



**Quarterly Report on
Cyber Hate
(January, February
and March 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	-	-	-	-	-	-	-	-	-	5.
3. Tools	-	-	-	-	-	-	-	-	-	6.
III. Data Collection and Analysis	-	-	-	-	-	-	-	-	-	7.
1. Hate Type Analysis-	-	-	-	-	-	-	-	-	-	7.
2. The Prevalence of Cyber Hate on Different Platforms	-	-	-	-	-	-	-	-	-	10.
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 found that in January, in France, the context of the presidential elections (of April and May 2017) had some effects. French candidates from all the political parties were increasingly present in the media sphere. There was a polarized political debate on “far-right” ideas such as immigration, integration, “laïcité” and Islam. This debate reinforced the extremist political statements and the radical political “buzz” with a lot of example on Twitter. Indeed, there were a lot of pro-National Front or “pro-Marine” (Marine Le Pen was the Front national candidate) Twitter and Facebook accounts. Moreover, the “Théo affair” also had an impact. Théo, a 22-year-old black youth worker, claimed that the police sodomised him with a baton during a violent confrontation in Aulnay-sous-Bois, a Parisian suburb. He also stated that the four policemen who raped and beat him used racial epithets (“niger”, “bamboula”). The “Theo affair” has revived past controversies over police brutality in France with, for example, the case of Adama Traoré, a 24-year-old black man who died last July after he was arrested by the police. Furthermore, more cases emerged with regards to the same context. Indeed, young Black or Arab people from the suburbs were victims of violent identity checks made by policemen which led to questioning the “contrôle au faciès” (racial, ethnic profiling). In 2016, the French Court of Cassation condemned the French state for racial and ethnic profiling related to identity checks carried out by the police. Another consequence of the affair was the number of riots and demonstrations taking place in some Parisian suburbs, especially in the area of the “Seine-Saint-Denis”. The situation was especially explosive due to the upcoming presidential elections. In February, the presidential campaign was marked by

corruption and fraud scandals for François Fillon (leader of the Republicans – Right conservative party). He was accused of fraud due to employing his wife, Penelope, and their children in “fake” parliamentary jobs. Another scandal was the one regarding Marine Le Pen (National Front), who was summoned by judges over a fake EU jobs scandal. She was accused of creating jobs in the European Parliament (EP) for members of her staff. She refused to meet with the judge until after the election. She was also stripped of her immunity by the EP, allowing French prosecutors to investigate her over offensive tweets sent in December 2015. Both Mr. Fillon and Ms. Le Pen have attacked judicial investigations into their use of parliamentary aides as an attempt by outgoing President François Hollande and his Socialist Party to influence the vote. Mr Fillon has described an inquiry launched in January as an "institutional coup d'état" and has accused journalists of trying to carry out a "lynching" and an "assassination", remarks compared to Donald Trump's frequent attacks of the media. On other matters, the Mehdi Mehklat affair had impacts as well. Mehdi Meklat and Badrounine Saïd Abdallah — “Mehdi” and “Badrou,” — were stars of the Bondy Blog, a much-praised chronicle of suburban France sponsored by the mainstream press. However, presently, Mr. Meklat has been revealed as the semi-hidden author, under a pseudonym, of hateful and obscene tweets — antisemitic, misogynistic, pro-jihadist and homophobic. Mr. Meklat has gone from hero to pariah. The International League Against Racism and antisemitism has turned over Mr. Meklat's tweets to the Paris prosecutor.

MCI in Spain found that in January, limits of freedom of expression regarding hate speech was strongly debated. Since article 510 of Penal Code was made more concrete there were many court cases regarding the issue. The radical left was also frequently targeted, for the alleged glorification of terrorism. Lastly, anti-Muslim hate was still predominant. In February, a new driver appeared, the so called “transphobic bus” which was a marketing campaign promoted by an ultra-conservative catholic organisation www.hazteoir.org. This bus travelled all around the country showing the “truth” to the “poor Spanish children manipulated by the gay lobby”. The organisation was later charged for transphobic hate speech. After the prosecution Hazteoir started a campaign for freedom of speech.

UNIA, in Belgium, found that in January the hateful comments subsequent to the death of a young Belgian citizen of Turkish origin during the attack in a nightclub in Istanbul led to relatively tumultuous online discussions. UNIA itself was contacted by citizens, who were angry and disappointed that the institution defended the “muslim community”. In February, UNIA appeared in the media after an internal email evaluating a judgement had been transferred by mistake to a stakeholder. The latter published the email on its website and the Flemish nationalist party (NVA) took the subject over and criticised the institution openly. Two complaints received concerned UNIA's “one way” analysis. Other similar requests could appear in the following months as not all complaint files have been closed yet. In March, UNIA received several complaints concerning the activities of the Turkish AKP campaigning in Belgium and neighbouring countries. Amongst other things, it was about a poster where the Muslim crescent hit the Christian cross. The origins

of the poster was not identified at the time. Furthermore, the construction or plan to build mosques have led to cyberhate inciting people to place “pig heads” on the construction grounds.

2.Trends

Regarding the new **trends**, jugendschutz.net (Germany) found that in January, one year after the events in Cologne on New Year's Eve 2015, when women were sexually assaulted allegedly by people of North African descent, there was a big debate in the media about the role of the police during the incident. In order to prevent repeated attacks on New Year's Eve 2016, they controlled and detained mostly "North African looking" men and dubbed them "North African multiple offenders", which partially evoked criticism from politicians and civil society. The event and the following discussion was picked up and used by far-right activists to create an atmosphere of hatred against migrants and politicians. In February, in the course of the Memorial Day for the victims of the Dresden bombings in 1945, many far-right extremist groups took part in demonstrations. An installation of three buses positioned upright in front of Dresden's Frauenkirche, which was meant to remind people of the war in Syria and to send out a message against war as a piece of art, received great response. Subsequently, there were physical protests from PEGIDA members (Patriotic Europeans against Islamization of the Western World) and from almost all far-right groups on the Internet against the "scrap pile" and the wrong culture of memory. The Identitarian Movement placed a banner with the slogan "Hypocrites. Your politics is scrap. No interventions, no arms delivery → no migration". Pictures of the action were shared thousands of times in social media.

LICRA in France denoted that in January, anti-black racism related to the Théo affair on Twitter and Facebook became prevalent. In these instances of racist cyber hate, black people (including the former Minister of Justice, Christiane Taubira) were compared to monkeys. There was also anti-Arab and anti-Black racism present related to the “denunciation of the suburb scum” (this idea is one of the ideological point of the National Front). In February, the far-right and right-wing parties promoted the conspiracy theory that “Islamism” was present in the French justice system and in the French government.

According to ZARA in Austria in January, the Austrian Integration Minister put forth a proposal at the beginning of the year to ban civil servants from wearing the Hijab and to ban full-face veils (burka) in public places. At the end of the month, the government (a coalition of the Social Democrats and the Conservatives Party) announced the plan of law amendments and also a new “integration law”, including the ban of wearing the burka in public and an order for certain civil servants to wear “neutral clothing”(addition to the already existing official uniforms for these professions), meaning that judges, prosecutors and police officers would not be allowed to wear a headscarf. At the same time, members of the government and political parties – e.g. the Justice

Minister and the Integration Minister – emphasized several times that the neutrality law does not concern the crosses/crucifixes in Austrian court rooms and other public rooms and that they will stay. These populist measures - adding to the existing anti-muslim climate - are seen (by some) as a confirmation that “Islam does not belong to Austria”. In March, after the attack in London, a picture circulated on far-right social media sites (also Austrian and German ones), which showed a woman wearing a hijab and looking at her phone on Westminster Bridge as people gathered around an injured person nearby. The picture was posted with texts like “Muslim woman pays no mind to the terror attack, casually walks by a dying man while checking her phone” and incited a vast number of appalling hate comments against the woman and Muslims in general. Lastly, articles and reports about the referendum in Turkey led to hate speech against people of Turkish origin (living in Austria and Germany) in general.

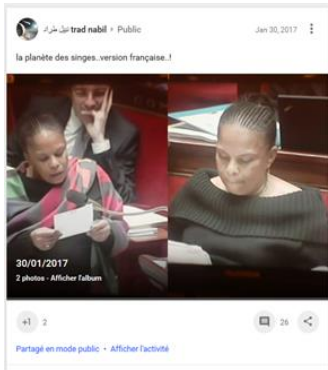
3. Tools

Regarding the new **tools**, LICRA noted the following photos; this Facebook post (which Facebook refused to delete), with the following text; “We have some news from Théo, enormous inflamed haemorrhoids crisis, it is an epidemic in all the closed suburbs, need to put them in quarantine illico presto!! Lol”.

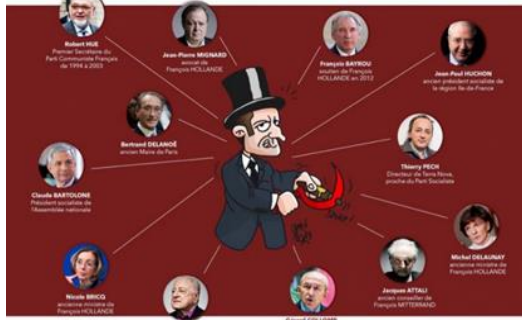


haemorrhoids crisis, it is an epidemic in all the closed suburbs, need to put them in quarantine illico presto!! Lol”.

Moreover, this Google+ post shows two photos of Christiane Taubira when she was the French Ministry of Justice with this caption “Planet of the Apes, French version...!”



In February, Emmanuel Macron, one of the presidential candidates (now President), was caricatured in a picture tweeted by the political party of the Republicans (right-wing party). This caricature reminded the public of the antisemitic pictures of the 30s.



Coming from the “fachosphere” vocabulary, the word “islamo-gauchisme” also became more mainstream. The translation could be “Islamism-leftism”. The idea behind this “trendy” online and offline insult is to create links between Islamic movements and people on the left or far-left end of the political spectrum. Lastly, this image was used as a tool for cyber hate; “If Fillon were

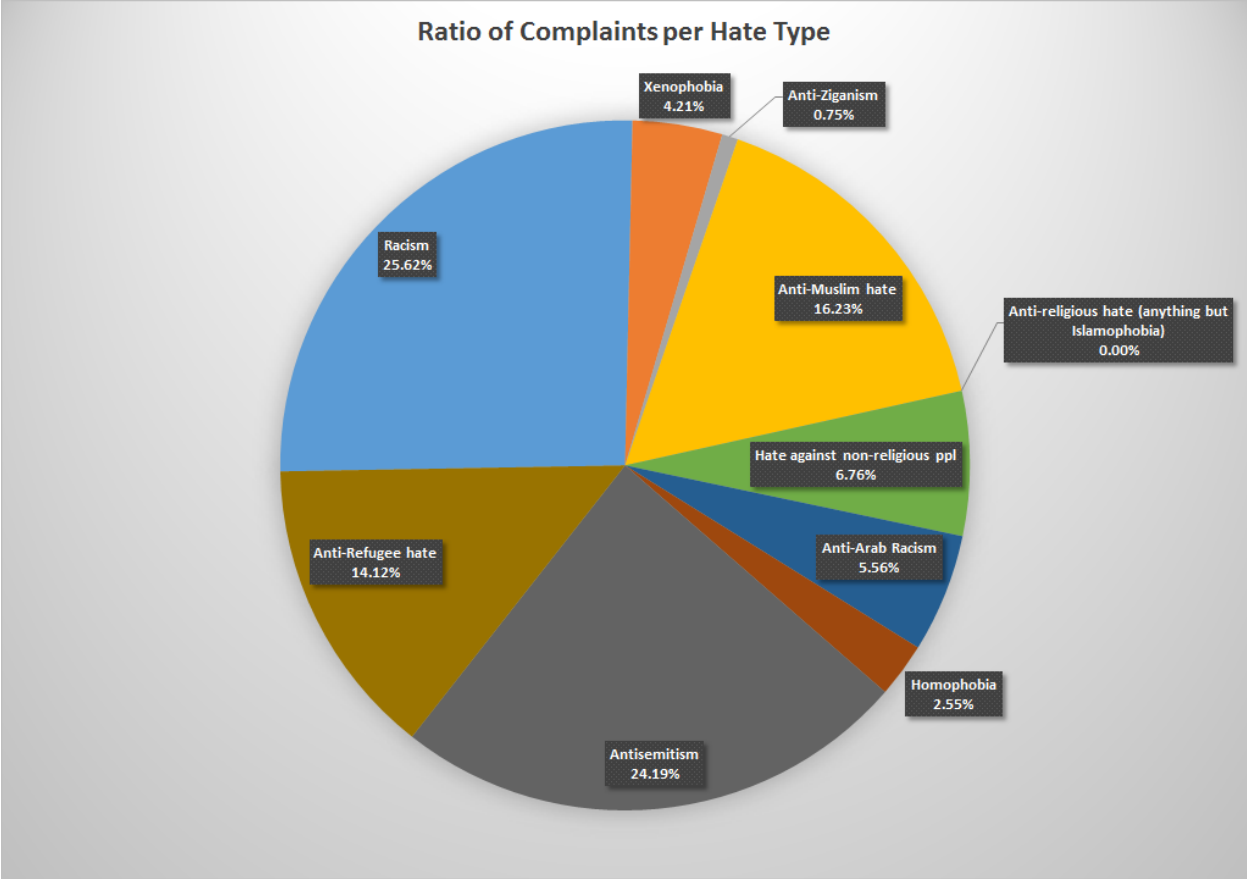


politically murdered by the red Islamo-leftist judges guess who will replace him. These mayors who promote Islam: Alain Juppé very helpful with the Muslim Brothers” (François Fillon is the Republican candidate and Alain Juppé is member of the same political party).

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 show that the trend of falling numbers in anti-Muslim hate continued in 2017. 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 quarter of 2017, with the numbers falling further to 16.23 per cent. Thus, the ratio of anti-Muslim instances of cyber hate collected by INACH fell even further below one-fifth of all cases, making it the third most prevalent hate type.

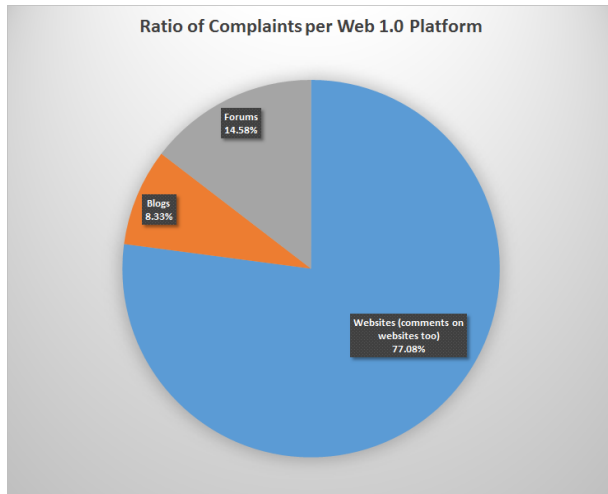
General racism took the first place from anti-Muslim hate in the 3rd quarter of 2016, by rising to 20.71 per cent from 20.02 per cent (due to our methodology, antisemitism, anti-Arab racism, anti-Ziganism and anti-refugee hate are all excluded from these numbers). This trend also continued as the hate 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 continues in 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. are now responsible for more than a quarter of the cyber hate cases handled by INACH and its partners. 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. This means that racist and antisemitic cases of online hate speech are responsible for half 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, 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. However, this trend has gone through a complete turn, where the ratio of this hate type among all monitored hate types has fallen to 6.76 per cent, an almost 100 per cent fall. The ratios of all other hate types are either just above or 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. 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, virtually raising the hate type's ratio back to virtually its highest level.

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.



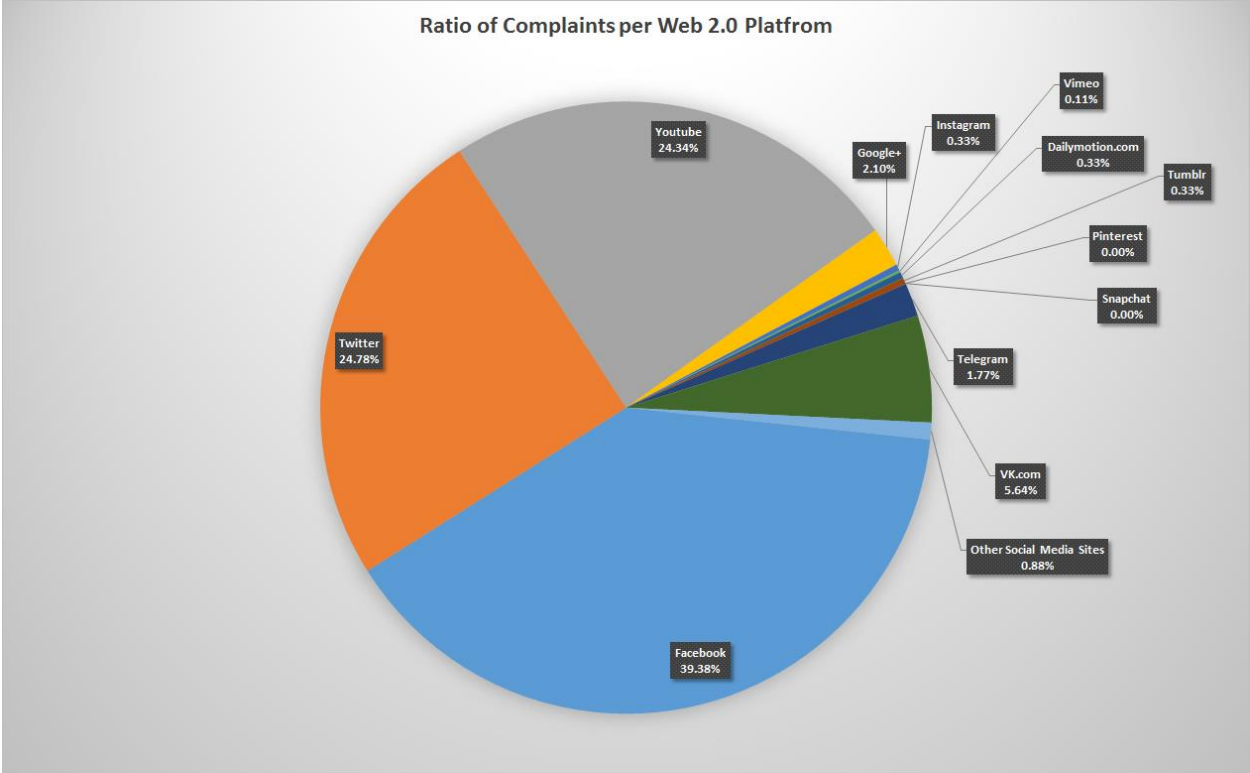
Websites are still the most widely used platforms among Web 1.0 platforms to spread cyber hate, either by producing it or by providing a platform for people to post vile, violent and hateful comments. During the third quarter of 2016, 69.35 per cent of recorded cases of online hate (that appeared outside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

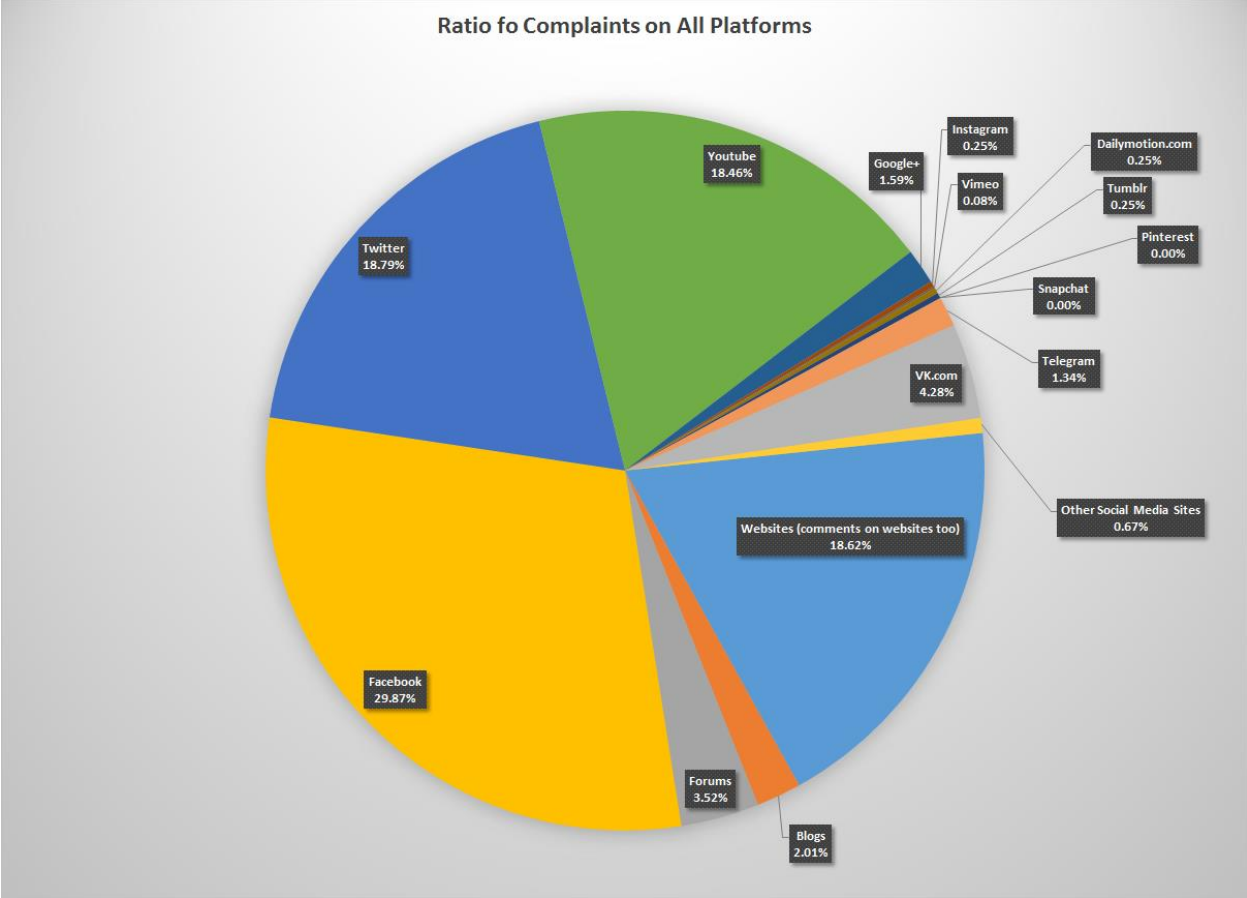
Cases on blogs fell from 19.01 per cent to 12.9 per cent in the third quarter of 2016 and fell even further in the fourth to 9.94 per cent, and then fell further to 8.33 per cent in the first three months of 2017, solidifying another trend in the data. Forums kept their second place. After seeing a sharp 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, 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 in 2016 and this ratio has hardly changed in the first quarter of 2017 by rising to 39.38 per cent, a negligible change. However, nothing can threaten the first place of Facebook among the social media platform, not even the two other major member of the dominating triumvirate. Twitter kept its second place in 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. Finally YouTube stayed third with 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 rising further to 24.34 per cent in the beginning of 2017. All other platforms

are completely dwarfed by the three major players. All of them are around or under 2 per cent. This means that both Google+ and Instagram have fallen quite a lot, making the hegemony of Facebook, Twitter and YouTube even more staggering than it used to be.



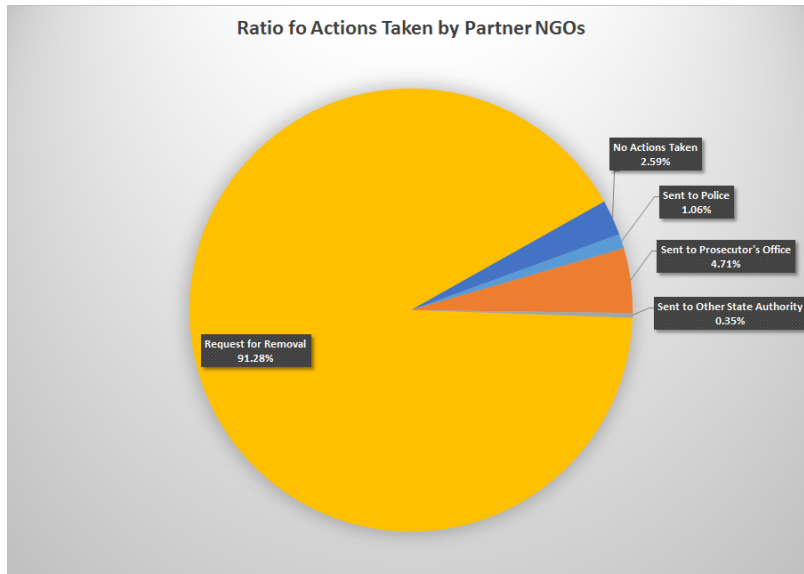
If the numbers of Web 1.0 and 2.0 platforms are merged, it becomes crystal clear that social media sites are still the pivotal platforms when it comes to the spreading of cyber hate. These platforms provide a cheap or even free tool for people and extremist groups to deliver their message to a gigantic 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 18.62 per cent (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 around or below 3 per cent (these ratios have not changed substantially since the last two months of 2016).



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 the unquestionable leader with 91.28 per cent (a minor 4 per cent change since the previous 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.

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 has not changed since and the ratio of these cases still hovers around 6 per cent in the first three months of 2017.

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. 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 is signaling 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 the first quarter of 2017, even though Facebook's removal rates have risen to 58.66 per cent. This removal rate is still atrocious, however, its upward trend does give some hope that Facebook will reach the high 70s again.

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 2017 and Twitter’s removal rates have risen to a whopping 74.6 per cent. Hopefully this will not change during the second quarter of the year.

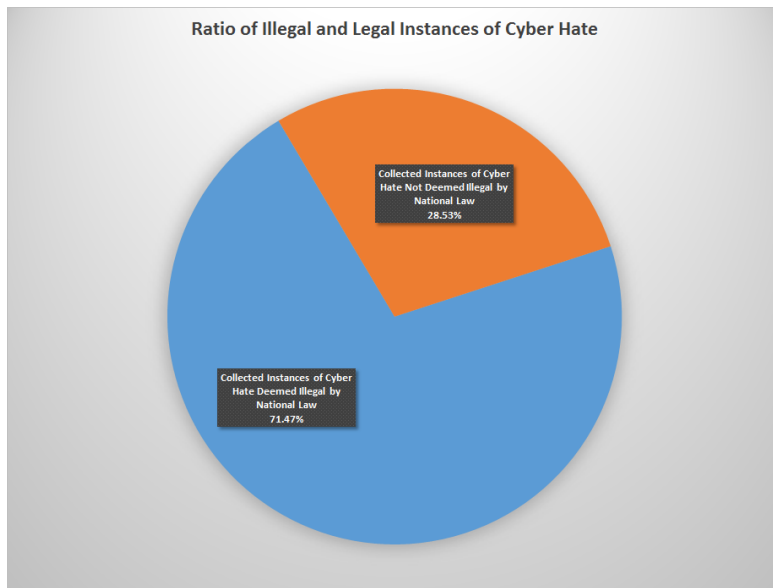
The biggest fall in 2016, however, was seen by YouTube. The platform’s removal rate was very close to Facebook’s with an exceptionally high 86.27 per cent during the second quarter of 2016. As we mentioned it in our previous report, that high ratio was fairly unusual for the platform and that statement has been underpinned by YouTube’s abysmal numbers in the third quarter of 2016. 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. However, adding to the aforementioned positive changes within trends observed in the beginning of 2017, YouTube’s removal rates have risen to a staggering 83.33 per cent. Another welcome change that will hopefully become a trend during this year.

“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 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. Sadly, this positive development did not become a tendency. Google+’s removal rates have 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.

Name of Platform	Percentages of Cases Removed	Percentages of Cases Not Removed
Websites (comments on websites too)	38.03%	61.97%
Blogs	37.50%	62.50%
Forums	72.73%	27.27%
Facebook	58.66%	41.34%

Twitter	74.60%	25.40%
YouTube	83.33%	16.67%
Google+	5.26%	94.74%
Instagram	0.00%	100.00%
Vimeo	0.00%	100.00%
Dailymotion.com	66.67%	33.33%
VK.com	48.00%	52.00%
Tumblr	100.00%	0.00%
Telegram	43.75%	56.25%
Other Social Media Sites	25.00%	75.00%

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 59.38 per cent removal rate on Facebook in January, whilst ZARA in Austria only had 29.41 per cent success rate in removals and LICRA in France produced similar numbers to ZARA’s (37.14 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 March, whilst LICRA had an 47.62 per cent success rate, but, for instance, MCI in Spain had 100 per cent removal success (although, based on only 2 cases), and finally ZARA in Austria did not manage to get anything 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.

This means that - in the first quarter of 2017 - 28.53 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

Reaching extensive conclusions based on numbers collected so far in 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 be published in late 2017.