Quarterly Report on Cyber Hate (July-September 2017)

By Tamás Berecz and Charlotte Devinat

Project Research - Report - Remove: Countering Cyber Hate Phenomena

INACH

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

Table of content

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 observed that during July and August, small social events took place; A "France insoumise" deputy (from the far left party), Danièle Obono, born in Gabon, was criticized after statements she made about the support of the author of a song titled "Nique la France" ("Fuck France"). She then became the target of the National Front and the fachosphère. She was also compared to a monkey on the social networks. Moreover, a racist comment on a picture was posted on Twitter against the famous actor Omar Sy and his wife Hélène Sy. The picture was a photo of the couple with this caption: "The racial mixture is an abomination regarding the Divine creation". Hélène Sy directly asked Twitter to react.

The murder of Sarah Halimi was also a driver. During the night of the 3rd of April, a neighbour of Sarah Halimi, 27-year-old Kobili Traoré was accused of having broken into her flat. Traoré

allegedly beat her to death and hurled her body off the balcony into the courtyard below. In the days that followed, French authorities classified Halimi's killing as an isolated incident. However, Jewish leaders immediately protested, especially after other neighbors testified that they heard Traoré scream "Allahu akbar," while allegedly attacking Halimi. In September, a psychological evaluation of Traoré was released. It said that the suspect was under the influence of illegal substances at the time, which was a contributing factor to the crime, but that the act was nonetheless clearly antisemitic. Ten days after, the Paris public prosecutor recognized the murder of Sarah Halimi as an antisemitic hate crime. The prosecutor's decision was announced based on interviews conducted by psychiatrist Dr. Daniel Zagury.

Furthermore, the #Couscous Gate was also considered as a driver. A picture was published online of Florian Philippot, Marine Le Pen's deputy, dining at a North African restaurant with other Front National officials (eating couscous) in Strasbourg which horrified FN supporters who found their dinner "unpatriotic". One week later, Florian Philippot, the FN's vice-president in charge of communication and strategy since 2012, has announced to the French TV that he was leaving the FN.

Another example of drivers is the one of the 10th of September, where a French Jewish family was assaulted in their home near Paris, at Livry-Gargan. In the attack, three men, two of whom were wearing masks, broke into the home of Roger Pinto, the president of Siona, a group that represents Sephardic Jews. One of the attackers said: "You Jews have money," according to the family members.

Other than that, the administrative court of Nice rendered a decision on the 4th of September, by which it condemned the Prefect of Alpes-Maritimes for infringement of the right to asylum, it being the second time in less than 6 months. The case regarded the arrest of three Sundaneses in Nice during the formalisation of their asylum request. The three men were hosted in Cédric Herrou's house, who is the main representative of "Royal Citoyenne", an NGO specialized in human rights residing near the French Italian border. Cédric Herrou has been sentenced to a four months suspended sentence for helping 200 migrants crossing the border. He is the French symbol of the "offense of solidarity" (délit de solidarité).

Further, The 15th of September, the publishing house Nathan, specialized in textbooks, has been at the center of polemics because of the publication of a mathematics textbook where a mathematics problem illustrated migrants on a boat (the problem used "migrants from war" going to a Mediterranean island as ecamples, "the 1st week, they are 100, and every week the number of migrants increases of a 10% (…) How many arrived (…)"). Nathan Editions has apologised and promised to publish a new mathematics book quite soon.

Yet another example is the one of France's parliament who approved a new counter-terrorism bill, putting several controversial measures in place under a nearly two-year-long state of emergency. It will allow the authorities to confine suspected jihadist sympathisers to their neighbourhoods, close places of worship accused of condoning terror and carry out more on-the-spot identity checks, all without the prior approval of a judge. The legislation has encountered little resistance from a public traumatised by jihadist attacks, despite criticism that it will undermine civil liberties. It is expected to become law before the state of emergency, declared after the 2015 Paris attacks, elapses on November 1 after being extended six times.

France's gender equality State Secretary has also set up plans for a new law countering sexual violence and harassment. Marlène Schiappa's proposals include on-the-spot fines for catcalling and lecherous behaviour in public. A taskforce of politicians will work with police and magistrates to establish what sort of behaviour constitutes sexual harassment. The escalating scandal about

Harvey Weinstein has prompted women and men in France to share their own experiences of sexual harassment by using #balancetonporc - "rat on your dirty old man (pig)" - to encourage women to name and shame those responsible.

Finally, eight people suspected of belonging to a French ultra-right extremist group have been given preliminary terror charges for allegedly plotting attacks against mosques, politicians and migrants. The suspects were charged with criminal terrorist association. Three of the eight charged were minors and seven were in custody. The suspects were allegedly linked to the founder of a newly formed ultra-nationalist group identified as Logan N., who was arrested in southern France in June. The official says that investigators believed that the group was planning to extort money from business leaders and use it to buy weapons. Some of its members had trained in shooting and stolen a vehicle for the alleged plot.

UNIA received two cases of hate-speech in September related to press articles concerning abortion. Those articles lead to discussions among the supporters and the opponents of abortion. The discussion had a religious connotation.

MCI found out that, in July, the Catalonian process for independence caused a dangerous polarisation within the Catalonian society. Expected collateral damages raised strong forms of hate speech. A discriminatory word was used repeatedly for nationalistic Catalonians: "charnego". Charnego stands for how "pure", "rich" and "cultured" "well educated" people refer to Spanish migrants from southern spain (mainly Andalucia and Extremadura. The "deep south") after a harsh debate in the Catalonian parliament. The following facebook post denotes the tensions of the situation; "I know I will have a lot of critics from different sides. I know that what I am about to say is "machinist"…. But just listening to Arrimadas in the debate in T5 (private national television) my desire is that she should be raped collectively tonight. She is a disgusting bitch". Arrimadas is the opposition leader in the Catalonian parliament. She stands for Spanish constitution and against the illegal referendum.

2.Trends

Regarding the new **trends,** Jugendschutz.net denoted that, in July, the adoption of same-sex marriage in Germany at the end of June continued to trigger homophobic hate speech. In August, in the wake of the electoral campaigns in Germany hate speech against politicians increased online and offline. In September, Islamists exploited the prosecution of the Muslim minority Rohingya in Myanmar for their propaganda. They used shocking images and videos to reach a wide audience and to incite hatred against Buddhists. Right-wing extremist groups also tried to influence the parliamentary election with concerted online campaigns using memes to defame politicians. Moreover, they used hashtags to influence public debates and employed social bots and fake accounts to boost their reach.

LICRA listed a few trends in the third quarter. In July the trends were the following: anti-black racism, anti-refugee hate speech (for example, videos with the far-right music band "Les Brigandes"), anti-Roma and anti-travellers hate speech, New popular movement of identitarian/white supremacist rap music and neo-nazi accounts on Google+. In August, there were many conspiracy theories that emerged. Thez are the following: Jews organized the arrival of

millions of migrants / Catalonia is "connected to the International Jewry", Muslims were associated with terrorism and Migrants with thieves, terrorists. In September, a YouTube video of Vincent Reynouard led to the spreading of Holocaust denial. There was also anti-Roma racism making them responsible for making France dirty, there was catcalling linked to people from Maghreb, refugees and sometimes Muslims, and homophobia emerged due to the idea that HIV is an illness only for gays.

3. Tools

Regarding the new **tools**, Jugendschutz.net found that in August there was a new edition of Al Qaeda's magazine "Inspire" featured a guide on how to derail trains.

LICRA found that in July, a new identitarian and White supremacist rap music movement emerged with three leaders "Kroc Blanc", "Amalec" and "Mc Amor" that were very close to the fachosphère, including Dieudonnée and Alain Soral. Their last song was a remix of an old song titled "At the time of the colonies": Kroc Blanc was filmed singing this song just behind the refugee camp of La Chapelle in Paris. There was also a picture that was promoted on Twitter by the ex-leader of the National Front's youth committee about the link between refugees and the terrorist attack against soldiers in Levallois-Perret near Paris. The caption reads "It is a foolish to let thousands of young sexually frustrated men enter Europe, without diploma or abilities and from another   culture".

In August, LICRA found some examples of anti-migrant hate speech: "One day, a migrant will drive over us with this kind of bus… That is the reason not to welcome them". It has 37 retweets. This account "JeuneAthéna" (Young Athena) is quite popular in the online fachosphère (21K followers). The description of the profile: "European and pagan – Anti Great Replacement multiculturalism and equal-feminist – I do amalgams"

Jeune Athéna
@JeuneAthena

Suivre

Un jour, un migrant nous roulera dessus avec ce genre de bus...
Voilà la raison de ne PAS accueillir.
ouest-france.fr/bretagne/douar...

Douarnenez. « Et vous, quelle est votre raison d'accueillir un réfugié ? »

00:00 - 23 août 2017

In September, after the attack in Marseille, an officer close to the Mayor of La Courneuve (a suburb of Paris), Sonia Nour, posted the following on her Facebook account: "When a martyr strangled a woman and stabbed another it makes a lot of noise. Terrorism, blood, civilization, blah, blah, blah, etc. However when patriarchal terrorism kills us [women] every day we don't hear any noise". This message may be considered illegal under French law on the apology of terrorism.

Sonia Nour
3 h ·

Quand un martyr égorge une femme et poignarde une autre là ça fait du bruit. Terrorisme, du sang, civilisation Bla Bla Bla.. Par contre que le terrorisme patriarcal nous tue tous les 2 jours on l'entend moins votre grande gueule.

43   10 commentaires · 3 partages

J'aime   Commenter   Partager

After this Facebook post, a lot of Internet users spread anti-black racist comments against Sonia Nour. Some users created fake accounts of Sonia Nour on Twitter spreading "anti-white racism" and "anti-colonialist"comments. The screenshot below is of a fake account.

There was also anti-black racism. Even if this post is not considered racist per se, it underlines an anti-black trend: It is a picture of the Deputy of Parliament, Danièle Obono with this caption "Yes, It is a real picture from the French National Parliament".
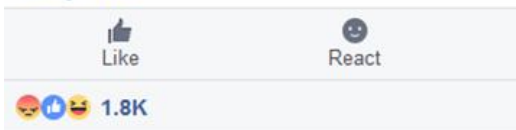

Anti-Roma racism also emerged when Stephane Ravier, the Mayor of the 7th arrondissement of Marseille, from the Front National, published a post on Facebook promoting stereotypes about Roma people. Extract: "On Saturday night I was […] close to "The Jungle of the Arnavons" in the 14th arrondissement of Marseille. A new fire started in the camp where a hundred Roma people have been living for 5 years! FOR 5 YEARS! In a place which used to be a park for families you can see tons of scrap, car and motorbike wrecks and garbage… Children are being used for begging at traffic lights; children under 5 years old are slaloming between the cars at the highway entry and exit […]. The inhabitants are frequently subjected to attacks. […] Rats have invaded the neighborhoods. We cannot talk anymore about hygiene it is more an open landfill site." The post has 1,4K likes + 880 shares + 296 comments and a lot of illegal comments were reported. For example: "We need to make a big chimney fire with all the people inside. It is scandalous to see this I am disgusted. We need to tie up Mr. Gaudin [the Mayor of Marseille] during a night in the midst of these people with all the rats, the dirt and the roaches they used to collect in the garbage and then discharge in their camp, what a scandal".

Lastly, Anti-Arab hate speech was illustrated by the actions of Christophe Arfeuillère, Mayor of a small city in the South-West of France named Ussel, who complained about "always the same gang, known and recognized, that spreads chaos" and added that "Unfortunately the City does not have the power to stop them" on his Facebook page. The post collected 350 likes + 387 shares + 190 comments and also generated a lot of hateful sentiment. For instance: "Regarding the non-willingness of action of the police force we need to organize a "ratonnade" [in French, "ratonnade" refers to a historical term: during the War of Algeria in the 60's, the French police was involved in racist attacks against people from Maghreb: there is no precise figure but hundreds of people from Algeria were killed and some died drowned in the Seine, the river of Paris". Or "The hunting season is open".

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.
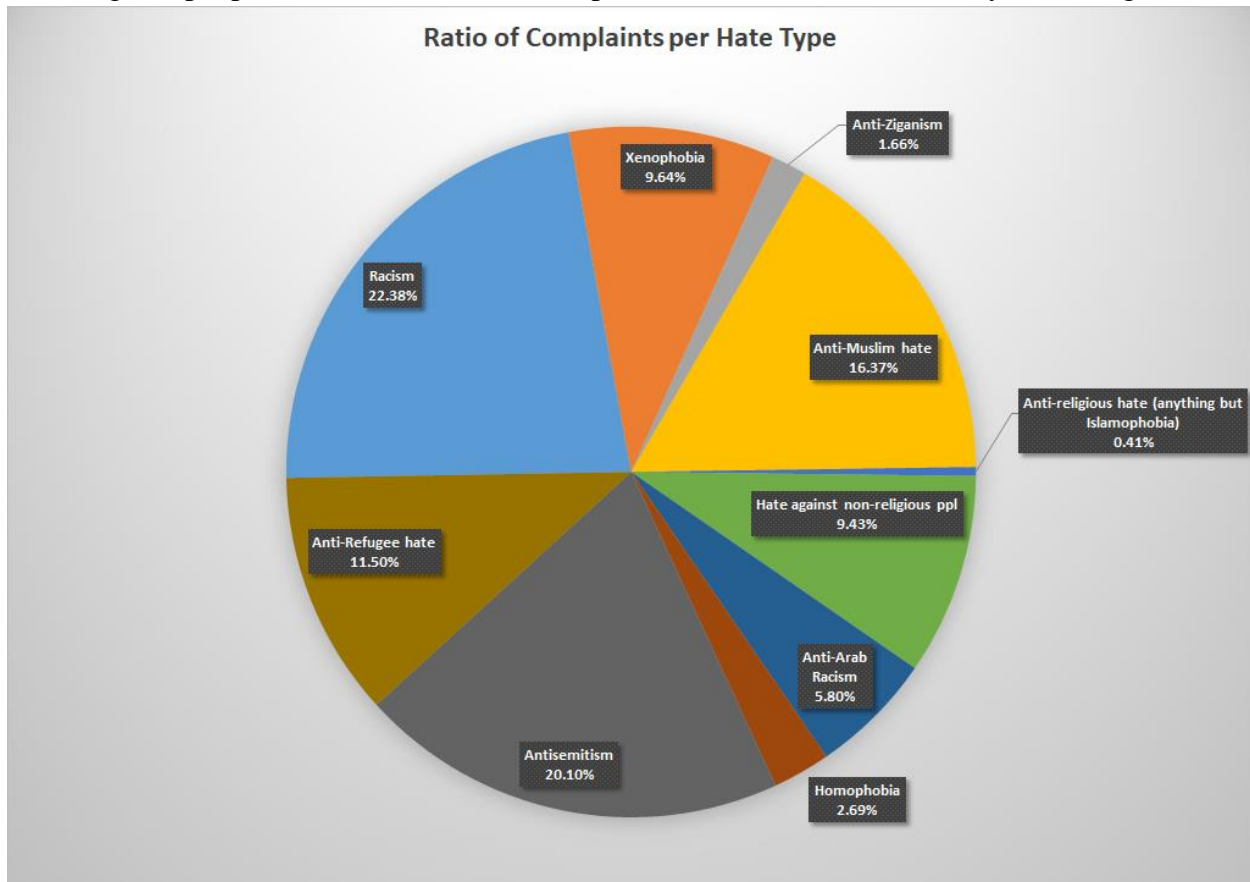
These hate types fluctuate immensely from month to month. Some hate types are very prevalent in some countries, while they are scarce in others. Also, the differences between INACH's project partners adds to this variegation. LICRA, in France, mainly focuses on antisemitism, for instance, therefore they 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. The third quarter did not bring around major change. A minor rise to 16.37 per cent was observed in anti-Muslim hate. Thus, the ratio of anti-Muslim instances of cyber hate collected by INACH did not change fundamentally. 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 turned around during the second quarter of the year and the ratio of racist cases fell to 21.12 per cent, relegating racism to the second position after more than nine months spent in first place. This drop in the second quarter turned into a minor rise in the third, with racist cases reaching 22.38 per cent, retaking their first place on the podium of shame.

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 rose to 24.19 per cent. Unlike the trend in racist cases, this trend did not abate. The ratio of antisemitic cases rose 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 put antisemitic cases into the first place for the first time since INACH started the data collection for this project. However, the first place of antisemitism only lasted for a quarter. By falling to 20.10 per cent, antisemitism once again has become the second most prevalent hate type in our partner countries.

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 fell 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. The minor, 1 per cent, rise to 11.50 per cent in the third quarter did not change anything related to hate directed at refugees.

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 fell to 6.76 per cent during the first quarter, an almost 100 per cent fall, just to rise to 8.99 per cent again. This rise continued into the third quarter with such cases reaching 9.43 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, 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 podium. This major rise, however, has proven to be a only a blip with anti-Arab cases falling back to 5.80 per cent, which could be argued is their normal level based on INACH's data collection so far.

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. However, the third quarter saw one of the largest rises in the ratio of such cases, with xenophobic cases reaching 9.64 per cent. It will be interesting to see if this trend continues into the winter months. The ratios of all other hate types are either just above or way below 2 per cent.

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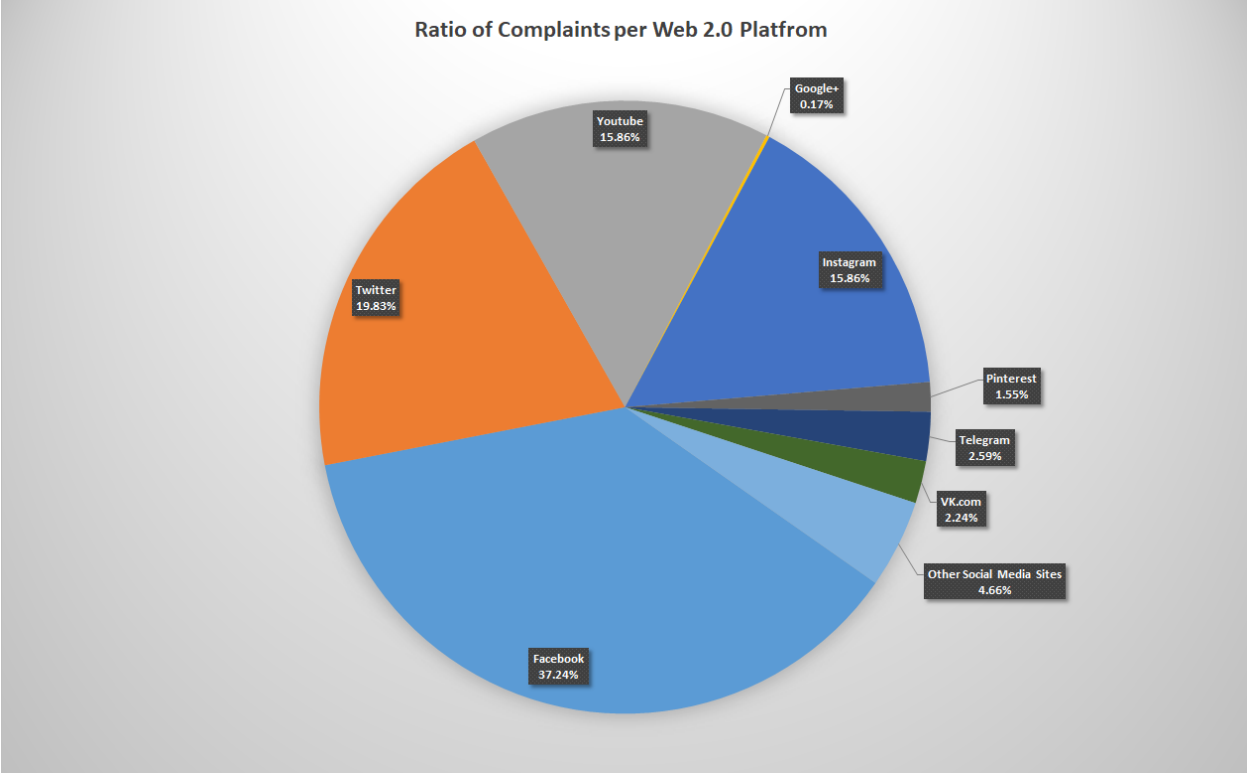
Websites are still the most widely used tools among Web 1.0 platforms to spread cyber hate, either by producing it or by providing a platform for people to post vile, violent and hateful comments. During the third quarter of 2016, 69.35 per cent of recorded cases of online hate (that appeared

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 in the second quarter just to drop even further to 65.03 per cent in the third. Still, this substantial drop has done nothing to endanger the first place of this platform among web 1.0 platforms.

Ratio of Complaints per Web 1.0 Platform

Forums 30.07%

Blogs 4.90%

Websites (comments on websites too) 65.03%
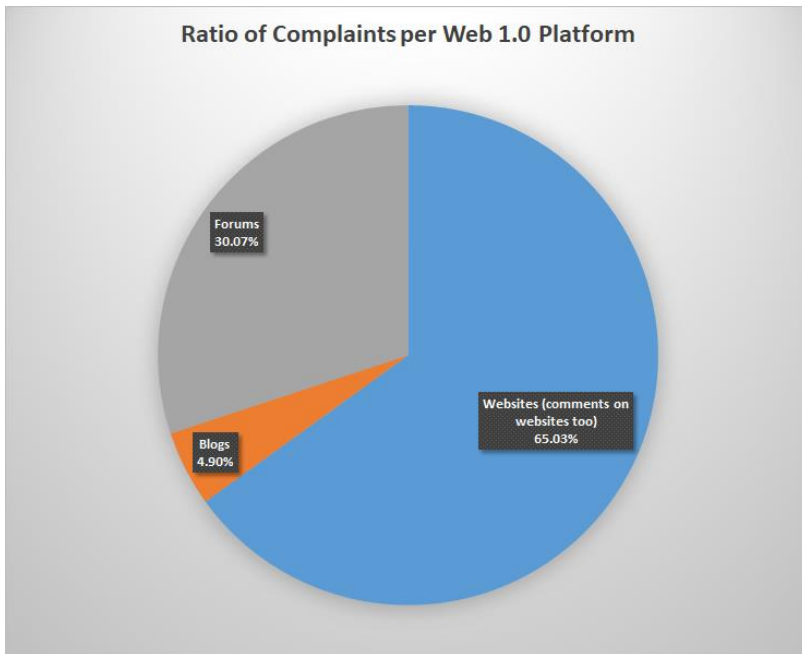
Cases on blogs fell from 19.01 per cent to 12.9 per cent in the third quarter of 2016 and fell even further in the fourth to 9.94 per cent, and then fell further to 8.33 per cent in the first three months of 2017 and to 5.49 per cent during the second quarter, solidifying another trend in the data. This solidification continued into the third quarter with recorded cases on blogs falling further 4.90 per cent. 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 turned around in the second quarter, in which the ratio of registered cases in forums rose to 23.48 per cent. This substantial trend continued into the third quarter with cases on forums rising to 30.07 per cent; not just keeping their second place among traditional platforms, but rising to a previously unseen level.

Ratio of Complaints per Web 2.0 Platfrom

- Google+ 0.17%
- Youtube 15.86%
- Twitter 19.83%
- Instagram 15.86%
- Pinterest 1.55%
- Telegram 2.59%
- VK.com 2.24%
- Other Social Media Sites 4.66%
- Facebook 37.24%

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. In the third quarter of 2017 Facebook observed a minor fall in cases recorded on the platform to 37.24 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 and third quarters, with the social media platform dropping further to 21.95 per cent and then to 19.83 per cent. Finally, YouTube rose to second place for one quarter only to drop back to third in the third quarter. 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. The third quarter, however, saw the largest drop in cases registered on the platform, by YouTube falling to 15.86 per cent. A major, almost 10 per cent shift. The other major change the INACH observed in the third quarter is a gigantic rise in cases registered on Instagram with the platform rising to the level of YouTube at 15.86 per cent. We will see whether this shift is a harbinger of a trend or just an outlier.

All other platforms are completely dwarfed by the three (or in this quarter four) major players. All of them are around or under 2 per cent.

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 21 per cent, which is the highest ratio ever reached by such platforms since the start of our data collection. The ratio of cases on blogs and forums falls to negligible levels, although cases on forums carved out a higher than usual ratio in this quarter by rising to 10 per cent. The major rise in cases on websites and forums means that web 1.0 platforms closed in on the three major social media platforms. So much so that websites took over both Twitter and YouTube, which was previously unseen and we never thought it would happen. It is to be seen whether this is something that stays more permanent or not.
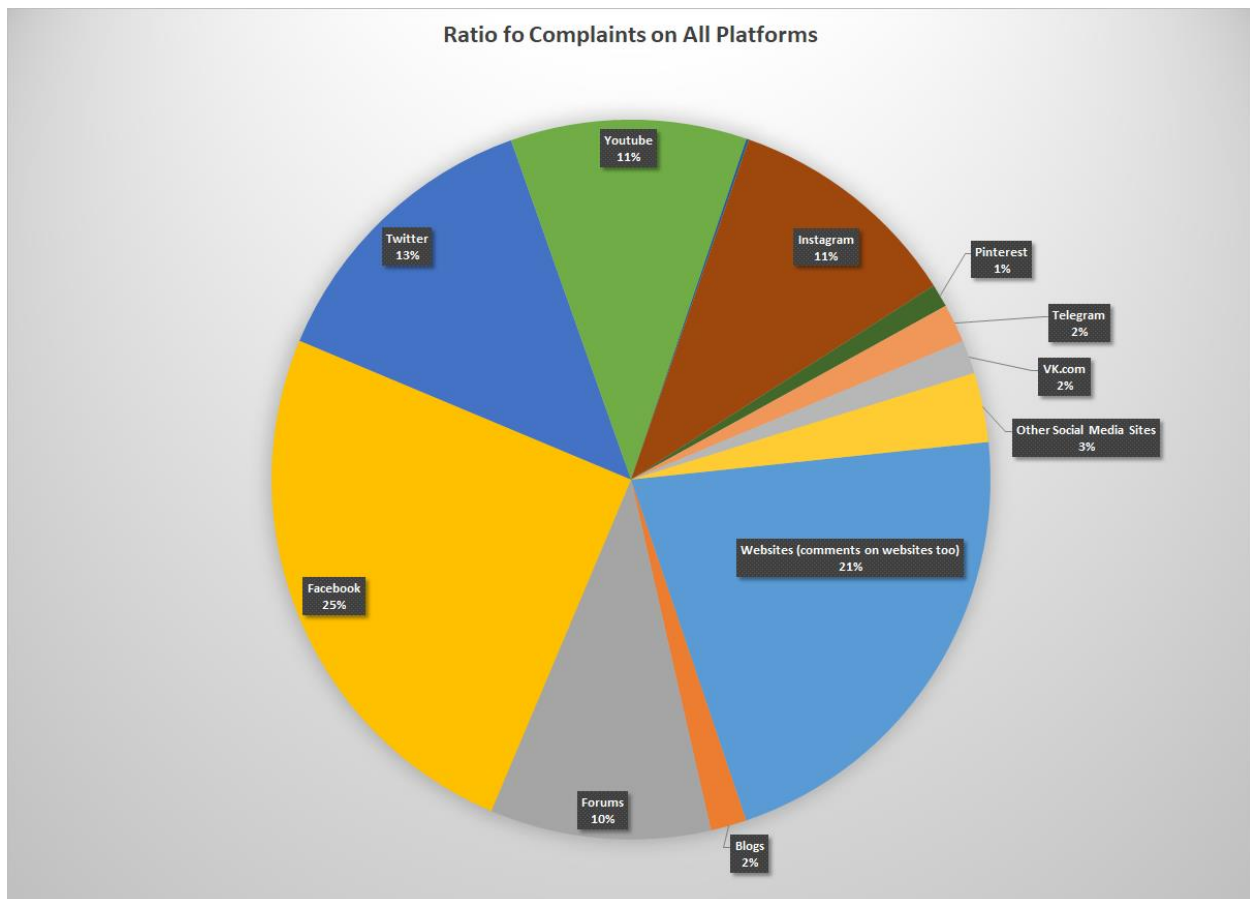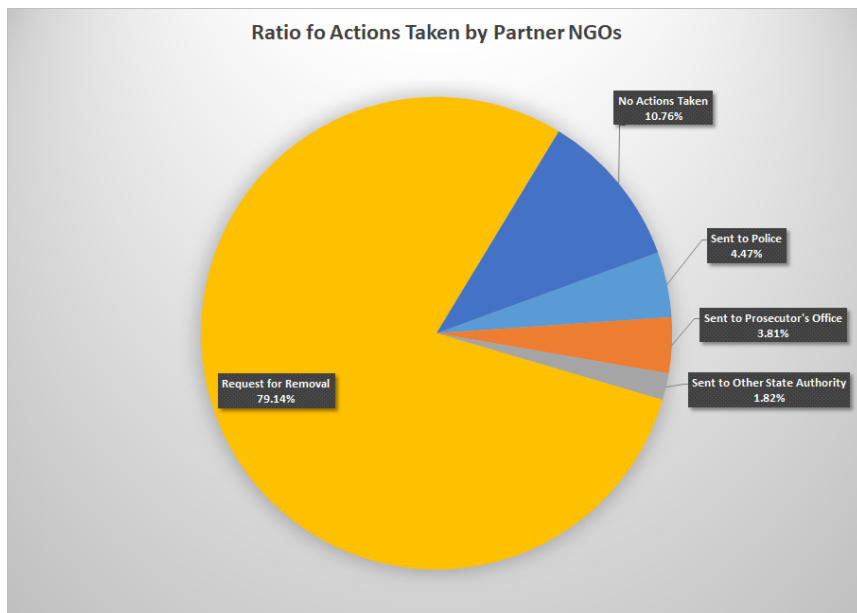

Even with the rise of forums and websites in this quarter, these numbers still clearly show that social media sites have completely taken over and fundamentally changed the landscape of cyber hate by letting their users spew hateful and violent content against minority communities in the form of memes, 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 79.14 per cent, which is a minor drop from 82.92 per cent in the second quarter, that was 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. The third quarter of this year, however, saw a major rise in such cases again to 10.76 per cent. Hopefully another outlier and not a trend starter.

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. Because most cases that are

sent to the prosecutor's office come from Spain due to the legal environment there, it is not surprising that their numbers fell to 3.81 per cent, since we did not receive data from Spain in August. This also meant that the ratio of cases sent to state authorities had a minor fall to 10 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, 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. It is positive news that the platform's removal rate bounced back to 62.32 per cent in the third quarter, a major rise, but still far away from their highest levels, therefore this removal rate is still fairly bad.

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 fell to 54.50 per cent. Just like Facebook, however, Twitter also observed a bump and rose to 67.95 per cent.

The biggest fall in 2016, however, was seen by YouTube. The platforms 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 in the second quarter. Happily, this fall was short lived, and YouTube rose back to 75.32 per cent in the third quarter, one of the best removal rates ever reached by the company.
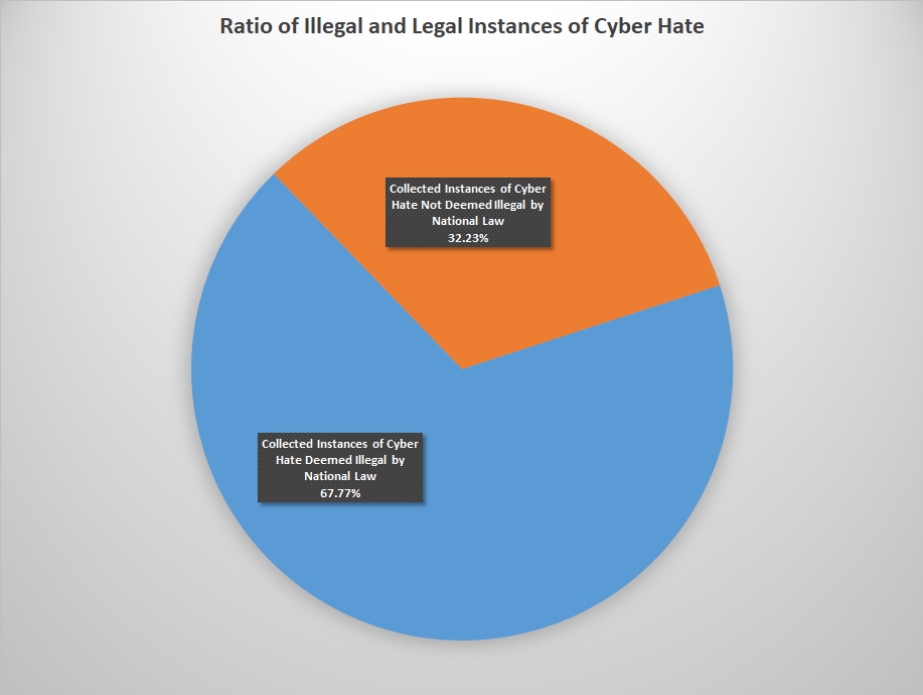
Name of Platform	Percentages of Cases Removed	Percentages of Cases Not Removed
Websites (comments on websites too)	62.50%	37.50%
Blogs	18.18%	81.82%
Forums	75.00%	25.00%
Facebook	62.32%	37.68%
Twitter	67.95%	32.05%
YouTube	75.32%	24.68%
Google+	100.00%	0.00%
Instagram	89.01%	10.99%
VK.com	33.33%	66.67%
Pinterest	100.00%	0.00%
Telegram	100.00%	0.00%
Other Social Media Sites	54.55%	45.45%

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.

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

Ratio of Illegal and Legal Instances of Cyber Hate

Collected Instances of Cyber Hate Not Deemed Illegal by National Law
32.23%

Collected Instances of Cyber Hate Deemed Illegal by National Law
67.77%

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, then to 70.22 per cent and finally 67.77 per cent in the third quarter; the lowest level of illegal cases ever to be recorded by INACH since the start of the data collection.

This means that - in the third quarter of 2017 - 32.23 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. Furthermore, due to multifaceted issues we did not receive data from Spain for the month of August, hence those numbers are missing from the aggregate data in that month.