



# Quarterly Report on Cyber Hate (April-June 2017)

By Tamás Berecz and Charlotte Devinat

Project Research - Report -  
Remove: Countering Cyber Hate  
Phenomena

INACH



Supported by the Rights, Equality and Citizenship (REC) Programme of the European Union

## **Executive Foreword**

This publication was written within the framework of the ***Research – Report – Remove: Countering Cyber Hate Phenomena*** project of the International Network Against Cyber Hate (INACH); funded by the European Commission Directorate-General for Justice and Consumers. The duration of the project is 2016-2017, and its aim is to study, document and report on online hate speech in a comparative and comprehensive way; and to establish structures for a transnational complaints system for instances of cyber hate.

Hate speech is intentional or unintentional public discriminatory and/or defamatory statements; intentional incitement to hatred and/or violence and/or segregation based on a person's or a group's real or perceived race, ethnicity, language, nationality, skin colour, religious beliefs or lack thereof, gender, gender identity, sex, sexual orientation, political beliefs, social status, property, birth, age, mental health, disability, disease.

This report was completed with the participation of the different members of the Network and partners in the project, namely the Zivilcourage und Anti-Rassismus-Arbeit (ZARA) from **Austria**, the Movimiento contra la Intolerancia (MCI) from **Spain**, jugendschutz.net from **Germany**, the Ligue Internationale Contre le Racisme et l'Antisémitisme (LICRA) from **France**, the Inter-Federal Centre For Equal Opportunities and Opposition to Racism from **Belgium** (now called Unia), and the Magenta Foundation from the **Netherlands** (MDI); who provided most of the data this report is based upon.

## **Legal Disclaimer**

This publication has been produced with the financial support of the Rights, Equality and Citizenship (REC) Programme of the European Union. The contents of this publication are the sole responsibility of the International Network Against Cyber Hate and can in no way be taken to reflect the views of the European Commission.

## Table of content

I. Introduction -	-	-	-	-	-	-	-	-	-	3.
II. Drivers, trends and tools	-	-	-	-	-	-	-	-	-	3.
1. Drivers	-	-	-	-	-	-	-	-	-	3.
2. Trends	-	-	-	-	-	-	-	-	-	4.
3. Tools	-	-	-	-	-	-	-	-	-	4.
III. Data Collection and Analysis	-	-	-	-	-	-	-	-	-	7.
1. Hate Type Analysis-	-	-	-	-	-	-	-	-	-	7.
2. The Prevalence of Cyber Hate on Different Platforms	-	-	-	-	-	-	-	-	-	9.
3. Actions Taken by Partner Organisations Against Instances of Cyber Hate	-	-	-	-	-	-	-	-	-	12.
4. Removal Rate	-	-	-	-	-	-	-	-	-	13.
5. Legality of Instances of Cyber Hate	-	-	-	-	-	-	-	-	-	16.
IV. End Remarks	-	-	-	-	-	-	-	-	-	17.

## I. Introduction

As a fundamental part of the Research - Report - Remove: Countering Cyber Hate Phenomena project, INACH collects data from all project members from multiple countries (Austria, Belgium, France, Germany, the Netherlands and Spain) on a monthly basis. We collect and merge these pieces of data in order to synthesise a comprehensive and extensive picture of cyber hate in Europe in the 21st century. In this report, the data collected between November and December 2016 will be explored and discussed. Furthermore, INACH also - with the help of the project partners - collects information on drivers, trends and tools that lie behind online hate speech.

As it will be denoted later on with the data collection regarding hate types, antisemitism, racism, anti-Muslim hate and anti-refugee hate were the highest on the list. Regarding the possible justification and explanation as to why those hate types were so high within those three months, the exploration of those drivers, trends and tools that were reported by each of the project partners will be observed in the first place to enable a better understanding of the phenomena.

## II. Drivers, trends and tools

### 1. Drivers

Regarding the new **drivers**, LICRA considered the legislative elections of June and the post presidential elections as drivers. Emmanuel Macron ran against Marine Le Pen in the second round of the French presidential elections. It is the second time that the National Front was in the second round throughout the Fifth French Republic. The first time, Jacques Chirac ran against Jean-Marie Le Pen and he won with 82, 2 %, whilst this time Macron won with 66 %. Another key-element was the high percentage of blank and null vote, which was 9%. Moreover, this election had the highest number of abstentions since 1969 (33, 4%).

Unia in Belgium found that, in May, after having posted a discriminatory cartoon on his Facebook page, the youth president of the Flemish National Alliance (NVA) resigned. Unia was also contacted by an author of hateful messages that had been informed by the police of his inappropriate behaviour, who wanted to know more about the limits of freedom of speech. Further, in June, Unia noted an interesting case law. On 30 June 2017, the criminal court of Antwerp convicted a person who created and administrated a public Facebook page “for the protection and the promotion of the Flemish identity” called Vlaamse Verdedigings Liga (Flemish Defense Group). In its decision, the criminal court of Antwerp ruled that the messages and the liking of the reactions could only be interpreted as hatred against people of different origin. As administrator of the page he had the possibility to remove some comments but he did not. On the contrary, he liked them. The court considered that the comments, cartoons and messages posted by other users

on the page are pointed at people of different origin by naming them apes and insects and expressing the opinion that they “should go home”. The posts provoked very violent reactions in their expression of racism and have been “liked” by the administrator. As the page is public, everyone can access and post comments or reactions. By establishing such an open page, he created a forum and thus a group that repeatedly incited to racism. When Unia filed a complaint against the administrator, the page had 6.502 followers (April 2015). At the time the page was put offline (August 2016) the number had gone up to 23.000 followers. The administrator was convicted for incitement to hatred, discrimination or violence (article 20, 4° anti-racism law) and membership of a group that repeatedly incites to racism (article 22 anti-racism law). The sentence is an imprisonment of 10 months (the maximum sentence would have been 12 months) and a financial penalty of 100 euros, both with adjournment during a period of 3 years. This means that if he is acting this way again within 3 years the sentence will be executed. Additional information is accessible on Unia’s website (Dutch and French only).

## 2. Trends

Regarding the new **trends**, Jugendschutz.net found out that in June, the passing of a bill legalising same-sex marriage in Germany triggered a wave of homophobic hate speech online.

LICRA denoted that in May, trends such as anti-Arab sentiment and hate speech related to the “suburb scum”, “prayers in the streets”, and “Asiatic prostitution” emerged during the legislative elections, as well as denial of the Holocaust and of the Oradour-sur-Glane massacre and anti-Semitic rap. In June, the trends were anti-Muslim hate with the use of words such as “muzz”, racism against the black community with the use of colonialist vocabulary; Transphobia and anti-gay marriage and adoption sentiment, and anti-Semitic insults against Simone Veil, a Holocaust survivor and human rights fighter.

MCI explained that in May a trend of provocative posts on the glorification of terrorism was used as a way to protest against Security Citizen Law.

## 3. Tools

Regarding the new **tools**, Jugendschutz.net in Germany found out that, in April, the fake news Facebook page "Die Wahrheit '24h News'" (The truth '24h news'") was shut down. However, pages with the same name were later still available on YouTube and other Social Media platforms. "Die Wahrheit '24h News'" and other similar pages use a mixture of fake news, rumours and reports by reliable news websites, which have been taken out of context. This mixture was used to give the websites a guise of credibility. Supposed "revelations" of crimes committed by Muslims or migrants promoted the image of a weak state and law enforcement, bad asylum policy and a defenceless "German people". In May, so-called "orcpostings" emerged. They are Memes derived from the Lord of the Rings movies and franchise combined with posts showing refugees as orcs.

The idea of "orcpostings" originates in the American "Alt-right movement" but quickly spread to Germany. In a symbolic campaign the identitarian movements of Germany and several other countries blocked the way of a rescuing ship in the Mediterranean Sea with a dinghy. They claimed that NGOs rescuing refugees from drowning are in fact human traffickers and that less refugees would come to Europe if more would drown in the sea. The campaign was designed to resemble campaigns of environmental activists like Sea Shepherds or Greenpeace. In the 9<sup>th</sup> of May, the edition of the "IS" magazine "Rumiyah" was published in German. Amongst anti-Christian and anti-Shia content it also contains a step-by-step guide for truck attacks.

LICRA, noted that, in May, during the legislative elections in France, one of the candidates, Nicolas Rousseaux (Les Républicains, right-wing party), proposed this electoral poster (and then an electoral flyer): "You are fed up of being duped by NKM? (Nathalie Kosciusko-Morizet, a right-wing political candidate who is criticized by other members of the right and far-right wing to be a "bobo")". "Immigration is a chance for France!"



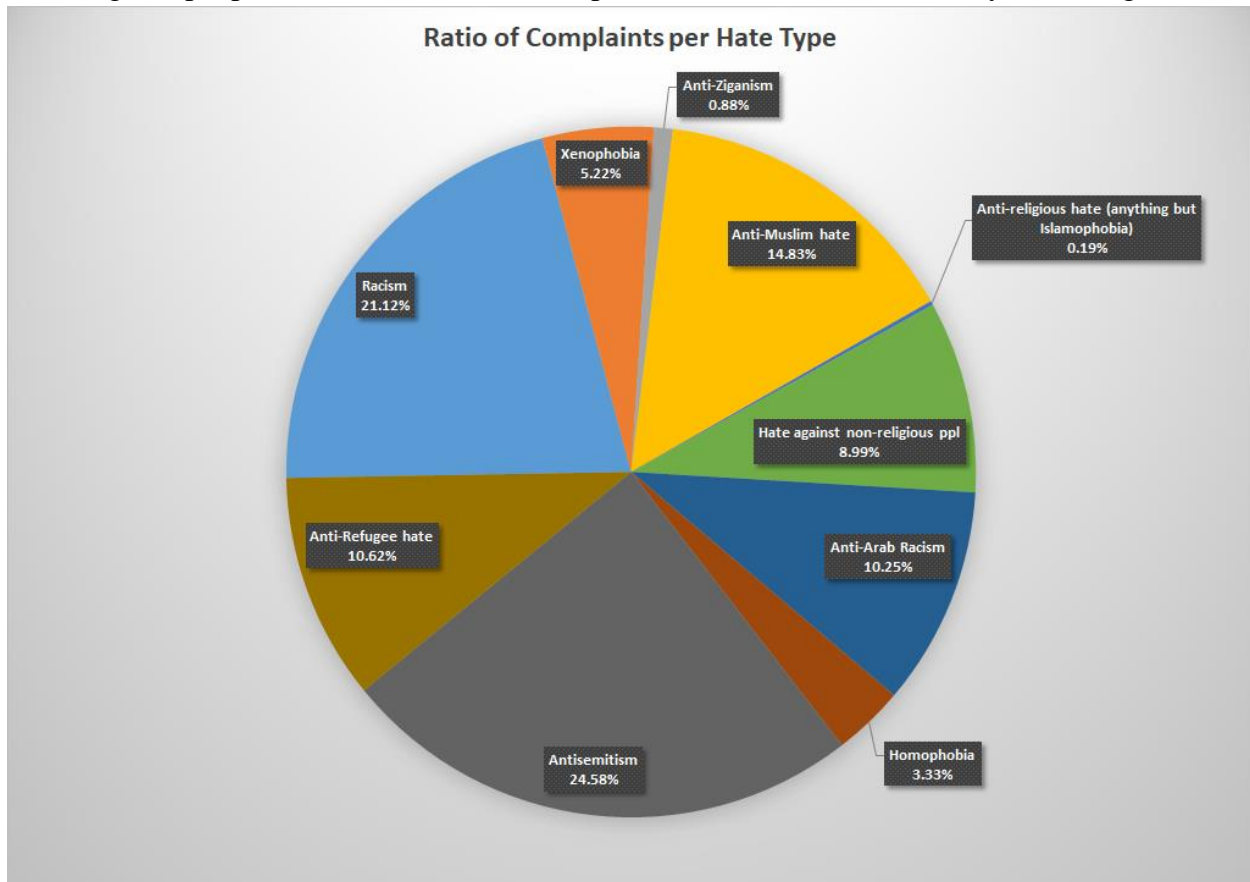
LICRA also made a new referral to the prosecutors against Vincent Reynouard for denial videos on Youtube. These videos were posted after the visit of Emmanuel Macron, in Oradour-sur-Glane in June. On 10 June 1944, the village of Oradour-sur-Glane in Haute-Vienne was destroyed, when 642 of its inhabitants, including women and children, were massacred by a Nazi Waffen-SS company. Vincent Reynouard is “a specialist” in the denial of the massacre of Oradour-sur-Glane. He has already been sentenced many times by the French courts. He also made videos about the denial of the Holocaust. LICRA has also made a referral against the far-right and anti-semitic rapper, Amelek. Some titles of his songs are: “rioters of faggots”, or “whores for Jews”. Amelek is close to the ultra-royalists and ultra-Catholics such as Civitas. His father is an executive of the National Front in Picardie. He is also known for taking videos of himself making the gesture of the “quenelle and for insulting the Prefect as a “Zionist sucker”. Amalek is a member of this new trend, the identitarian rap, promoted by the fachosphere as, for example, Dieudonné or Alain Soral. Tweet posted by Amalek (answer to Jacques Attali): “You of course prefer Kaaris (a French rapper), this bitch of Bourimech (a young entrepreneur from the suburb of Bondy) who calls people to “gun blaze on the Aryans”#FuckingDirtyJew” #AtTheOwen. In June, the same day Simone Veil died, some public figures have tweeted anti-semitic insults and anti-abortion theories. Henry de Lesquen, already sentenced for a tweet against Simone Veil in August 2016 “the boisterous Simone Veil, “Auschwitz survivor”, is 88. As far as I know, she is fine” (in a previous tweet, he wrote: “I am amazed of the longevity of Holocaust survivors dead at more than 90 years old. Did they really live through what they said?”). He tweeted on the 30<sup>th</sup> of June: “Because of the law of 1979, Simone Veil is responsible for the Holocaust of French children: 8 millions dead children in 40 years”. He also wrote on Twitter that Simone Veil was an “anti-French, anti-Catholic and cosmopolite rubbish”. Henry de Lesquen who was the director of a far-right radio channel has been dismissed a few days after these new anti-semitic tweets. A Front National regional councillor, Samuel Potier, tweeted: “May she rest in peace and contemplate the millions of victims her law has caused”.



### III. Data Collection and Analysis

#### 1.Hate Type Analysis

Now that a background information about drivers, trends and tools was outlined, it is possible to move on to the data collection and analysis part of this report, with a better understanding of the general atmosphere in Europe. During the monthly data collection INACH put particular focus on 10 different hate types, due to their prevalence and pervasiveness on the internet. These hate types are the following: racism, xenophobia, anti-Ziganism (hate against the Roma community), anti-Muslim hate (ie. Islamophobia), anti-religious hate (everything but Islamophobia), hate against non-religious people, anti-Arab racism, homophobia, antisemitism and finally anti-refugee hate.



These hate types fluctuate immensely from month to month. Some hate types are very prevalent in some countries, while they are scarce in others. Also, the differences between INACH's project partners adds to this variegation. LICRA, in France, mainly focuses on antisemitism, for instance, therefore they always deliver a high number of cyber hate cases against the French Jewry. Other partners focus more on anti-Muslim hate or other types of racism, so their numbers tend to be higher in different hate types. The last factor affecting the numbers is the difference in size and funding amongst the project partners. Jugendschutz.net is a major organization in Germany with a lot more manpower and resources than, for instance, ZARA in Austria or MCI in Spain. Hence, the number of cases we receive from Germany tend to be a lot higher than from other countries



where our project partners reside. However, altogether, the numbers received from all partners give a fairly extensive and wide insight into cyber hate in Europe.

The collected numbers in the first quarter of 2017 showed that the trend of falling numbers in anti-Muslim hate continued in 2017 and the numbers in the second quarter substantiated this trend. In 2016, anti-Muslim cases were the first among the hate types in the second quarter, then fell to 18.28 per cent from 22.32 per cent in the third quarter and kept falling to 17.68 per cent in the final quarter of the year. This tendency has bled into the first and second quarters of 2017, with the numbers falling further to 16.23 per cent and then to 14.83 per cent. Thus, the ratio of anti-Muslim instances of cyber hate collected by INACH fell even further below one-fifth of all cases. However, it is still third most prevalent hate type.

General racism took the first place from anti-Muslim hate in the 3rd quarter of 2016, by rising to 20.71 per cent from 20.02 per cent (due to our methodology, antisemitism, anti-Arab racism, anti-Ziganism and anti-refugee hate are all excluded from these numbers). This trend also continued as the hate type has risen from 20.71 per cent to 21.81 per cent, keeping its unflattering first place among the hate types in the final quarter of 2016. And this tendency continued in the first quarter of 2017. Racism kept its first place and its ratio has risen to 25.62 per cent, which means that instances of cyber hate involving general racist slurs, ideologies, etc. were responsible for more than a quarter of the cyber hate cases handled by INACH and its partners. However, this trend has turned around during the second quarter of the year and the ratio of racist cases has fallen to 21.12 per cent, relegating racism to the second position after more than nine months spent in first place.

The third trend, a rise in antisemitic cases, also continued into the fourth quarter of 2016. Antisemitism first rose from 17.1 per cent to 18.1 per cent in between the second and third quarters of 2016, taking the third place from anti-refugee hate (any kind of cyber hate that attacks people solely based on the fact that they are refugees or migrants) and kept on rising to 20.67 per cent, taking the second place among hate types in the last two months of 2016. Just like the other tendencies discussed above, this trend also continued in the first three months of 2017. Hence, antisemitism kept its second place among the hate types, while its ratio has risen to 24.19 per cent. Unlike the trend in racist cases, this trend have not abated. The ratio of antisemitic cases have risen to 24.58 per cent in the second quarter of 2017; a minor rise, but still very meaningful since, with the fall in racist cases, it puts antisemitic cases into the first place for the first time since INACH started the data collection for this project. This means that racist and antisemitic cases of online hate speech are responsible for almost half of the cases handled by INACH and the project partners.

Anti-refugee hate also stayed on the same track as between the second, third and fourth quarters of 2016. It fell from 18.94 per cent to 16.42 per cent during the autumn of 2016 just to diminish further to 15.04 per cent during the winter. In the first quarter of 2017 the ratio has fallen further to 14.12 per cent. A trend that continued into the second quarter of the year with this hate type

falling further to 10.62 per cent, solidifying its off the podium place after being one of the most prevalent hate types in the beginning of 2016.

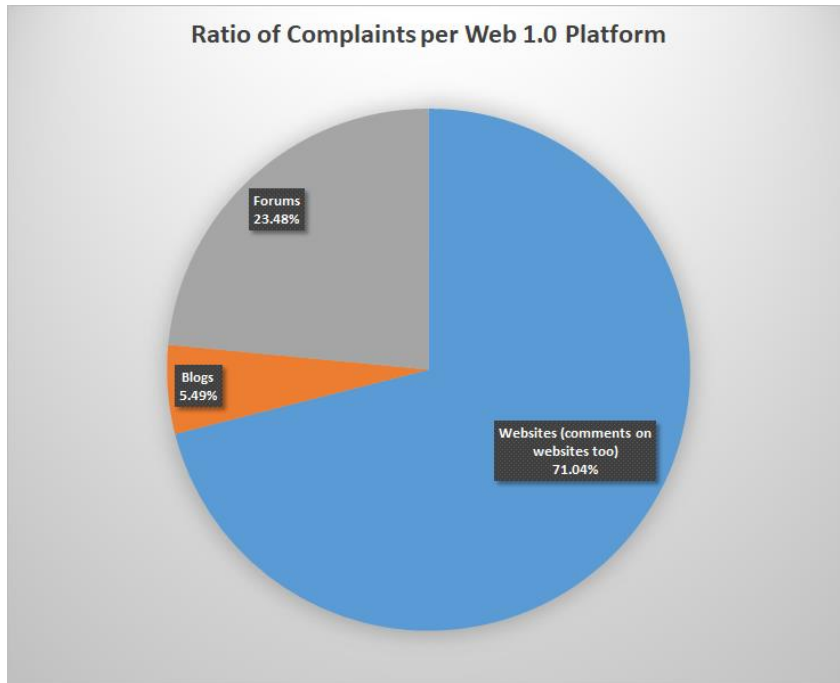
These hate types were followed by hate against non-religious people that saw a sharp rise between the 2nd and third quarters of 2016 from 4.83 per cent to 8.21 per cent, a whopping 100 per cent increase (this is based mainly on data received from Germany). This rise continued into November and December, reaching 11.60 per cent. However, this trend has gone through major ups and downs since the start of 2017, where the ratio of this hate type among all monitored hate types has fallen to 6.76 per cent during the first quarter, an almost 100 per cent fall, just to rise to 8.99 per cent again.

Anti-Arab racism - after a sharp rise between the second and third quarters of 2016 - saw a minor drop in the fourth from 5.6 per cent to 4.59 per cent. This trend is among the few that has changed in the first quarter of 2017, since such instances of online hate speech reached 5.56 per cent, raising the hate type's ratio back to virtually its highest level, until the second quarter of the year, during which this hate type showed a whopping - almost 100 per cent - rise to 10.25 per cent; making it a contender for the dubious podium. The ratios of all other hate types are either just above or way below 5 per cent. Xenophobia - after a minor fall from 6.06 per cent to 4.48 - basically stayed the same by falling a bit further to 4.21 per cent in the first quarter of 2017 and then rising to 5.22 per cent in the second.

## **2. The Prevalence of Cyber Hate on Different Platforms**

When it comes to platforms where cyber hate is flourishing, maybe it is not surprising that social media trumps Web 1.0 platforms by a magnitude. However, there are still websites, blogs and forums on the internet that spew hate or provide a platform for people to post hateful messages.

Websites are still the most widely used tools among Web 1.0 platforms to spread cyber hate, either by producing it or by providing a platform for people to post vile, violent and hateful comments. During the third quarter of 2016, 69.35 per cent of recorded cases of online hate (that appeared outside of social media) were posted on websites, a minor fall from 72.73 per cent. However, this trend has turned around in the fourth quarter, during which the percentage of online hate posted on websites rose to 74.53 per cent. Just like most of the trends that started in 2016, this tendency also continued in the first quarter of 2017. In other words, the ratio of complaints about hateful content posted on websites has grown even further to 77.08 per cent. However, just like some other trends in hate types, this pattern has also turned around somewhat, with websites observing a 6 per cent drop to 71.04 per cent. Still, this substantial drop has done nothing to endanger the first place of this platform among web 1.0 platforms.

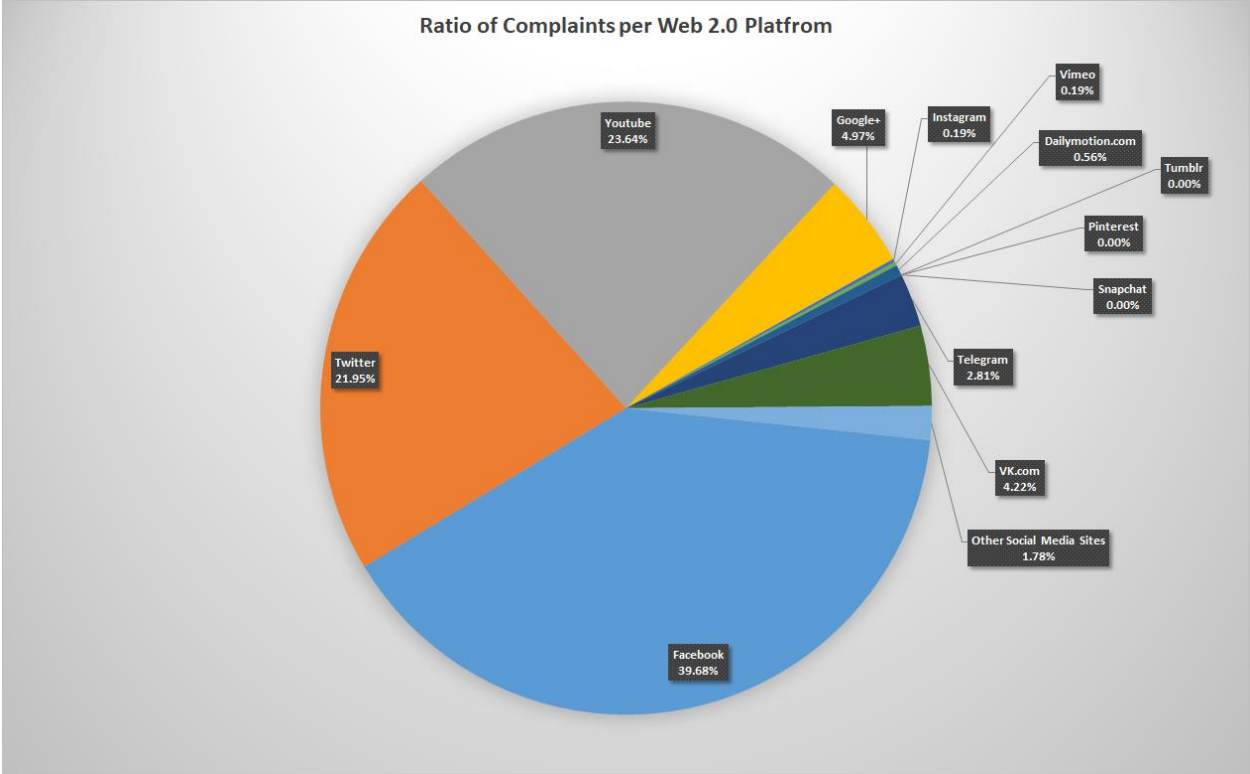


Cases on blogs fell from 19.01 per cent to 12.9 per cent in the third quarter of 2016 and fell even further in the fourth to 9.94 per cent, and then fell further to 8.33 per cent in the first three months of 2017 and to 5.49 per cent during the second quarter, solidifying another trend in the data. Forums kept their second place. After seeing a sharp rise from 8.26 per cent to 17.74 per cent in the third quarter of 2016, cases on forums fell to 15.53 per cent

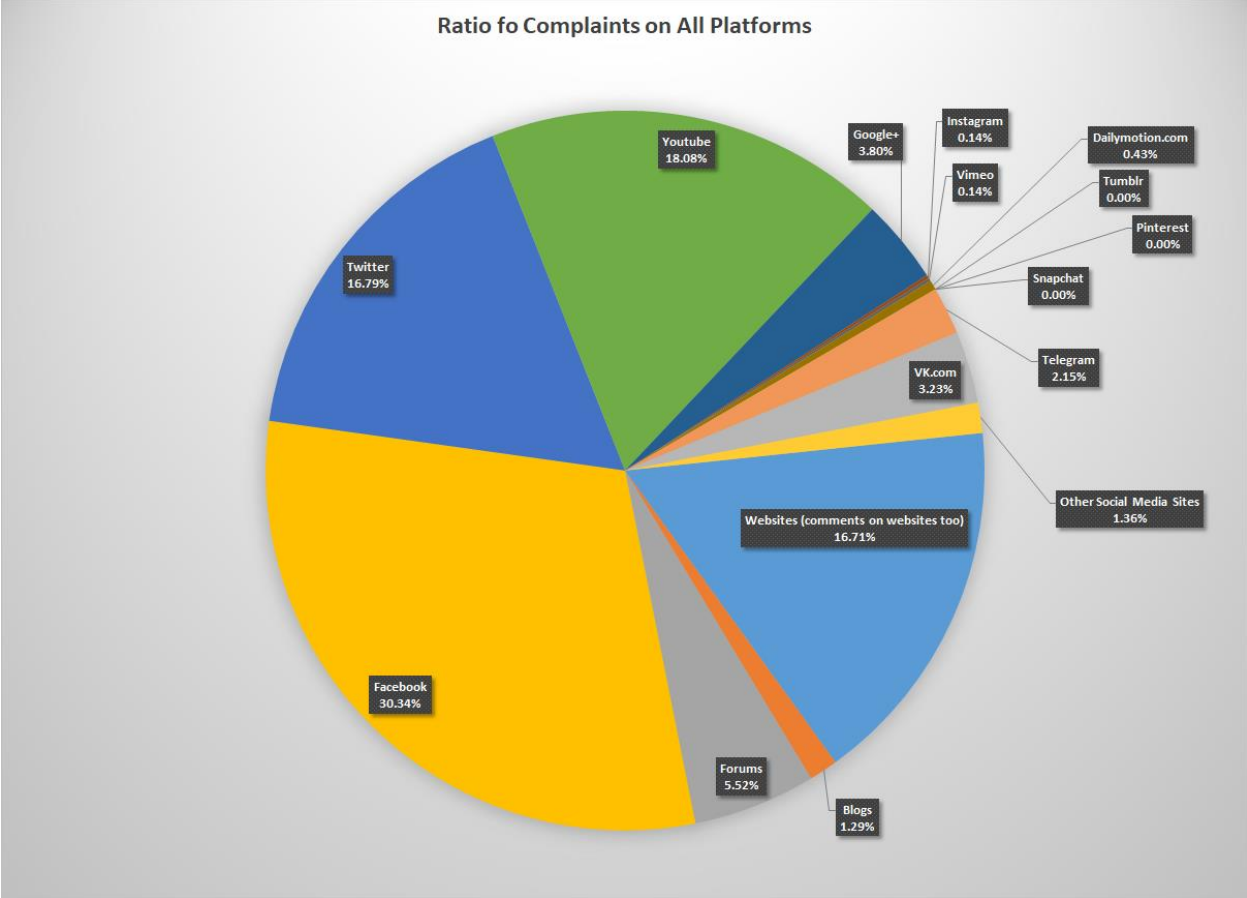
at the end of 2016 and then further to 14.58 in the beginning of 2017. However, this pattern has turned around in the second quarter, in which the ratio of registered cases in forums rose to 23.48 per cent; not just keeping their second place among traditional platforms, but rising to a previously unseen level.

When examining social media sites, the three giants discussed in our previous quarterly reports, still rule the online sphere of cyber hate. Facebook, however, has seen a sharp fall from 46.06 per cent to 37.54 per cent in 2016 and this ratio has hardly changed in the first quarter of 2017 by rising to 39.38 per cent, a negligible change that continued into the second quarter with a rise to 39.68 per cent. However, nothing can threaten the first place of Facebook among the social media platforms, not even the two other major members of the dominating triumvirate. Twitter kept its second place in the first quarter of 2017, by rising from 23.25 per cent to 28.7 per cent in the third quarter of 2016 and then basically staying there at 28.07 per cent, whilst observing a minor 3 per cent drop in 2017, bringing it to 24.78 per cent. This downward trend continued, however, into the second quarter, with the social media platform dropping further to 21.95 per cent, becoming third for the first time since the start of our data collection. Finally, YouTube has risen to second place. The platform observed a minor, approximately 3 per cent rise, from 17.89 per cent to 21.40 per cent at the end of last year and then it rose further to 24.34 per cent in the beginning of 2017. The second quarter of the year did not bring much change, YouTube's ratio fell less than 1 per cent to 23.64 per cent. All other platforms are completely dwarfed by the three major players. All of them are around or under 2 per cent. There are two exceptions to this, however. Both Google+ and VK.com have shown a remarkable rise during the second quarter of this year. Google+ has risen to 4.97 per cent, and VK has also reached above 4 per cent (4.22%). Telegram is the third among

the less popular platforms that has shown a surprising rise, reaching 2.81 per cent. These minor platforms took a bite from the ratios of all three major platforms, however they are still far away from endangering the hegemony of Facebook, Twitter and YouTube.



If the numbers of Web 1.0 and 2.0 platforms are merged, it becomes crystal clear that social media sites are still the pivotal platforms when it comes to the spreading of cyber hate. These platforms provide a cheap or even free tool for people and extremist groups to deliver their message to a gigantic audience. Hence, the dominance of the three aforementioned giants remains intact in the same order previously described, but the ratio of cases recorded on websites falls to 16.71 per cent (18.62 per cent in the first quarter of 2017; 7.84 per cent in second quarter of 2016, 9.54 per cent in the third and 16.42 per cent in the fourth) and the ratio of cases on blogs and forums falls to negligible levels.



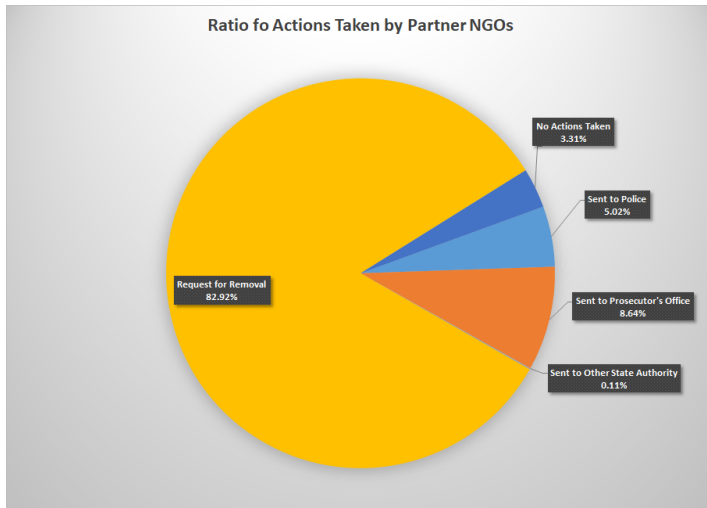
These numbers clearly show that social media sites have completely taken over and fundamentally changed the landscape of cyber hate by letting their users spew hateful and violent content against minority communities in the form of memes, conspiracy theories, fake news and other viral content. Even more alarmingly, these platforms made it possible to extremist groups and individuals to deliver such content to users who do not actively seek it out, paving the way for radicalization among adolescents and young adults.

**3. Actions Taken by Partner Organisations Against Instances of Cyber Hate**

Partner organizations that participate in the project mainly focus on getting instances of cyber hate removed from social media and other platforms. Therefore, it is not surprising that, among the reported actions that had been taken by our partners, request for removal is still the unquestionable leader with 82.92 per cent, even though it has gone through a major, almost 10 per cent drop from 91.28 per cent in the first quarter.

The number of cases where no actions were taken rose between the 3rd and fourth quarters of 2016 from 3 per cent to 7.50 per cent, which was a major change in the previous trend. However, it seems that this change was not part of a developing tendency, just a temporary jump, since the

ratio of such cases has fallen back to 2.59 per cent; a welcome development that continued into the second quarter, with such cases only rising to 3.11 per cent in the second quarter of 2017.



Finally, sometimes INACH discovers hate speech online that is so serious that it is not enough to just report it to the platform where it had been posted, but the case has to be reported to state authorities too. This can be the police, the prosecutor's office or any other law enforcement agency. Altogether, cases forwarded to these authorities counted for 8 per cent of all cases in the second quarter of 2016 and in the 3rd quarter they have seen a 2 per cent rise, reaching

10 per cent of all cases. In the fourth quarter, this trend changed with such cases making up less than 6 per cent altogether. This did not changed, since and the ratio of these cases still hovered around 6 per cent in the first three months of 2017. However, the second quarter brought a major shift, with such cases rising to more than 13 per cent, with cases forwarded to the prosecutor's office rising to previously unseen levels at a whopping 8.64 per cent.

#### **4. Removal Rate**

Removal rates can be very varied and inconsistent when it comes to the three big social media platforms. INACH's project partners received most of their complaints on Facebook, Twitter and YouTube. All other platforms are way below the numbers of the aforementioned triumvirate.

The removal rate of Facebook was fairly high on average. Between May and July 2016, across all six partner countries, the platform's removal rate was 78.57 per cent. However, this ratio has fallen in the 3rd quarter by more than 5 per cent to 73.4 per cent, which signalled a bad trend that did not turn around in the fourth quarter either. In November and December of 2016 Facebook's removal rates have fallen to an abysmal 50.90 per cent. Sadly, this trend has not gone through a major change in 2017, even though Facebook's removal rates have risen to 58.66 per cent in the first quarter, they quickly fell back to 49.25 per cent during the second. This removal rate is absolutely atrocious.

Twitter used to do even worse than Facebook, with its removal rate falling from 67.01 per cent to 66.34 per cent in the third quarter of 2016 and to 24.68 per cent in the fourth. This horrendous trend, however, has turned around in the first quarter of 2017 and Twitter's removal rates rose to

a whopping 74.6 per cent. We were hopeful that this sudden upward trend will not change during the second quarter of the year. Sadly, our hopes were premature and Twitter’s removal rates have fallen to 54.50 per cent.

The biggest fall in 2016, however, was seen by YouTube. The platform’s removal rate was very close to Facebook’s with an exceptionally high 86.27 per cent during the second quarter of 2016. As we mentioned it in our previous reports, that high ratio was fairly unusual for the platform and that statement was underpinned by YouTube’s abysmal numbers in the third quarter of 2016. The platform’s removal rate fell to 62.22 per cent, and then to 52.63 per cent in the final quarter of 2016, sadly solidifying this trend. However, in the beginning of 2017, YouTube’s removal rates rose to a staggering 83.33 per cent. Just like the positive changes in Twitter’s numbers, this positive shift was also short-lived. YouTube’s removal rates have fallen to 64.57 per cent, an abysmal showing for the company that is also sadly their average.

“Google+ is used by a lot less people than the previous three platforms, and the number of complaints on the platform is a lot lower, but these are still not sufficient excuses for the very low removal rate by the site, which is a meagre 30 per cent.” We wrote this in our first quarterly report in 2016. The grim picture described in that report became even grimmer in the second quarter. Google+’s removal rates almost halved by falling to 16.67 per cent. However, as arguably the only positive development in the fourth quarter, the platform’s removal rate rose to 47.06 per cent in the winter of 2016. Sadly, this positive development did not become a tendency. Google+’s removal rates dropped to the basement in the first three months of 2017, reaching 5.26 per cent, which is the worst removal rate for the company, since the start of INACH’s data collection. What is even more worrying is that, in the second quarter of 2017, the company was able to do even worse by not removing any content based on the complaints of our partners; something that had never happened before.

Name of Platform	Percentages of Cases Removed	Percentages of Cases Not Removed
Websites (comments on websites too)	33.33%	66.67%
Blogs	50.00%	50.00%
Forums	88.00%	12.00%
Facebook	49.25%	50.75%

Twitter	54.50%	45.50%
YouTube	64.57%	35.43%
Google+	0.00%	100.00%
Instagram	0.00%	100.00%
Vimeo	100.00%	0.00%
Dailymotion.com	66.67%	33.33%
VK.com	50.00%	50.00%
Telegram	63.33%	36.67%
Other Social Media Sites	33.33%	66.67%

The issues NGOs and users face - sadly - still have not changed either. If we take a look at individual removal rates in different partner countries in different months, we can see the biggest problem NGOs that fight cyber hate have with these sites. They are outrageously inconsistent in their removal rates between countries and in cases that are very similar to each other. It is understandable that these companies' community guidelines are interpreted in relation to given countries national laws, but the guidelines are the same globally, therefore, the same infractions should be removed everywhere. However, that is most definitely not the case. Removal rates are highly influenced by the amount of complaints given social media site receives about an instance of online hate, and by who the complainer is. If it is an authority or a very well established local NGO, or other civil society organization that is a trusted reporter or flagger, it is much more likely that the hateful content will be removed; just like when a lot of people complain about a certain content. This should not be the case. Illegal content and content that violates the guidelines should be removed globally and universally, irrespectively of the number of complainers or who the flagger is.

Taking all this into account, it is very aggravating that removal rates variegate vastly between countries. For example, jugendschutz.net in Germany had a 38.71 per cent removal rate on

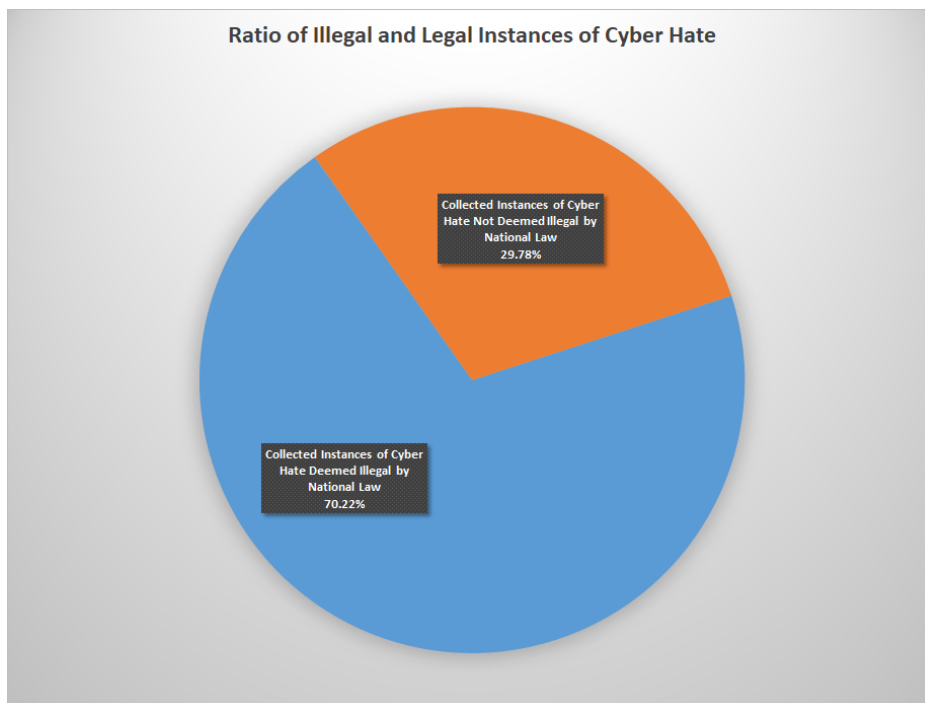


Facebook in June, whilst ZARA in Austria had a zero per cent success rate in removals and LICRA in France produced similar numbers to jugendschutz.net's (41.67 per cent) in that month.

There are very similar problems with Twitter. Jugendschutz.net had a 0.00 per cent removal rate on Twitter in June, whilst LICRA had a whopping 87.50 per cent success rate, but, for instance, ZARA in Austria had zero removal success, and finally LICRA only managed to get 30 per cent removed from the platform in that month.

The list could be continued, but the point is already clear. There are major differences in removal rates on a monthly basis and between countries. This insinuates that social media companies interpret their own rules and guidelines subjectively and arbitrarily. This arbitrariness makes the job of NGOs and other organizations extremely hard and frustrating, whilst it also nurtures an enabling culture online towards extremist groups and people who hold extreme ideas and ideologies. Highly illegal, violent, hateful and vile contents are left online for months without any real explanation from social media giants, whilst minor and benign infractions are removed within hours. This attitude and the companies' modus operandi must change, if we are ever to have an online community that respects the human rights of all of its members.

### **5. Legality of Instances of Cyber Hate**



What is mainly noted by the data collected by INACH is that, although some cases might be considered hate speech by the public or by INACH members, they might not always be considered illegal. As presented in one of our previous reports, in the second quarter of 2016, 89.58 per cent of reported or discovered instances of cyber hate were deemed illegal by the complaints officers

of our partner organizations. This ratio did not fundamentally change during the third quarter, in which the ratio of cases deemed illegal by our experts was 89.1 per cent, a very minor decrease. It seems, however, that this minor decrease was a harbinger of a major fall during the winter months,

where cases deemed illegal by our experts fell to 78.57 per cent. This negative trend continued into 2017, where these numbers fell further to 71.47 per cent and then to 70.22 per cent.

This means that - in the second quarter of 2017 - 29.78 per cent of cases (the highest this number has ever been) assessed by our officers fell into a murky field, in which the inspected speech is highly offensive, dangerous, demeaning and/or goes against human dignity, yet it does not fall into what given nation state considers as illegal hate speech. Even though this ratio is still not very high - although much higher than the previous 10-11 per cent -, EU member states should pay more attention to hate speech that falls through the cracks of legislation in order to be able to stand up against this destructive social phenomenon even more effectively.

#### **IV. End Remarks**

We do not draw extensive conclusions based on numbers collected during a three-month period. Some minor changes and shifts in the data can be observed, but these minor fluctuations are still not enough to draw up far-reaching conclusions. Such trends cannot be drawn up based on such a relatively small sample size. Therefore, we discuss trends, shifts in the data and the conclusions that can be drawn from them in our extensive and comprehensive yearly report that was published in July 2017.