



**Quarterly Report on
Cyber Hate
(November and
December 2016)**

By Tamás Berecz and Charlotte Devinat

**Project Research - Report -
Remove: Countering Cyber Hate
Phenomena**

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.

Table of content

I. Introduction -	-	-	-	-	-	-	-	-	-	3.
II. Drivers, trends and tools	-	-	-	-	-	-	-	-	-	3.
1. Drivers	-	-	-	-	-	-	-	-	-	3.
2. Trends	-	-	-	-	-	-	-	-	-	3.
3. Tools	-	-	-	-	-	-	-	-	-	4.
III. Data Collection and Analysis	-	-	-	-	-	-	-	-	-	5.
1. Hate Type Analysis-	-	-	-	-	-	-	-	-	-	5.
2. The Prevalence of Cyber Hate on Different Platforms	-	-	-	-	-	-	-	-	-	7.
3. Actions Taken by Partner Organisations Against Instances of Cyber Hate	-	-	-	-	-	-	-	-	-	10.
4. Removal Rate	-	-	-	-	-	-	-	-	-	11.
5. Legality of Instances of Cyber Hate	-	-	-	-	-	-	-	-	-	13.
IV. End Remarks	-	-	-	-	-	-	-	-	-	14.

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 synthesize 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**, MCI found that in December a new anti-arab bias appeared in online hate speech around the anniversary of the so called “Toma de Granada”, i.e. the campaign during which the Catholic Kings’ army “re-conquered” the southern parts of the Iberian Peninsula.

UNIA noted that there were no new drivers as such, but that there had been a slight diversification amongst the types of discrimination. The number of complaints based on sexual orientation increased, as well as a slight increase regarding disability and wealth also being the subject of one of the complaints. The high number of reports concerning hate on websites in December was due to the closing of complaints relating to chain mails at the end of the year. 31 chain mails were closed at the end of that month.

ZARA noted that the presidential election campaign was still going on until December 4th, including hate inciting online comments of politicians of the FPÖ (right-wing freedom party) mainly on Facebook.

2. Trends

Regarding the new **trends**, Juschu found that in November, during Donald Trump’s election campaign, that was supported by the so called “Alt-Right”, one could notice increasing connections and collaborations between the (mainly European based) Identitarian movement and the Alt-Right in the US. Some of the latter called themselves “Identitarian” and referred to the same ideological groundworks. Statements of IM-members emerged on popular Alt-Right websites and profiles,

e.g. leading figure Martin Sellner was quoted on Breitbart, the biggest hub of Alt-Right thought online. In turn, Sellner increasingly published postings and videos in English. Moreover "Pepe the frog", one of the most popular symbols of the Alt-Right, was used by IM-members and German right-wing extremists as part of their social media strategy. Regarding December, on the 15th of the month, a new activist group called "Matefaschisten" (mate fascists, "mate" is a kind of tea from South America) streamed a live video via Facebook that showed them baking a mate cake. Referring to a very popular beverage among young people in Germany was a strategy to address a young target group with right-wing propaganda coming along in an unsuspecting format. During the baking, the activists answered questions from the audience. Their statements alternated between fooling around and neo-Nazi propaganda, as one participant said "Bombs on Israel" for instance. The format "FB-live" was problematic in regard to countermeasures as well; although live videos can generally be reported, reaction time to them are highly restricted by the medium itself. In this specific case, the activists removed the content after the broadcast.

LICRA recorded that François Fillon, an ultra-conservative politician, had been elected in December as the candidate of the right-wing party (Republicans) in the presidential election.

MDI looked at incidents surrounding a new political party, 'Denk', in November. The party focuses on fighting against discrimination and racism. The two Turkish founders and Sylvana Simons were victim of hatred and discrimination. Especially Sylvana was a target for many on- and offline instances of hatred and discrimination. More specifically, there was a video made and pictures posted online in which one could see black people hanging from a tree photoshopped with a picture of Sylvana's head on them. The video was accompanied with an upbeat carnival song about the fact that Sylvana needed to leave Holland. In December, the Black Pete/Zwarte Piet debate started again, as it has for the last 3 or 5 years, during which Zwarte Piet is becoming less and less black. Demonstrations, online and offline were held, both by people against Zwarte Piet because of black face and all the racism that comes with it; and by the people pro Zwarte Piet, as it is - according to them - a holiday tradition. Incitement to hate and violence took place and MDI received several complaints from both sides. Lastly, Geert Wilder's trial for his "we want less Moroccans" expression ended. It lasted 2 years and he was found guilty. The whole trial and the outcome were biased according to Wilders and his followers.

3. Tools

Regarding the new **tools**, Juschu found that the live video streaming feature of Facebook was increasingly used in order to promote hate. Dubious surveys with provocative questions like "Should child abusers be punished harder?" facilitated wide distribution with thousands of likes, comments and shares. FB-live was also problematic since comments could not be reported.

LICRA noted that new platforms have been reported; **RussiaToday or RT news**, the conspiracy and propagandist Russian channel is one. In France, the information on RT's website were mainly

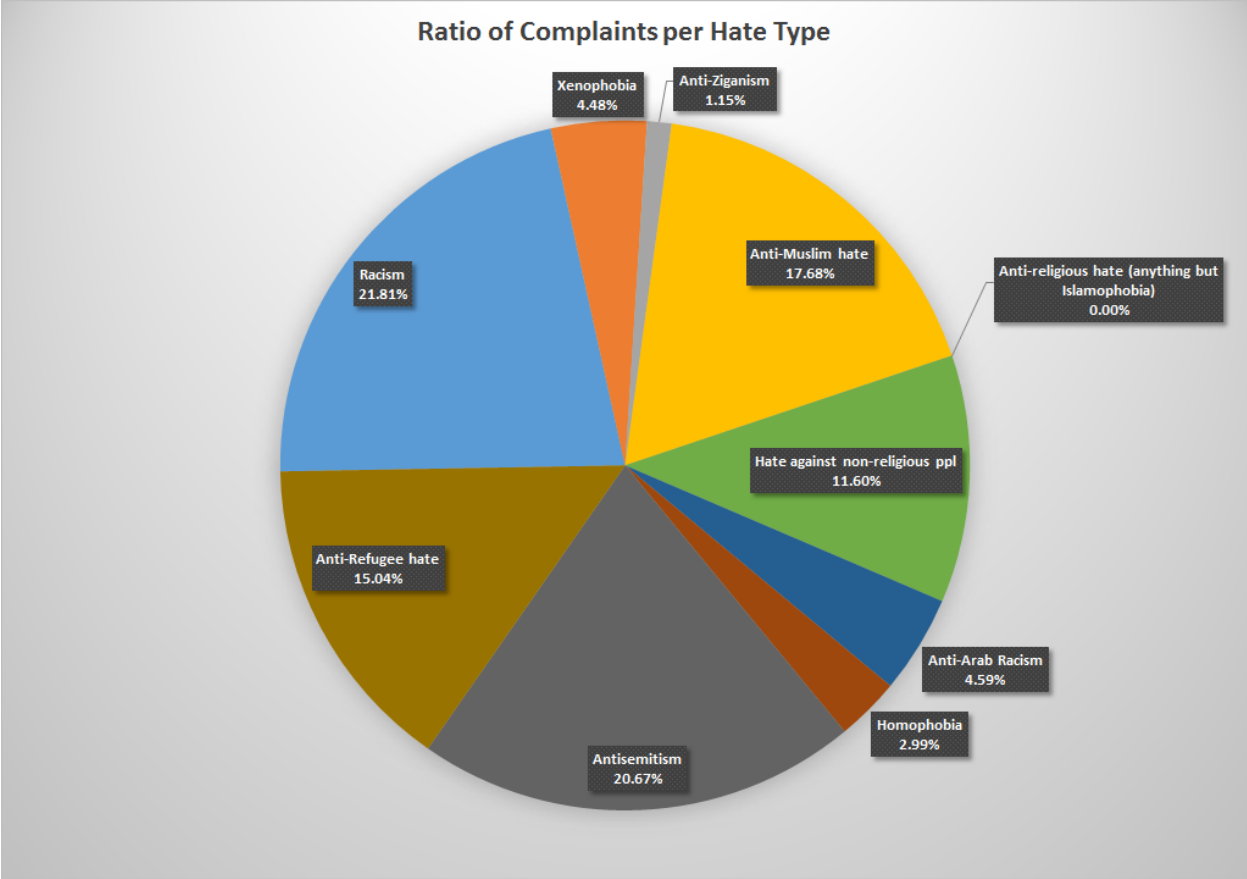
promoted by the “fascist-sphere” (Dieudonné, Alain Soral, Robert Ménard, etc.) Because of the presidential campaign in France, the “Russia” political and diplomatic position was becoming a real political topic; some candidates like Marine Le Pen (National Front) and François Fillon have expressed their sympathy for Vladimir Putin’s government. The development of the popularity of these kinds of platforms (Russia Today, Sputnik News) is in relation with this new topic. Mainly antisemitic comments under articles about Israel’s situation were reported. **Twitch TV** (live streaming video platform and community for gamers – same system as Periscope), is the other one. The issue was that comments and interactions were in live. Moreover, Licra proposed on its website a report form for any user; it appeared that people close to the ideas of the “fascist-sphere” used their form for reporting “anti-White” or “anti-European” racist videos. They all report the same video. Furthermore, in an article on the website bvoltaire.fr (the website of Robert Ménard, Mayor of Béziers very close to the National Front and already sentenced for racist offences), a journalist asked the readers to report a video. The video, titled “A French man humiliated by the scum”, was very likely misinformation.

MCI reported that fake news was a persistent issue.

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 fourth quarter of 2016 (due to the starting point of our data collection within the project framework, the final quarterly report on 2016 focuses on only two months, November and December) show that the trend of falling numbers in anti-Muslim hate continued between the third and fourth quarters of the year. 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. Thus, the ratio of anti-Muslim instances of cyber hate collected by INACH fell even further below one-fifth of all cases. General

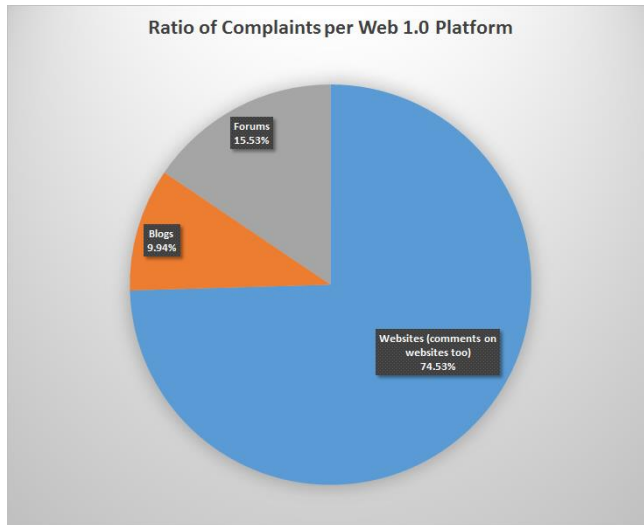
racism took the first place from anti-Muslim hate in the 3rd quarter of 2016, by rising to 20.71 per cent from 20.02 per cent (due to our methodology, antisemitism, anti-Arab racism, anti-Ziganism and anti-refugee hate are all excluded from these numbers). This trend also continued as the hate type has risen from 20.71 per cent to 21.81 per cent, keeping its unflattering first place among the hate types.

The third trend, a rise in antisemitic cases, also continued into the fourth quarter. Antisemitism first rose from 17.1 per cent to 18.1 per cent in between the second and third quarters, 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.

Anti-refugee hate also stayed on the same track as between the second and third quarters. It fell from 18.94 per cent to 16.42 per cent during the autumn of 2016, only to diminish further to 15.04 per cent during the winter; solidifying its off the podium place after being one of the most prevalent hate types in the beginning of 2016. These hate types are still followed by hate against non-religious people that has seen 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. No other hate type reaches 5 per cent after these hate types. Xenophobia has seen another minor fall from 6.06 per cent to 4.48 per cent that arguably put it on par with anti-Arab racism that - after a sharp rise between the second and third quarters - saw a minor drop in the fourth from 5.6 per cent to 4.59 per cent.

2. The Prevalence of Cyber Hate on Different Platforms

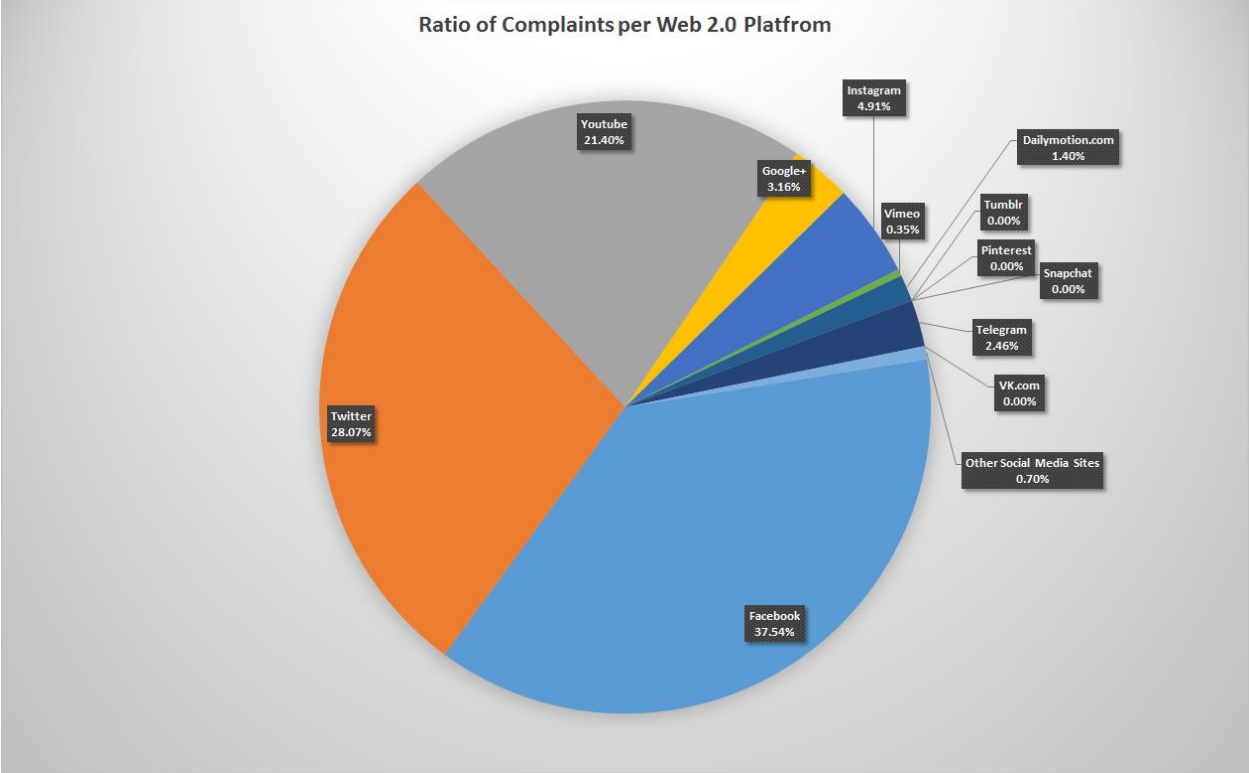
When it comes to platforms where cyber hate is flourishing, it is arguably unsurprising that social media trumps Web 1.0 platforms by a magnitude. However, there are still websites, blogs and forums on the internet that spew hate or provide a platform for people to post hateful messages.



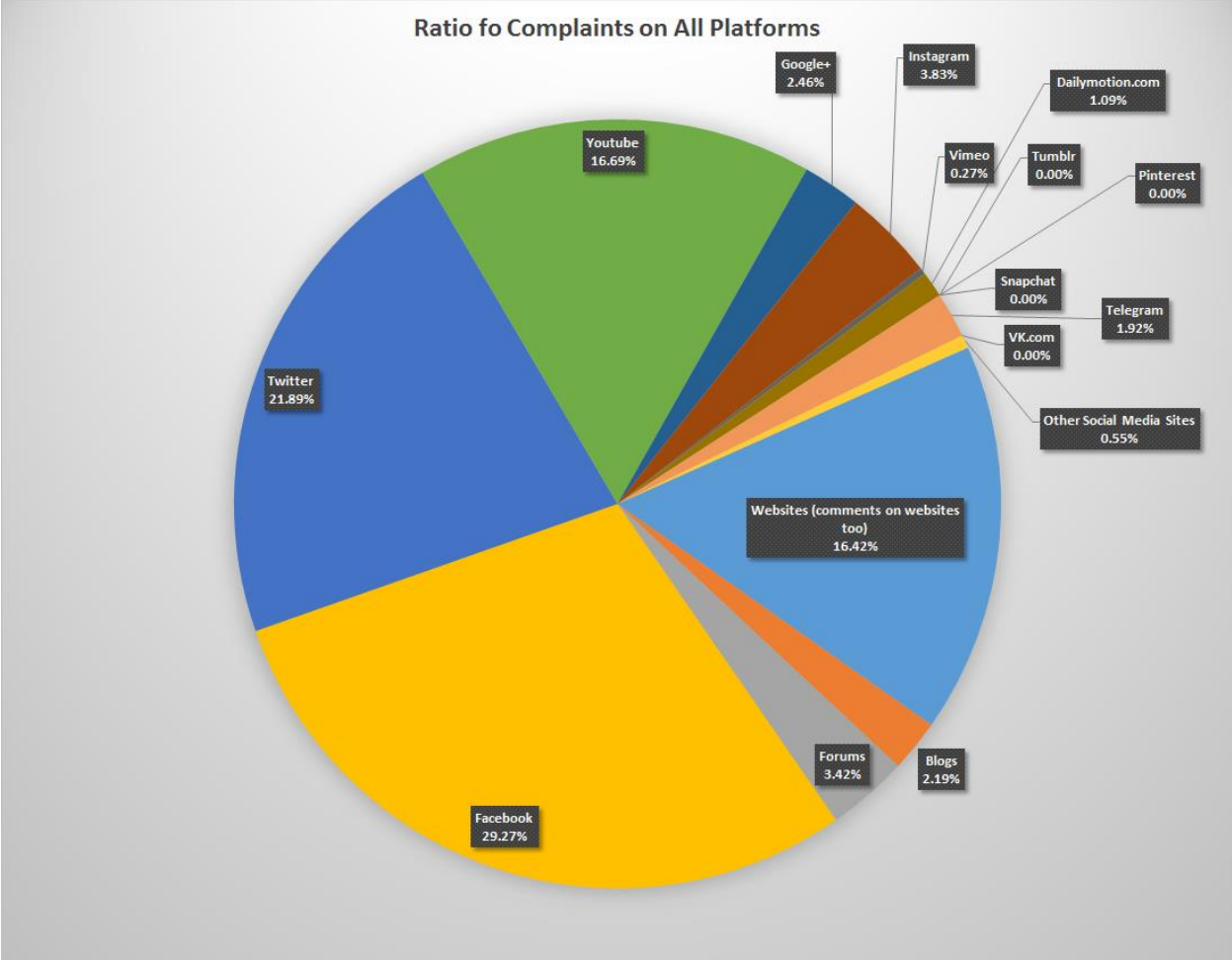
Websites are still the most widely used platforms among Web 1.0 platforms to spread cyber hate, either by producing it or by providing a platform for people to post vile, violent and hateful comments. During the third quarter of 2016, 69.35 per cent of recorded cases of online hate (that appeared outside 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; solidifying the first place of websites among web 1.0

platforms. Cases on blogs fell from 19.01 per cent to 12.9 per cent in the third quarter and fell even further in the fourth to 9.94 per cent, whilst forums kept their second place. After seeing a sharp rise from 8.26 per cent to 17.74 per cent in the third quarter, cases on forums fell to 15.53 per cent, but still keeping their second place among traditional platforms.

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. Although, this 9 per cent fall still has not endangered the platform's first place among social media sites. Twitter kept its second place by rising from 23.25 per cent to 28.7 per cent in the third quarter and then basically staying there at 28.07 per cent. Finally YouTube stayed third with a minor, approximately 3 per cent rise, from 17.89 per cent to 21.40 per cent. Google+ lost its third place to Instagram (that has produced a substantial rise from below 2 per cent to almost 5 per cent) by falling to a minuscule 3.16 per cent from its already meagre 3.22 per cent.



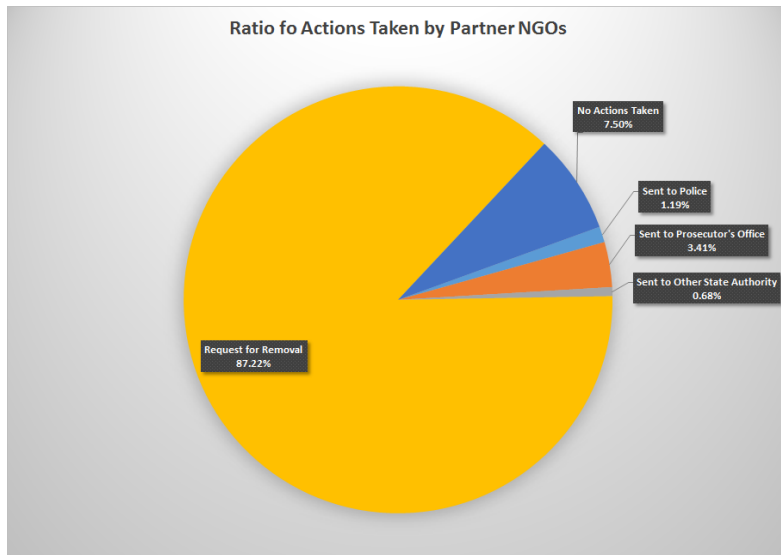
If the numbers of Web 1.0 and 2.0 platforms are combined, it becomes crystal clear that social media sites are still the pivotal platforms when it comes to the spreading of cyber hate. These platforms provide a cheap or even free tool for people and extremist groups to deliver their message to a gigantic 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 16.42 per cent (7.84 per cent in second quarter and 9.54 per cent in the third) and the ratio of cases on blogs and forums falls to around or below 3 per cent (not a major change from the previous data, unlike websites that have observed a major 7 per cent rise).



These numbers clearly show that social media sites have completely taken over and fundamentally changed the landscape of cyber hate by letting their users spewing out 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 Organizations 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 the unquestionable leader with 87 per cent (no change from previous quarter).



This is followed by cases where no actions were taken. The number of these cases has risen between the 3rd and fourth quarters from 3 per cent to 7.50 per cent, a 180 degrees turn in the previous trend (see previous report). 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.

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, followed by Twitter and YouTube. The fourth highest number of complaints in the fourth quarter was received on Instagram, but that number is dwarfed by 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 is signalling a bad trend that did not turn around in the fourth quarter either. In November and December Facebook's removal rates have fallen to an abysmal 50.90 per cent. Twitter is still doing even worse than Facebook, with its removal rate falling from 67.01 per cent to 66.34 per cent in the third quarter and to 24.68 per cent in the fourth. The biggest fall, 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. As we mentioned it in our previous report, that high ratio was fairly unusual for the platform and that statement has been underpinned by YouTube's abysmal numbers in the third quarter. The platform's removal rate has fallen to 62.22 per cent, and then to 52.63 per cent in the final quarter of 2016, sadly solidifying this trend. "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 in our first quarterly report. The grim picture described in that report has become even

grimmer in the second quarter. Google+'s removal rates have 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 has risen to 47.06 per cent in the winter of 2016. Although, there is still obvious room for progress.

Name of Platform	Percentages of Cases Removed	Percentages of Cases Not Removed
Websites (comments on websites too)	52.38%	47.62%
Blogs	50.00%	50.00%
Forums	75.00%	25.00%
Facebook	50.90%	49.10%
Twitter	24.68%	75.32%
Youtube	52.63%	47.37%
Google+	47.06%	52.94%
Instagram	78.57%	21.43%
Vimeo	0.00%	100.00%
Telegram	57.14%	42.86%
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.

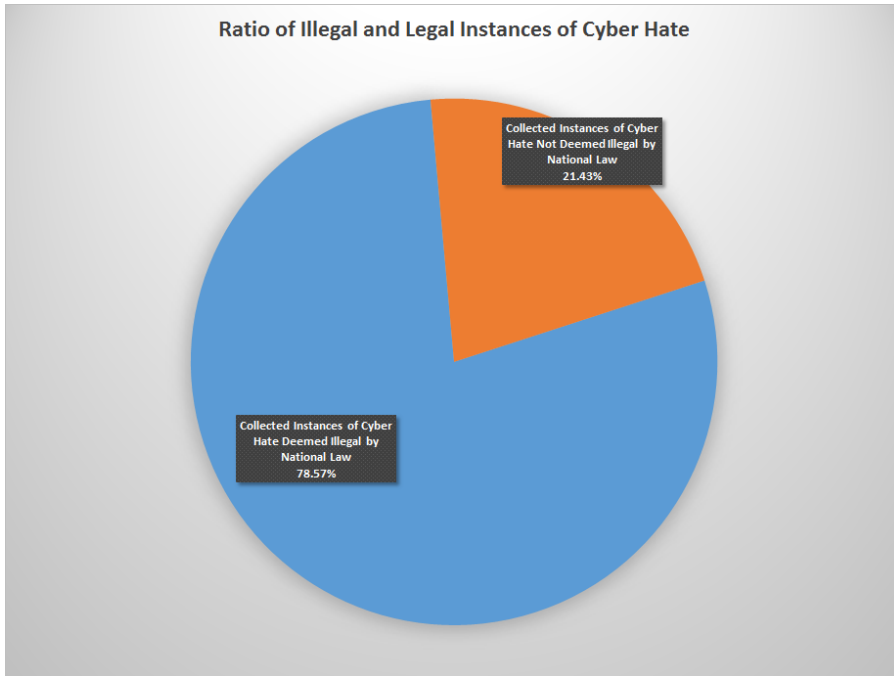
Taking all this into account, it is very aggravating that removal rates vary vastly between countries. For example, jugendschutz.net in Germany had a 82.46 per cent removal rate on Facebook in November, whilst ZARA in Austria only had 21.71 per cent success rate in removals and Licra in France faced similar numbers to ZARA's (21.74 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 December, whilst Licra had an 71.43 per cent success rate, but, for instance, MCI in Spain and ZARA in Austria did not manage to get anything removed from the platform in that month. In the meantime Magenta had a 100 per cent success rate in the same time period.

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, it might not always be considered illegal. As presented in our previous report, 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 means that - in the fourth quarter - 21.43 per cent of cases 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

Reaching extensive conclusions based on numbers collected in the first and second three months of our data collecting period would be still fallacious and premature. Some minor changes and shifts in the data can be already 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 will discuss trends, shifts in the data and the conclusions that can be drawn from them in our extensive and comprehensive yearly report that we will publish in late 2017.