GA #: 101084457

D-2.5 SafeNet Final Report

Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. The European Union cannot be held responsible for them. GA #: 101084457

Description of the SafeNet project

The 24-month project Monitoring and Reporting for Safer Online Environments seeks to apply a comprehensive and intersectional approach in prevention and fight against intolerance, racism, and xenophobia online. It joins 21 partners, members of the International network against cyber hate (INACH) and the roof organisation itself. All the partners are trusted flaggers, and many have taken part in the monitoring exercises within the scope of the Code of Conduct on countering illegal hate speech online. The project focus on two priorities being 1. continuous monitoring and reporting hate speech content to the IT companies and responsible authorities and 2. awareness raising by regular advocacy towards the social media companies, providing consolidated and interpreted data to national authorities as well as running national bi-monthly information campaigns involving different stakeholders, including IT Companies, public authorities, civil society organisations and media. The project tasks have been organised in 3 work packages consisting of management and organisational framework; monitoring of content deemed illegal under national laws transposing the EU Framework Decision 2008/913/JHA using the methodology from the past monitoring exercises conducted by the European Commission; and dissemination of gathered data to the relevant stakeholders and the general public. Up to 20 000 of cases have been reported, 10 infosheets in English and 170 in other EU languages produced, online training run for the monitoring partners, standards for trusted flaggers reached for all partners, advocacy roundtables and closing conference have been organised. The project fights for targets of online hate based on grounds of racial or ethnic origin, colour, religion, sexual orientation, or gender identity. The second primary target group involves IT companies, national and European authorities, CSOs and media. A wide public will benefit from a kinder internet due to a better and faster removal of hate speech. Project funded by the European Union's CERV-2022-EQUAL.

Table of contents

D-2.5 SafeNet Final Report 1

Description of the SafeNet project 2

Introduction 4

Methodology and Approach 5

1- Training activities 5

2- Why it is important to monitor online hate 6

Protecting Individuals and Vulnerable Groups 6

Promoting a Safer Online Environment 6

Preventing Escalation to Violence 6

Ensuring Compliance with Laws and Policies 6

Gathering Data for Advocacy and Policy Development 7

Empowering Civil Society and Counter-Speech Efforts 7

SafeNet project activities and results 7

WP 1 - Project management and coordination 7

1- Regular meetings 7

2- Continuous evaluation 8

WP2 - The continuous monitoring activities 8

WP3 - Advocacy, awareness raising and dissemination activities 14

1- Production of Factsheets and trend analysis 14

2- Dissemination strategy and implementation - social media campaigns 17

3- Results of the advocacy roundtable 19

4- The SafeNet final conference 20

Future Perspectives with the DSA 21

Conclusion 22

Introduction

The SafeNet project objective is to analyse how IT platforms moderate illegal hate speech. To pursue that goal, the Consortium has followed the basic guidelines of the EU Framework Decision 2008/913/JHA. With the adoption and entry into force of the DSA (UE) 2022/2065, the SafeNet project has also evolved to take account of this European regulation. All 21 SafeNet Consortium members already have expertise in reporting illegal online hate speech, corresponding to each partner's local laws. With this project, the Consortium has gone further and reported harmful hate speech.

The Consortium monitored platforms that have signed the Code of conduct on countering illegal hate speech online[1] , which are:

- Facebook,
- Instagram,
- Microsoft,
- Twitter - X,
- TikTok,
- YouTube,
- Snapchat,
- Dailymotion,
- LinkedIn,
- Jeuxvideo.com,
- Rakuten Viber,
- Twitch.

[1] https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combatting-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en

It is important to note that the partners are not obliged to monitor all platforms listed above. Each monitoring officer prioritises monitoring platforms based on the following criteria:

- Platforms that are relevant in individual countries,
- Have public content (not the content in private or secret groups) and
- Are included in the European Commission Monitoring Exercise.

Methodology and Approach

1- Training activities

To prepare all participants, INACH organized two 2-hour training sessions on January 25, 2023. These two webinars aimed to present the SafeNet project and answer all the partners' questions about its methodology. The coordinators of the project initiated the webinar by explaining the purpose of the project and reminded the Consortium of INACH's official definition of "*hate speech*," "*racism*," "*antisemitism*," "*anti-Muslim racism*," "*antigypsyism,*" and other concepts. The Consortium adopted the relevant definitions and concepts, which enabled consistency in monitoring.

They presented the methodology: how to label and classify the cases. They introduced the data collection and explained how to fill out the data in the INACH platform. They also explained how to comply with the GDPR by removing or blackening all private data except if it is helpful to understand the context of the alleged hate speech.

At the end of the webinars, the participants could ask questions, and they discussed how to improve this monitoring. Following these training activities, LICRA also provided the definition of the Criteria, Parameters and Definitions Guide, which is available in Appendix 1 of the deliverable D2.1.

2- Why it is important to monitor online hate

Monitoring online hate speech is essential for multiple reasons, as it addresses critical social, psychological, and legal challenges. The SafeNet Consortium wants to highlight the importance of continuous monitoring of social media.

Protecting Individuals and Vulnerable Groups

Hate speech disproportionately targets vulnerable populations, such as minorities, women, LGBTQ+ individuals, ethnic or religious groups. Continuous monitoring can help identify and mitigate threats before they escalate into harmful real-world actions like harassment, violence, or discrimination. Victims of online hate speech often suffer from psychological harm, including anxiety, depression, and fear for personal safety. Monitoring allows for timely intervention and support for these individuals.

Promoting a Safer Online Environment

Digital platforms often serve as spaces for public discourse. If left unchecked, hate speech can create hostile environments, discouraging free expression and participation, especially for marginalised communities. Monitoring and addressing hate speech reinforces societal norms of respect and inclusivity, fostering safer online interactions.

Preventing Escalation to Violence

Research has shown that hate speech can incite real-world violence. It normalizes harmful rhetoric, dehumanizes groups, and can mobilize individuals toward extremist actions. By identifying and mitigating hate speech early, it's possible to disrupt pathways that lead to radicalization or organized violence.

Ensuring Compliance with Laws and Policies

Since the adoption of the EU Framework Decision 2008/913/JHA, monitoring is necessary to ensure compliance with and to hold individuals or organizations accountable for violations. And with the adoption of the DSA, online platforms are legally and no longer just ethically obligated to remove hateful content. Monitoring helps them meet these responsibilities while maintaining user trust.

Gathering Data for Advocacy and Policy Development

Continuous monitoring provides data on trends, hotspots, and the effectiveness of current measures. This data is invaluable for crafting targeted policies, improving moderation algorithms, and advocating for stronger legislative protections. It also allows authorities to assess the responsiveness of IT companies to hate speech reports and demand improvements where needed.

Empowering Civil Society and Counter-Speech Efforts

Monitoring hate speech creates opportunities for civil society organisations to counter it with educational campaigns and constructive narratives. This proactive approach addresses underlying biases and fosters greater understanding.

SafeNet project demonstrates that monitoring hate speech is a cornerstone for advocacy and intervention. Working with multiple stakeholders, such as the European Commission, civil society, and tech companies, strengthens the fight against digital hatred and ensures a balanced approach to safeguarding freedom of expression alongside user protection.

SafeNet project activities and results

During the project, all the goals have been achieved. This is due to the generally good coordination of the project. The Consortium has successfully carried out ongoing monitoring of online hate content and dissemination of the project.

WP 1 - Project management and coordination

Management structures have been established and shared with the entire Consortium. Regular meetings were needed to coordinate the project effectively. The coordinator identified the risks and envisaged problems with the online-only management of a big consortium. So, INACH carried out an ongoing evaluation of the organisations.

1- Regular meetings

In order to create a permanent link between the members of the Consortium and ensure the smooth running of the project's activities, a Consortium meeting was organised every two months. 14 consortium meetings were organised throughout the project.

These meetings enabled the members of the Consortium to provide regular updates on each Work Package, to discuss any problems they were encountering in implementing the project's activities, and to share information crucial to the development of reporting procedures on the platforms as part of the implementation of the DSA.

Each month, Work Package leaders' meetings were also organised between INACH, Licra and LGL. 25 Work Packages leaders' meetings have been organised. These meetings provided regular updates on the progress of activities in each Work Package, ensured that all deliverables were produced on time and that collaboration with other consortium members was going well.

2- Continuous evaluation

INACH, the project coordinator, shared the evaluation sheet with each partner and updated it frequently. It measured the quality of communication, meeting attendance, number of cases collected and the dissemination effort. The purpose of the evaluation was to provide support and suggest improvements. In addition, at every consortium meeting, past activities are evaluated. All partners were asked to keep up to date with their project teams, and their CVs were sent to the project coordinator of the new team members.

WP2 - The continuous monitoring activities

During the SafeNet project, the Consortium reported 21,470 cases of online hate speech on December 10, when this report was in preparation.

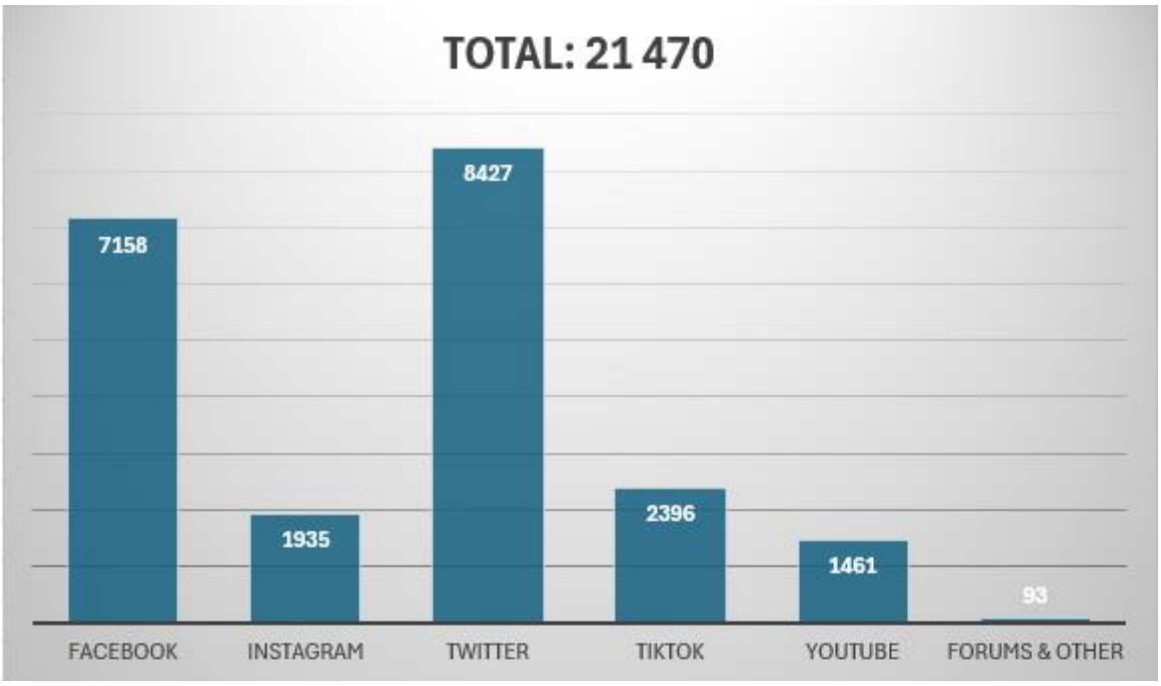
Table 1. Number of cases per partner, per year and total number

Name of the organisation	Total number of cases 2023	Total number of cases 2024	Total number of cases
Jugendschutz	662	926	1588
LCHR	799	815	1619
Never Again	701	772	1473
ITU	630	696	1326

Subjective Values	488	793	1281
DigiQ	577	689	1266
ETEPE	442	654	1096
CESIE	391	692	1083
LICRA	564	484	1048
LGL	400	500	900
INACH	476	346	822
HRHZ	246	579	825
Integro Ass.	406	364	770
Hatter Society	370	488	858
EHCR	281	793	1074
ROMEA	349	445	794
ZARA	257	645	902
Plataforma Khetane	240	557	800
CEJI	375	372	747
MCI	222	540	827
ILGA Portugal	137	233	370

Over these two years, consortium members have been very assiduous in implementing continuous monitoring, the vast majority on Facebook, Instagram, Twitter (X), TikTok and YouTube.

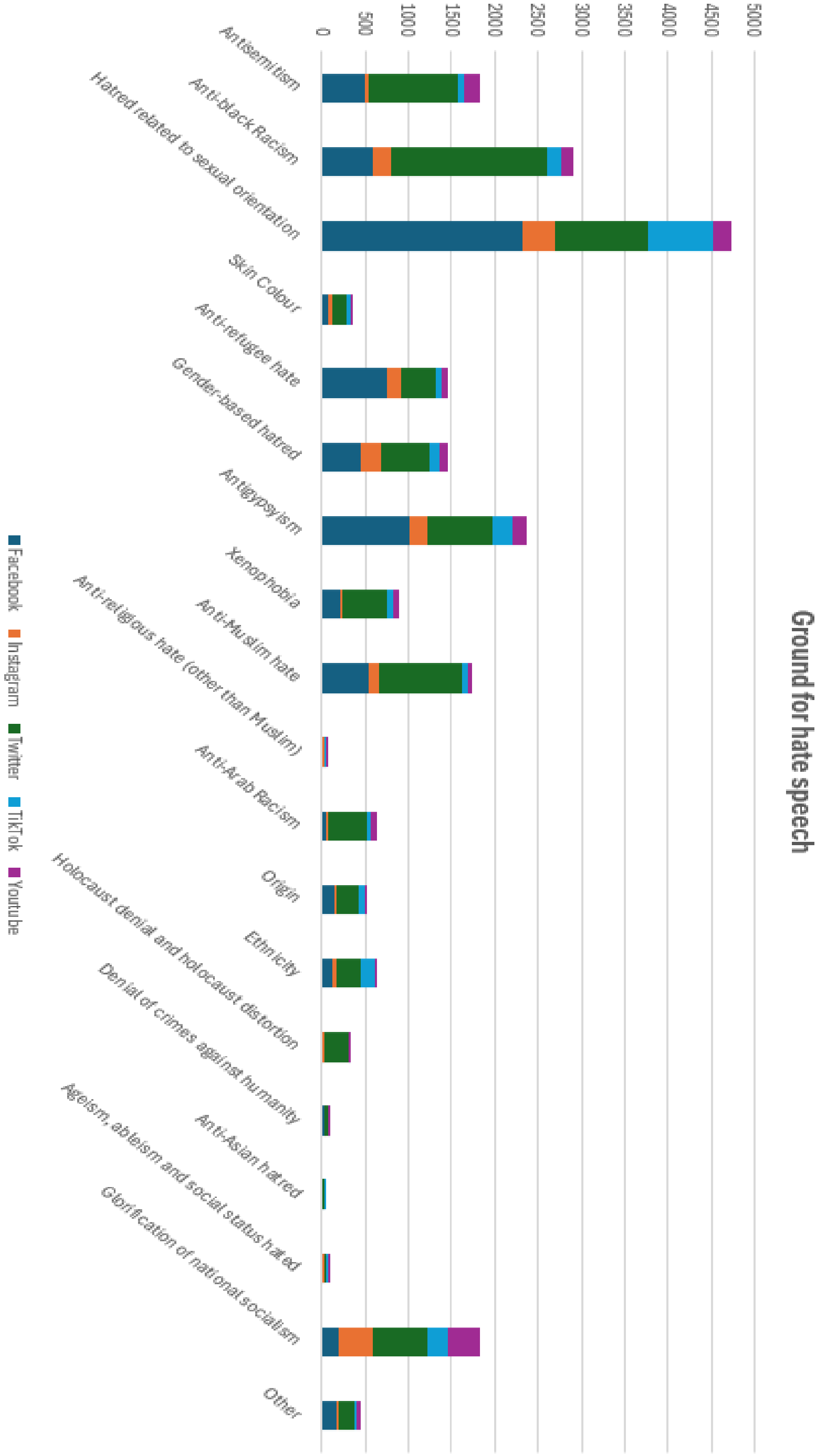
Graph 1. Number of cases per platform



Members of SafeNet's Consortium agreed to monitor online hate speech with an intersectional approach and all types and forms of online hate speech found on IT platforms, including videos, images, memes, and symbols.

It is important to note that partners are not obliged to report all types of hate speech, as some are specialized in specific types of hate. During the SafeNet project, 19 hate motives were monitored.

Graph 2. Grounds for hate speech per platform

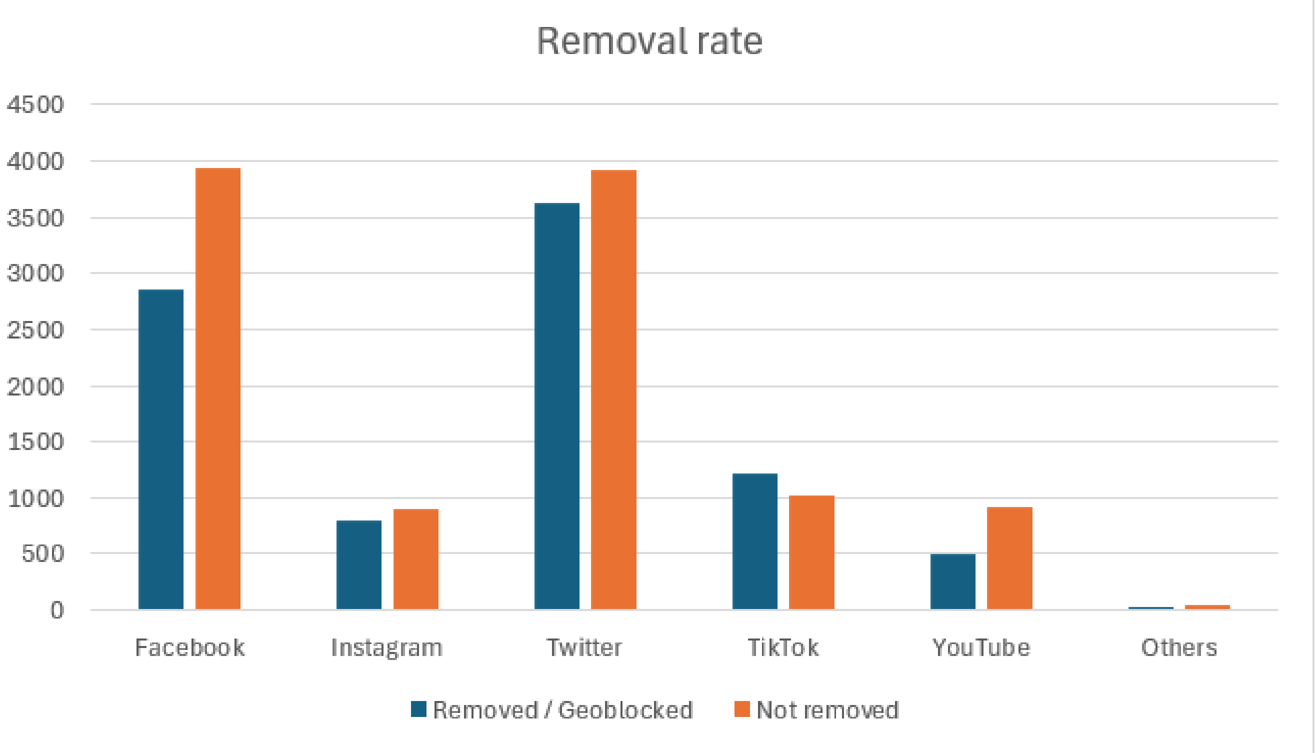

Ground for hate speech
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
Antisemitism
Anti-black Racism
Hatred related to sexual orientation
Skin Colour
Anti-refugee hate
Gender-based hatred
Antigypsyism
Xenophobia
Anti-Muslim hate
Anti-religious hate (other than Muslim)
Anti-Arab Racism
Origin
Ethnicity
Holocaust denial and holocaust distortion
Denial of crimes against humanity
Anti-Asian hatred
Ageism, ableism and social status hatred
Glorification of national socialism
Other
Facebook
Instagram
Twitter
TikTok
Youtube

The hatred related to sexual orientation, anti-black racism, antigypsyism, antisemitism and anti-Muslim racism were the most reported hate content.

Concerning the removal rate, Members of SafeNet's Consortium have expertise in reporting illegal online hate speech, corresponding to each partner's local laws. Members of the Consortium reported illegal hate speech and harmful hate speech. Note that of the 21 470 reports, 1,358 concerned harmful hate speech; it only represents 6.35% of the reports.

One of the aims of the continuous monitoring of online hate speech was to evaluate the reaction of platforms to reports of online hate. As a result, the members of the Consortium unanimously agree that the removal rate is not satisfying. There is still room for improvement for platforms. Despite the end of the project, the members of the Consortium remain available to discuss with the platforms and work together to find solutions to improve the removal of online hate content and encourage members of civil society to do the same.

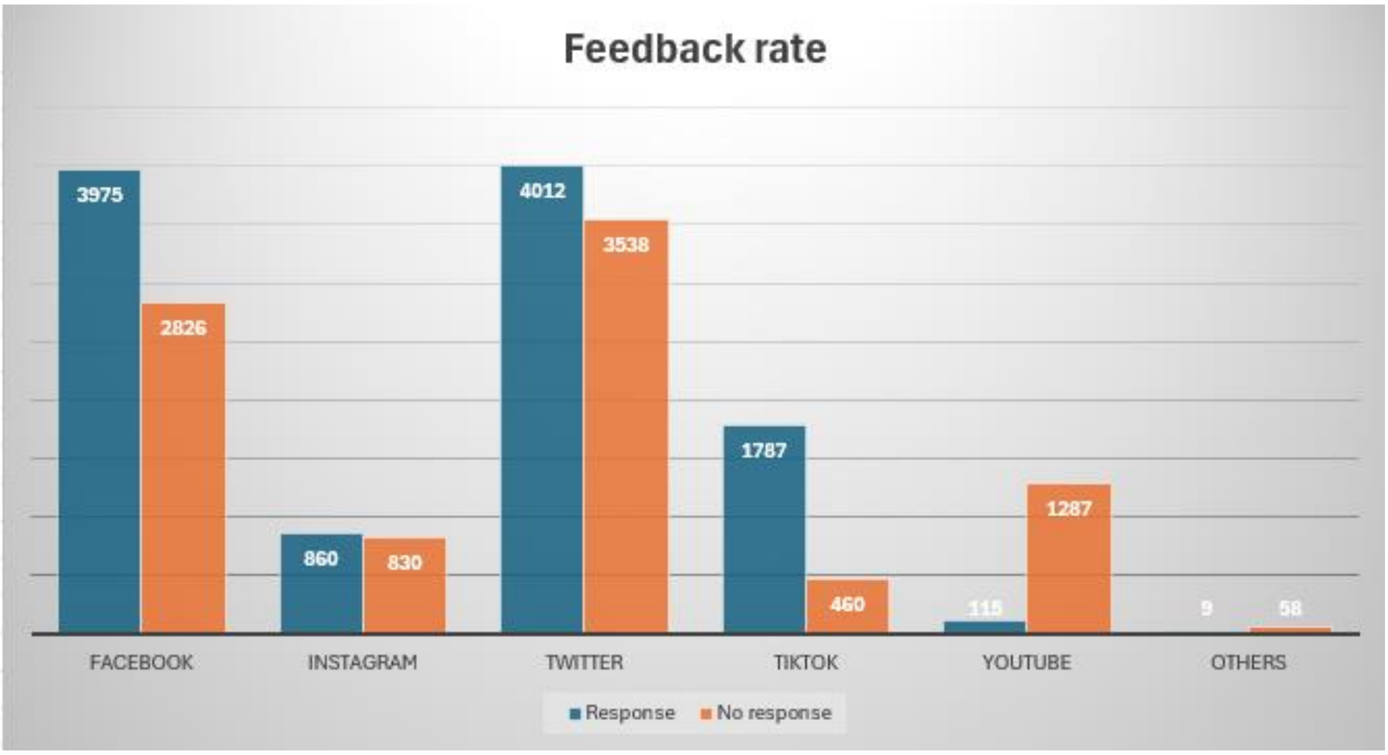
Graph 3. Removal rates per platform



Another aim was to evaluate the feedback received from the platforms after the reports: the responses received, and the rate of responses before 24 hours. The feedback get from the platforms is very important: 1) it allows follow-up on reports more effectively, 2) in case of disagreement, it enables to argue more effectively when reporting as a Trusted Flagger or in case of appeal, 3) it improves communication between the platforms and the members of civil society who make the reports and helps us to understand certain decisions.

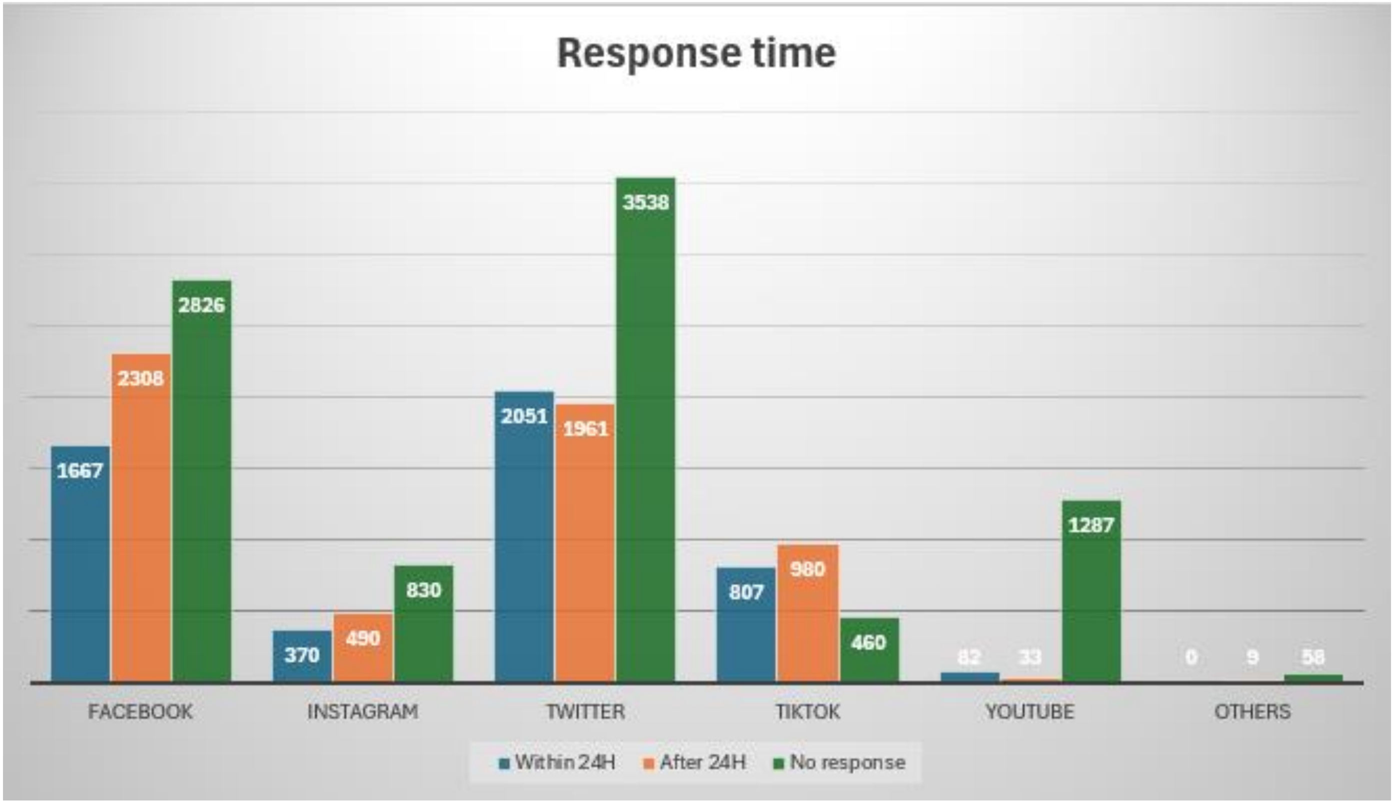
Since the last advocacy roundtable, the Consortium has noticed an improvement concerning the feedback rate and the time for response.

Graph 4. Feedback rates per platform



The time for response has also been increased since the advocacy roundtable.

Graph 5. Response time



WP3 - Advocacy, awareness raising and dissemination activities

The aim of the dissemination strategy was to share the results of the project through Infosheets/Factsheets but also to publicise the members of the SafeNet consortium through social media campaigns. By December 10, 2024, 1,796 publications had been produced by consortium members.

1- Production of Factsheets and trend analysis

To analyse all the reports made by the Consortium, 10 factsheets were produced, documenting the findings of continuous monitoring of hate speech on social media. These Factsheets have fewer cases because the data considered dates from November 2024 while the database considers the reports until December 10, 2024.

Since the beginning of the project, the Consortium analysed 19 486 reports.

Facebook: 6 744

Instagram: 1 752

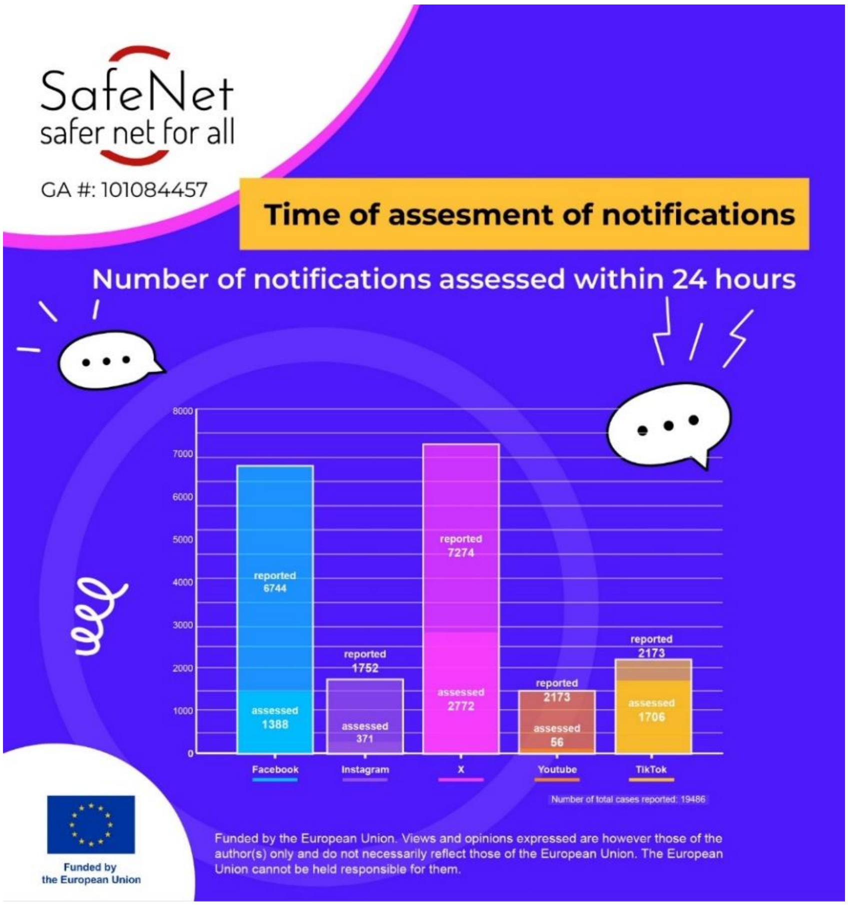
Twitter (X): 7 274

TikTok: 2 173

YouTube: 1 420

Other : 66

Here is an example of the Factsheet/Infosheet. All Factsheets in 19 national languages and one in English can be found at https://www.inach.net/safenet-fact-sheets/

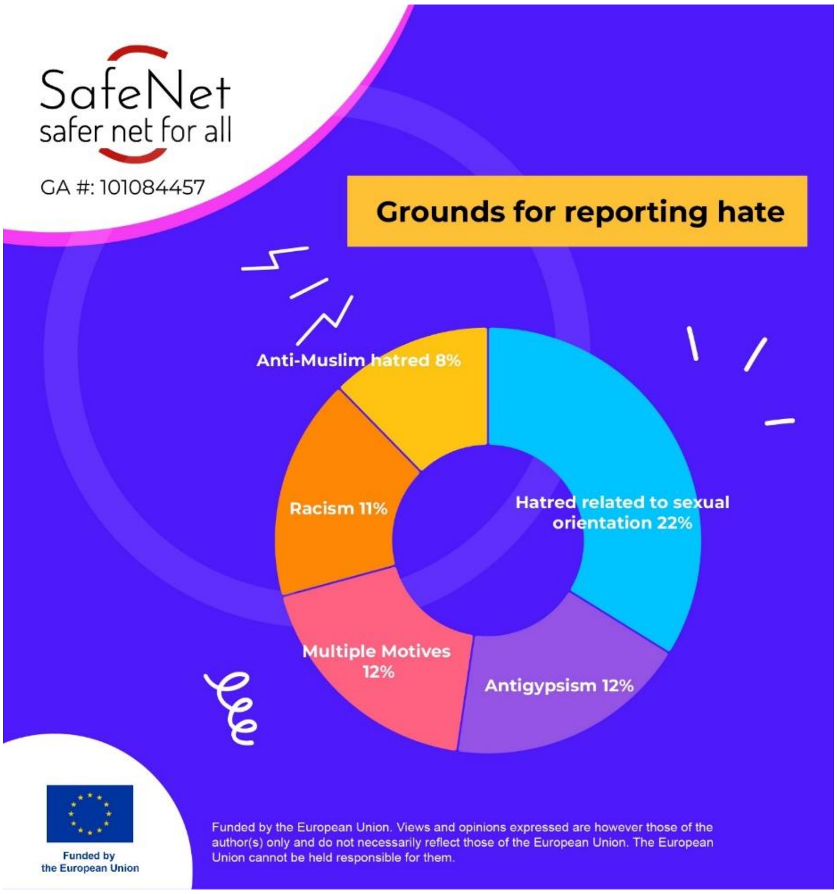


GA #: 101084457

Time of assesment of notifications

Number of notifications assessed within 24 hours

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. The European Union cannot be held responsible for them.



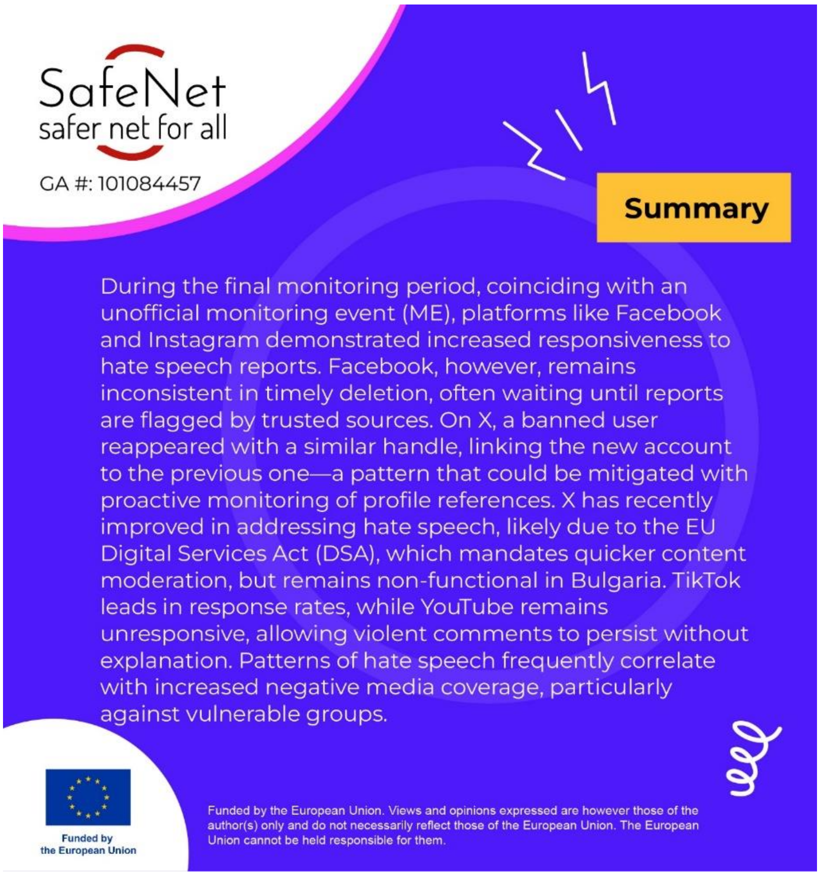
SafeNet
safer net for all

GA #: 101084457

Grounds for reporting hate

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. The European Union cannot be held responsible for them.



SafeNet
safer net for all

GA #: 101084457

Summary

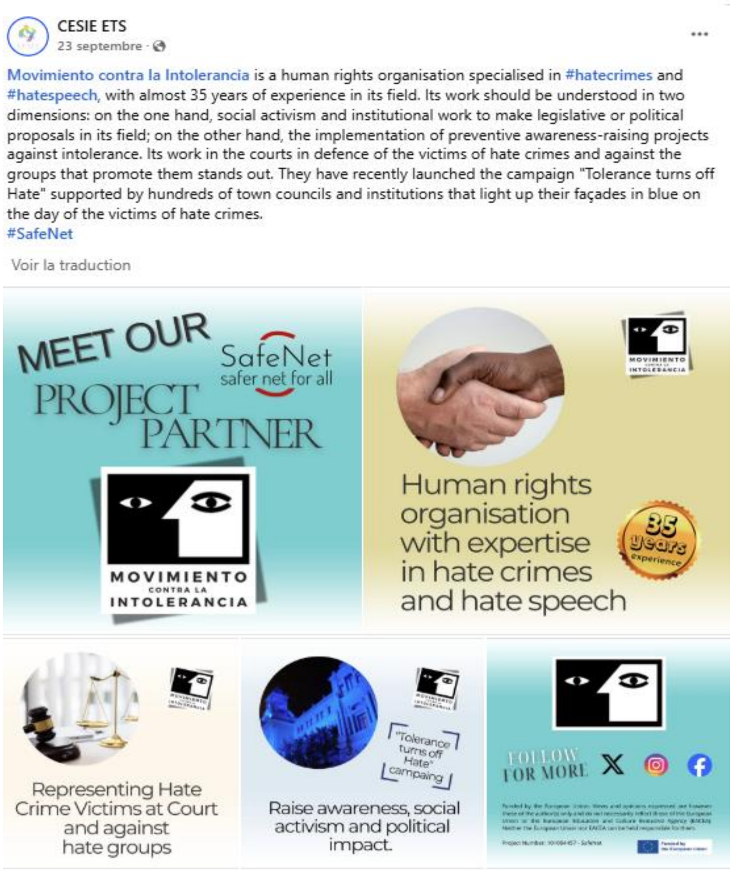
During the final monitoring period, coinciding with an unofficial monitoring event (ME), platforms like Facebook and Instagram demonstrated increased responsiveness to hate speech reports. Facebook, however, remains inconsistent in timely deletion, often waiting until reports are flagged by trusted sources. On X, a banned user reappeared with a similar handle, linking the new account to the previous one—a pattern that could be mitigated with proactive monitoring of profile references. X has recently improved in addressing hate speech, likely due to the EU Digital Services Act (DSA), which mandates quicker content moderation, but remains non-functional in Bulgaria. TikTok leads in response rates, while YouTube remains unresponsive, allowing violent comments to persist without explanation. Patterns of hate speech frequently correlate with increased negative media coverage, particularly against vulnerable groups.

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. The European Union cannot be held responsible for them.

2- Dissemination strategy and implementation - social media campaigns

The social media campaign had two elements. First is the presentation of different project partners and each organization's activities. Second, it is the presentation of the local legal context regarding hate speech. Since the Consortium is large and the project envisages using partners' social media channels rather than creating new social media pages for the project, the social media campaign's purpose is to introduce both different partners and different legal backgrounds of hate speech legal framework throughout the EU and present the project with engaging social media posts. 21 social media campaigns were disseminated. The Consortium entered their dissemination efforts into the project dissemination log every month. The social media campaign is estimated to have generated 972 posts and reached 261,982 users. Here are some examples of partners' social media campaign dissemination efforts:

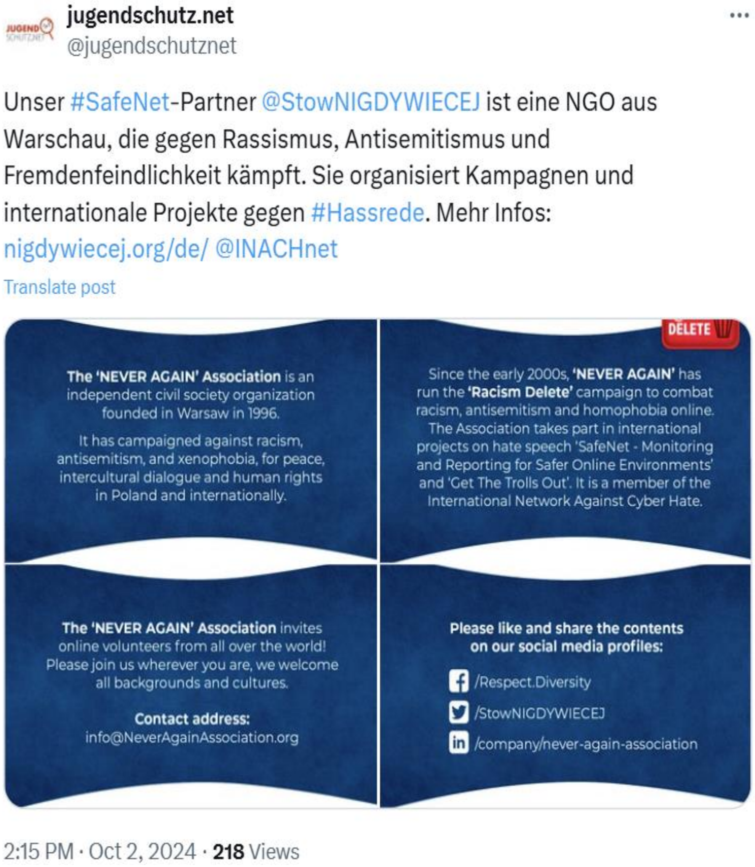


CESIE ETS
23 septembre ·

Movimiento contra la Intolerancia is a human rights organisation specialised in #hatecrimes and #hatespeech, with almost 35 years of experience in its field. Its work should be understood in two dimensions: on the one hand, social activism and institutional work to make legislative or political proposals in its field; on the other hand, the implementation of preventive awareness-raising projects against intolerance. Its work in the courts in defence of the victims of hate crimes and against the groups that promote them stands out. They have recently launched the campaign "Tolerance turns off Hate" supported by hundreds of town councils and institutions that light up their façades in blue on the day of the victims of hate crimes.
#SafeNet

Voir la traduction

MEET OUR
SafeNet
safer net for all
PROJECT PARTNER
MOVIMIENTO CONTRA LA INTOLERANCIA
Human rights organisation with expertise in hate crimes and hate speech
35 years experience
Representing Hate Crime Victims at Court and against hate groups
"Tolerance turns off Hate" campaing
Raise awareness, social activism and political impact.
FOLLOW FOR MORE



jugendschutz.net
@jugendschutznet

Unser #SafeNet-Partner @StowNIGDYWIECEJ ist eine NGO aus Warschau, die gegen Rassismus, Antisemitismus und Fremdenfeindlichkeit kämpft. Sie organisiert Kampagnen und internationale Projekte gegen #Hassrede. Mehr Infos: nigdywiecej.org/de/ @INACHnet

Translate post

The 'NEVER AGAIN' Association is an independent civil society organization founded in Warsaw in 1996.

It has campaigned against racism, antisemitism, and xenophobia, for peace, intercultural dialogue and human rights in Poland and internationally.

Since the early 2000s, **'NEVER AGAIN'** has run the **'Racism Delete'** campaign to combat racism, antisemitism and homophobia online. The Association takes part in international projects on hate speech 'SafeNet - Monitoring and Reporting for Safer Online Environments' and 'Get The Trolls Out'. It is a member of the International Network Against Cyber Hate.

The 'NEVER AGAIN' Association invites online volunteers from all over the world! Please join us wherever you are, we welcome all backgrounds and cultures.

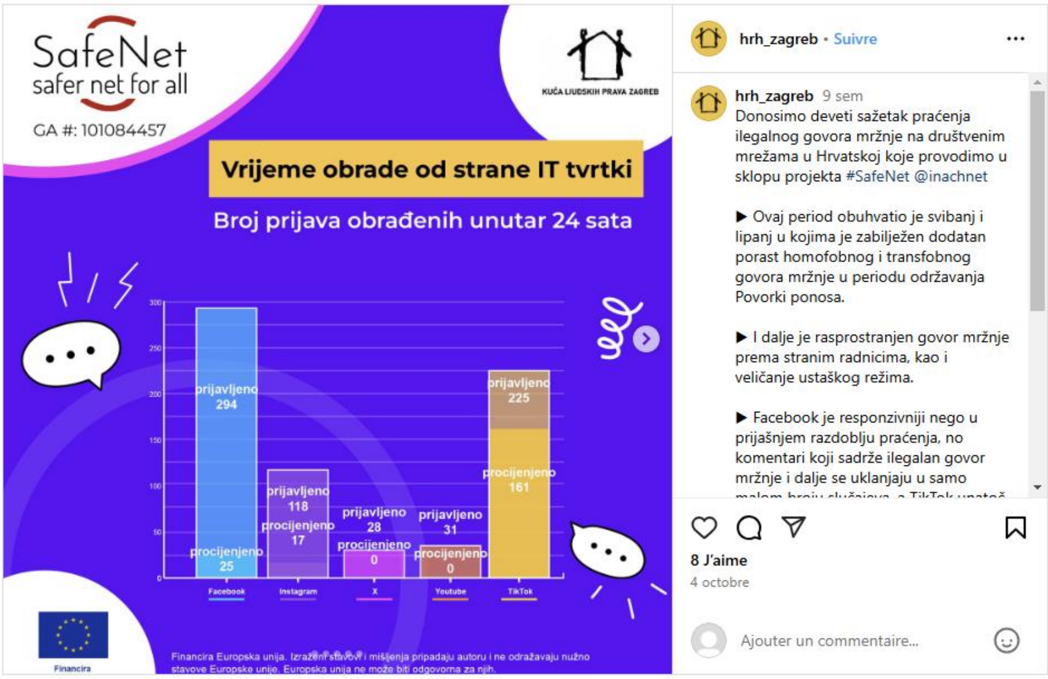
Contact address:
info@NeverAgainAssociation.org

Please like and share the contents on our social media profiles:

/Respect.Diversity
/StowNIGDYWIECEJ
/company/never-again-association

2:15 PM · Oct 2, 2024 · **218** Views

The Factsheets were also disseminated. Since the beginning of the project, the dissemination of fact sheets produced 448 posts and generated 107 261 users reached on social media. Here are some examples of fact sheet dissemination on social media:


SafeNet
safer net for all
GA #: 101084457
KUĆA LJUDSKIH PRAVA ZAGREB
Vrijeme obrade od strane IT tvrtki
Broj prijava obrađenih unutar 24 sata
300
250
200
150
100
50
0
prijavljeno 294
procijenjeno 25
prijavljeno 118
procijenjeno 17
prijavljeno 28
procijenjeno 0
prijavljeno 31
procijenjeno 0
prijavljeno 225
procijenjeno 161
Facebook
Instagram
X
Youtube
TikTok
Financira
Financira Europska unija. Izraženi stavovi i mišljenja pripadaju autoru i ne odražavaju nužno stavove Europske unije. Europska unija ne može biti odgovorna za njih.

hrh_zagreb • Suivre

hrh_zagreb 9 sem
Donosimo deveti sažetak praćenja ilegalnog govora mržnje na društvenim mrežama u Hrvatskoj koje provodimo u sklopu projekta #SafeNet @inachnet

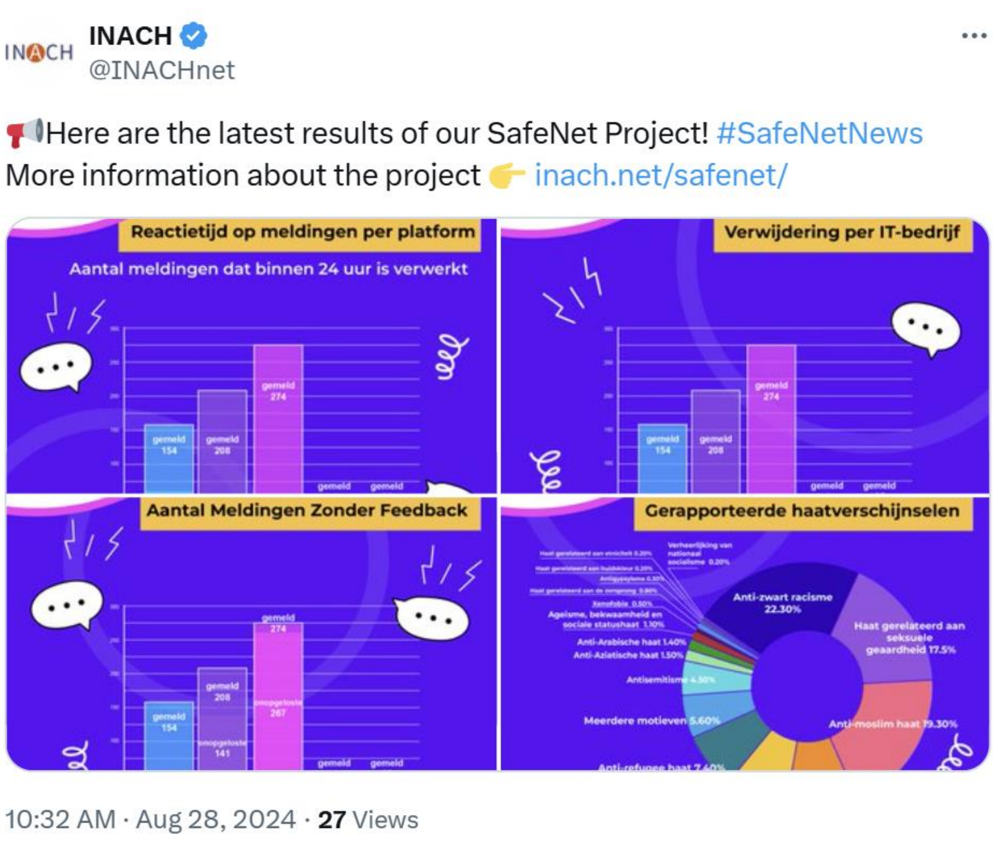
▶ Ovaj period obuhvatio je svibanj i lipanj u kojima je zabilježen dodatan porast homofobnog i transfobnog govora mržnje u periodu održavanja Povorki ponosa.

▶ I dalje je rasprostranjen govor mržnje prema stranim radnicima, kao i veličanje ustaškog režima.

▶ Facebook je responzivniji nego u prijašnjem razdoblju praćenja, no komentari koji sadrže ilegalan govor mržnje i dalje se uklanjaju u samo malom broju slučajeva, a TikTok unatoč

8 J'aime
4 octobre

Ajouter un commentaire...


Reactietijd op meldingen per platform
Aantal meldingen dat binnen 24 uur is verwerkt
gemeld 154
gemeld 208
gemeld 274
gemeld
gemeld
Verwijdering per IT-bedrijf
gemeld 154
gemeld 208
gemeld 274
gemeld
gemeld
Aantal Meldingen Zonder Feedback
gemeld 154
gemeld 208
onopgeloste 141
gemeld 274
onopgeloste 267
gemeld
gemeld
Gerapporteerde haatverschijnselen
Haat gerelateerd aan etniciteit 0.20%
Haat gerelateerd aan huidskleur 0.20%
Antigypsyisme 0.30%
Haat gerelateerd aan de oorsprong 0.80%
Xenofobie 0.50%
Ageisme, bekwaamheid en sociale statushaat 1.10%
Anti-Arabische haat 1.40%
Anti-Aziatische haat 1.50%
Antisemitisme 4.50%
Meerdere motieven 5.60%
Anti-refugee haat 7.40%
Verheerlijking van nationaal socialisme 0.20%
Anti-zwart racisme 22.30%
Haat gerelateerd aan seksuele geaardheid 17.5%
Anti-moslim haat 19.30%

INACH
@INACHnet

📢Here are the latest results of our SafeNet Project! #SafeNetNews
More information about the project 👉 inach.net/safenet/

10:32 AM · Aug 28, 2024 · 27 Views

3- Results of the advocacy roundtable

Advocacy, together with monitoring and dissemination, was one of the core activities of the SafeNet project. Two advocacy roundtables have been organised. The first was held on November 7, 2023, and it brought together all the partners and platforms: Microsoft, YouTube, Meta, X (Twitter), TikTok and Viber. The main objective was to discuss their removal policy and their reporting system. First, the platforms explained their efforts to combat illegal hate speech by updating their community guidelines and outlined how their removal policy of hate speech is managed. Secondly, the Consortium presented the monitoring statistics. The Consortium expressed dissatisfaction about the fact that there is no feedback about the reports, no way to follow, and no way to see if the content was removed or not. The Consortium also noticed that some platforms still allow some Nazi accounts. Finally, the Consortium reminded of the importance of removing hate speech. Indeed, the statistics showed that a great rate of hate content was not removed or was just "withheld in country X" or had "limited visibility."

The second advocacy roundtable was held on May 29, 2024, and it brought together YouTube, Meta, X (Twitter) and TikTok. By then, the DSA, Digital Services Act, had been implemented in several European Union member states and, some of the Consortium's members have either applied for or will be applying for official Trusted Flagger status. Therefore, there was more than one objective. 1) The Consortium discussed the implications of the Trusted Flagger status as it stands. It seems important to broaden our focus beyond just the Trusted Flagger designation. In some countries, this status can be problematic as it aligns organisations too closely with the government, restricting their ability to secure funding and collaborate with social media platforms. 2) The Consortium shared with the platforms the result of the continuous monitoring in order to compare them with the claims and figures of the platforms. The Consortium was, therefore, able to raise the issue of withholding hate speech rather than removing it, the fact that Nazi accounts are still active on some platforms, and the lack of response to reports made. 3) To compare the data since the first advocacy roundtable to evaluate the improvement. This advocacy roundtable showed a slight improvement in the number of responses received after the reports. However, the removal rate still needs to be improved by the platforms.

These two advocacy roundtables have had an impact on collaboration between platform representatives and consortium members, who have found key contacts for each platform. Any other problems encountered by any of the consortium members were resolved directly with the platform concerned. This improved the relationship and communication between the SafeNet consortium and the platforms.

4- The SafeNet final conference

The final SafeNet conference was held on November 6, 2024, in Brussels, bringing together 50 participants (24 online and 26 in person). A representative of the European Commission and representatives of Twitter and YouTube were also present among others.

The main results were presented during the conference: the overall grounds for hate speech and specific country analyses by social media platforms, the removal rates and response times and the development of issues most often encountered during the monitoring process. The goal was to foster a multisectoral approach and open dialogue. Participants had ample time to ask questions, provide feedback, and offer suggestions for the future development of continuous monitoring, which was significantly influenced by the new DSA.

Special time was dedicated to advocacy efforts reporting and experience. Only continuous monitoring by many partners can provide a picture of platforms' responses and policies toward hate speech online and allow partners to react quickly to a constantly changing hate speech environment.

Future Perspectives with the DSA

The Digital Services Act aims to regulate the activities of the platforms. The objective is to make digital actors responsible so that they fight against the propagation of illicit and illegal content on their services. As soon as it was adopted and implemented, the members of the Consortium immediately took the DSA into account in their continuous monitoring of online hate speech and identified several challenges.

Firstly, to ensure that all obligations arising from the DSA are enforced in each Member State, in particular:

- The obligation to report a serious illegal act by a user to the European Commission,
- The implementation of technical and organisational measures to process alerts issued by trusted Flaggers,
- The obligation to suspend, after prior warning, for a reasonable period, access to the service to a user who frequently provides manifestly illicit content.

The second challenge is to have effective and efficient sanctions against platforms that do not respect the obligations arising from the DSA. And the third challenge would therefore be to find how to discuss with the authorities of the Member States and the platforms to ensure compliance with the DSA. Civil society organizations have their place in all these challenges. It is thus crucial for CSOs to keep their independent status and have means to continue with this type of hate speech monitoring and advocacy.

Conclusion

The SafeNet project has more than achieved its objectives. The number of hate messages reported was 21,470, compared with 18,000 initially forecast in the project. The risks linked to the coordination and management of a large consortium were identified at an early stage and avoided. The dissemination of the project has been greatly successful, with more than 369,000 people reached by the publications of the consortium members. The Consortium produced 10 Infosheets in 19 national languages and one in English presenting them in a form visually adapted for social media. Thus, the Consortium spread awareness and relevance of online hate speech monitoring in the national and EU context. The Consortium significantly improved advocacy efforts and contact with platforms. The results of the project were presented also on various workshops and meetings. Infosheets, particularly important for the national context since produced in 19 national languages and covering cover States, had a very strong impact on raising awareness about online hate speech. Each Infosheet is visually adapted for social media use and dissemination. Each Infosheet also contains a summary in national languages and English. Consortium members collected 21 470 cases of reported online hate speech incidents across various social media platforms. The data collected present a unique and valuable foundation for analysing hate speech trends and social media platforms' reporting process and removal strategies. Continuous monitoring by civil society organizations is thus essential for the empowerment of regular users, and it also serves as the basis for fact-based policies. The two advocacy tables organised by the Consortium increased CSOs' leverage in negotiating with platforms. The final conference brought together a broad spectrum of stakeholders.

According to the results of monitoring hate speech on social media, we note that platforms have room for improvement in moderating hateful content online. When we compare our figures, we notice an improvement in platform reactions between the start and the end of the project, as well as better communication between the members of the Consortium and the platforms in resolving problems linked to the continuous monitoring of hate speech.

The members of the Consortium will continue their efforts in the fight for a safer Internet and, the results of the project are essential data for continuing this fight.