a counter-dialogue when hateful speech comments have been identified. It is therefore a reaction; #ichbinhier does not become active to change opinion in the comment columns 'a priori', as for example, Reconquista Germanica does.

In addition, #ichbinhier **make known its coordination with a hashtag** for the sake of transparency. This could be interpreted as a disadvantage to #ichbinhier, but #ichbinhier considers it morally imperative to disclose its coordinated approach. Campaigns from extreme right groups, however, deliberately try to disguise their character, origins and actions.

Interestingly, some hateful speech comments sometimes get likes from more active accounts on average than the co-ordinated #ichbinhier comments with many positive interactions. There is no other plausible explanation for this than an extremely high degree of coordination.

One conclusion from this data is that the activities of #ichbinhier are not directed against factual conservative opinions, but against coordinated extreme right actions on media outlets' Facebook pages. Coordinated right-wing extremist actions are intended to suggest a prevalence of opinions that does not exist in reality.



Coordinated right-wing extremist actions are intended to suggest a prevalence of opinions that does not exist in reality.



Conclusions

Hate speech among media articles on the large German-language news sites on Facebook is produced, pushed and distributed by a small group of accounts, measured by the number of users. The distribution is often coordinated in terms of content and timing. This coordination aims to place political content in the centre of society and to create the impression that the communicated attitudes towards this content correspond with the average user's attitude. This achieves a distorted representation of the opinion, which in turn affects the actual opinions of the user.

Joint campaigns are coordinated on "Discord" and in closed groups on Facebook, for example. The closed groups have a content and organisational proximity to groups and parties that are located at the extreme right boundary of the political spectrum. The individual users sympathise with the content of these organisations. In many cases, the actions are so well organised that they can promote the conduct of hate campaigns and "shitstorms".

The analyses clearly demonstrate a strategic calculation by extreme right organisations to shape public opinion and sentiments on social media. In particular, activists from the right-wing extremist spectrum are showing a strong desire to transfer extremist content to the centre of society via social networks such as Facebook and Twitter and to suggest that they represent the majority opinion via the sheer mass of comments and likes. Social media provides an ideal environment for these organisations to use covert actions to carry their marginal ideology into the centre of public discourse. Representatives of the rest of the political spectrum, pro-constitutional citizens and minorities are insulted, worn down and driven out of the comment sections. Politicians and the media sometimes pick up content from campaigns of extremist organisations and thus make them socially acceptable in public discourse.

The supposed quantitative superiority of extremist users results in two effects:

Misinformation, extremist opinions, but also degrading or insulting language can, in connection with very high and fast growing like numbers, be perceived as the socially predominant opinion.

The present analysis shows that 5% of hateful speech accounts generate 50% of the likes for degrading comments. Thus, in the online world, it is not only majority opinions that are suggested, but also a social consensus to the effect that people of other origins, other faiths or ethnicities can be insulted and devalued without having to expect sanctions. This manipulates the formation of opinion in the population. It is not yet possible to predict what damage the campaigns described above will cause in the long term, for example with regard to social ties or political culture.

Another consequence of the coordinated actions is the impairment or even suppression of the pluralistic discourse. Multiple accounts specifically combat counter-dialogue and contribute in supposedly numerical superiority to suppressing expressions of opinion and commitment to objective discourse. As a result, elementary principles of fundamental law, such as the protection of minorities and the freedom of religion and opinion, are coming under increasing pressure – until now largely unnoticed.



The present analysis shows that 5% of hateful-speech accounts generate 50% of the likes for hateful comments.



Recommendations: coordinated moral courage against coordinated hate

This publication addresses those in the media, institutions and political parties as well as citizens who cherish values such as democracy and pluralism. It is intended to sensitise people to the functioning and mode of action of coordinated hatred and the resulting consequences.

In order to counteract the resulting destabilisation and to protect our constitution and democratic values, countermeasures by society as a whole are necessary. From the authors' point of view, the following recommendations are derived from the above findings:

The increased occurrence of the right-wing extremist Identitarian Movement, in which the Federal Office for the Protection of the Constitution sees "evidence of efforts against the free constitutional order", shows that an increased identification, observation and analysis of such digitally organised illiberal groupings online is absolutely necessary.

Both the media and civic organisations have an obligation to inform the public thoroughly and comprehensively about these targeted attempts to influence public opinion and to extensively inform and sensitise the population. Similar analyses and initiatives are also needed in other countries in Europe. For example, there are other counter-speech initiatives in other European countries based on #ichbinhier that need support. The problem is not limited to Germany, but is a general problem in digital social networks.

Overall, the evidence requires better collaboration between social media operators and encrypted platforms and anti-hate initiatives. Concerted right-wing extremist actions must be identified early so that civil society can take countermeasures. Gathering information and educating people about the danger posed by extreme right organisations is paramount here. This is a task for society as a whole. Here, platform operators are responsible for creating the technical environment for developing early warning systems that could quickly identify hate accumulation on social media to protect NGOs, activists, journalists and other Internet users from online attacks. Access to publicly available information on Facebook should again be possible via an interface. The large proportion of right-wing extremist activity in social media must be meticulously documented and worked on.

First and foremost, however, users of social media and digital civil society are called upon to show moral courage on the Internet, to support the victims of shitstorms irrespective of their own political views, and to prevent the division of society through respectful dialogue – including with those very people who spread hate speech. Hate campaigns purposely disguise their authors in order to suggest the impression of widespread social anger. Organisers of grassroots social media campaigns should therefore call upon their sympathisers to clearly commit to their own campaign with the help of hashtags, and to follow the principles of respectful communication without exception.

Media-representatives, especially social media editors, must be sensitised to the problem. Active moderation can lead to a significantly improved discussion culture and opens the comment sections to a broader group of users. Moderators can respond to constructive and differentiated comments in a targeted manner and thus contribute to greater visibility of these comments. Like numbers that are easy to manipulate thus lose credibility, and media can prevent their comment sections and the reach of their brand from being misused as a free platform for anti-constitutional and extremist groups.

Last but not least, the Federal Office for the Protection of the Constitution, law enforcement agencies and the judiciary must focus on social media much more when combating extremist structures. The state must be put in a position to be able to protect its institutions and democratic civil society by means of the rule of law from targeted attacks on the Internet. The NetzDG has so far fallen short of its target in this respect. If the organisations responsible for coordinated hatred are found to be anti-constitutional, their scope for action must be consistently restricted by means of constitutional procedures.

Further reading

Baldauf, Johannes, Miro Dittrich et al. (2017)

Toxische Narrative: Monitoring Rechts-Alternativer Akteure

Bednarz, Liane and Giesa, Christoph (2015)

Dangerous citizens: The new right reaches for the middle

Bruns, Julian, Kathrin Glösel and Natascha Strobl (2014)

The Identitarians: Handbook on the youth movement of the New Right in Europe

Davey, Jacob und Ebner, Julia (2017)

The Fringe Insurgency: Connectivity, Convergence and Mainstreaming of the Extreme Right

Ebner, Julia (2018)

Anger: What Islamists and right-wing extremists are doing to us

Ley, Hannes and Görig, Carsten (2018)

#ichbinhier: Together against fake news and hate in the net

Nagle, Angela (2017)

Kill All Normies: Online Culture Wars From 4Chan And Tumblr To Trump And The Alt-Right.

Phillips, Whitney (2015)

This is why we can't have nice things: Mapping the relationship between online trolling and mainstream culture.

Rost, Katja, Lea Stahel, und Bruno S. Frey (2016)

Digital social norm enforcement: Online firestorms in social media.

Wagner, Thomas (2017)

The fear makers: 1968 and the New Right

Weiß, Volker (2017)

The authoritarian revolt: The New Right and the fall of the West

Wendling, Mike (2018)

Alt-Right: From 4Chan to the White House

Footnotes

1. Council of Europe, Committee of Ministers, Recommendation No. (97) 20
 2. See also Federal Agency for Civic Education: www.bpb.de/apuz/231307/medien-im-mainstream?p=all
 3. Mudde, Cas: "The populist zeitgeist." *Government and opposition* 39.4 (2004): 541-563
 4. "What is this hate speech?" in: "'Go and die' – hate speech and comment culture on the Internet", ed. Amadeu Antonio Foundation, 2015
 5. For further analysis on this topic: Jacob Davey und Julia Ebner, "The Fringe Insurgency: Connectivity, Convergence and Mainstreaming of the Extreme Right", ISD (October 2017). Available here: www.isdglobal.org/wp-content/uploads/2017/10/The-Fringe-Insurgency-221017_2.pdf
 6. www.kleinanfragen.de/bundestag/19/1994-aktivitaeten-des-rechten-internetnetzwerkes-reconquista-germanica
 7. www.fr.de/rhein-main/junge-alternative-auf-strammem-kurs-nach-rechtsaussen-a-1451025
 8. www.youtube.com/watch?v=-IM4wtEr4No
 9. www.verfassungsschutz.de/de/aktuelles/zur-sache/zs-2016-001-maassen-dpa-2016-08
-

