INACH

Bringing the Online In Line with Human Rights

INACH 20 ANNUAL CONFERENCE REPORT

2002 to 2022 – 20 years of combating cyber hate: lessons learned and future challenges

Compiled by

Adinde Schoorl

2022

TABLE OF CONTENTS

International Network Against Cyber Hate – INACH

INACH was founded in 2002 to use intervention and other preventive strategies against cyber hate. The member organisations are united in a systematic fight against cyber hate, for example as complaints offices, monitoring offices or online help desks. In their respective countries, they provide important contacts for politicians, internet providers, educational institutions, and users.

Funding for INACH is provided by its members, the European Commission, and other donors. The International Network Against Cyber Hate (INACH) unites multiple organizations from the EU, Israel, Russia, South America, and the United States. While starting as a network of online complaints offices, INACH today pursues a multi-dimensional approach of educational and preventive strategies.

Introduction

INACH's conference in 2022 took place on the 3rd and 4th of November. The title of the annual conference was: '2002 to 2022 - 20 years of combating cyber hate: lessons learned and future challenges'. With a mixture of academics, representatives of NGOs, other stakeholders and social media platforms, the theme was discussed from different perspectives. The conference took place in the Apollo Hotel in Amsterdam. Over 150 participants registered and joined both in person and virtually. This year was a special edition since INACH celebrated its 20th anniversary. Therefore, the conference lasted two days. During these two days the conference looked at how cyber hate has developed over the last twenty years, new legislation as the Recommendation on combating hate speech by the CoE and the DSA, and the obstacles in Russia when it comes to cyber hate. Also, there was room to focus on gender-based hate and LGBT+ hate with two keynote speakers and different workshops. For the second time the Ronald & Suzette Award was presented.

DAY 1

The conference was opened by the moderator, **Tamás Berecz**, and the Chair of the INACH Board, **Philippe Schmidt**. Philippe welcomed everyone and expressed how great it is to be all together in person again after having had two virtual conferences in the last two years. He thanked the Secretariat for making this conference possible, the co-board members, the EU and the sponsors Google and Twitter. Finally, he thanked the moderators and the speakers who will make this conference possible. INACH was founded by Jugendschutz.net and Magenta Foundation in 2002 and the start was marked by the beginning of social media. Nobody was aware of the effects of these new digital streets. INACH was a forerunner, and the idea of the network was unique. Today, INACH has 32 members from all over the world: in Europe, the US, Latin America, and Africa. INACH will continue to fight for human rights on the internet in the next twenty years.

The first speech was given by **Stefan Glaser** titled **'No soup for you'**. Stefan was there when INACH was founded and therefore he was the right person to give a historical overview of INACH and its origins. In March 2000, Stefan worked for Jugendschutz.net and he met Suzette and Ronald for the first time in the office of LICRA. The idea of a network became concrete because it was clear that online extremism required an international approach. The network was founded under Dutch law and one important question was what to name it. In the end they decided for INACH, but it turned out to be a name that was very hard to pronounce for many.

For many years, INACH had no money, no room, and no staff. Regardless, it always was ambitious to what it should develop into. INACH started with four full members and the reality of the first few years was that members left or joined. Suzette and Ronald often referred to the series Seinfeld and the 'Soup Nazi' who would say: 'no soup for you.' They often used that sentence when they did not agree with people. To be able to work with them, one had to be open minded and accept the special spirit of the INACH humour. Eating together was usually the most important part of every meeting. INACH held conferences in Paris, New York, Berlin, Warsaw, Kiev, Vienna, Jerusalem, Strasbourg, Brussels, Basel and many more. Both Suzette and Ronald are sorely missed as uncompromising fighters. Hate is still viral on the internet, maybe more than ever. but looking at INACH today gives all of us hope that something is done about it. The network is bigger and more diverse than it used to be, and Stefan called upon everyone to keep INACH running for another 20 years.

The first Keynote speech was given by **Karolin Schwarz: 'From websites to social media'**. At the beginning of the creation of internet there was a naive conviction that far right extremists in Germany could not use the internet and cooperate internationally because they do not speak English. That turned out to be an incorrect conviction. An example of the first tools to cooperate: info phones which were answering machines one could call and the machine would answer with information. After that, the first websites followed and webrings, those linked all the white supremacist pages.

At first, there was little awareness and little regulation. For example, domain names could get registered without being checked at all. Video games were used early on to address a

younger audience. One example: a game called 'ethnic cleansing'. These games were and are still full of insider jokes.

Twitter was in 2012 the first to ban an account when it was requested by the police to do so. At the time, the platform was only six years old. The talking strategies by extreme right groups to reach an audience have not changed since the beginning of the internet; they adapt their message to every platform to every format, in order to be able to repeat it as much as possible. It is not about who is right but who the audience thinks is right. The extremist groups still use the main social media platforms but there is also a range of spaces that are use due to the little moderation they conduct. And then there are the platforms that have been created exactly for that purpose; to receive these far-right groups.

The first presentation was a joint cooperation between **Menno Ettema** and **Tarlach McGonagle**, on **'The Introduction to Recommendation CM/Rec(2022)16 on combating hate speech'**. Menno introduced the Recommendation of the Council of Europe. The CoE takes the perspective of the victims of hate speech. The development of the Recommendation was a multi stakeholder process: member states,

intergovernmental institutions and CSOs were all involved. Public consultation rounds with social media companies and other actors were held as well. Therefore, it is a very broadly carried recommendation that was adopted last May in Italy, Turin. It is important to differentiate between hate speech that is prohibited under criminal law and subject to civil or administrative law or harmful speech that still merits action. Hate speech is a broad concept that has different levels of harm and effects and therefore needs a comprehensive approach that includes criminal action but also other essential efforts. CDADI & CDMSI monitor its implementation by member states, with review of promising practices at national level of the implementation of relevant aspects of the Recommendation and review of the state of implementation of the Recommendation by member states. And finally, promotional activities and capacity building efforts through cooperation projects.

Tarlach explained that freedom of expression is a powerful tool but can also be used as a tool to destroy freedom. Politicians often talk about freedom of speech but not about the obligations that come with it. What is the added value of the Recommendation? It is not legally binding, does that retract its usefulness? Is it a tool that can be put to effect in our professional activities? There are a lot of examples where the European Court of

Human Rights decided between what is free speech and what is hate speech. Politicians are crucial actors in combating hate speech, they have legal and moral obligations to condemn hate speech at every opportunity. It is also necessary to look beyond the usual suspects and pay attention to the psychological, emotional, and financial needs of the victims of hate speech. Monitoring and analysis are crucial as well as national and international coordination. And finally, it is so important for CSO to be involved in combating hate speech.

The second presentation of the day was given by **Josephine Ballon** of Hateaid about the '**The Digital Services Act – will it protect users from online violence?'.**
The DSA will be in force for the Very Large Online Platforms from 17 November 2022 and for the smaller platforms from February 2024. The DSA replaces national legislation, for the benefit of having horizontal regulation for the whole EU. The DSA is not a tool for criminal law enforcement, it is only for platform regulation. There are no fundamental changes in the liability regime and the country of origin principle stays, which means that platforms have to mainly adhere to the rules of the member state where they decide to have their headquarters. In the case of social media platforms that means Ireland. Finally, they are not liable for the content that is uploaded on their hosting services, only on the pieces of content they know about. The DSA does not include a definition of illegal content. Illegal is what is illegal in the member state. The DSA is not only for social media platforms, it is also for websites, copyright violations, consumer protection violations. So, it is not a special social media law. There is no takedown obligation in the DSA, only an assessment obligation, and platforms must be enabled to identify the illegality of the content without a detailed legal examination. Every EU member state will have a Digital Services Coordinator. So far, it is mostly not clear who that will be. Regarding the Trusted Flagger Status, NGOs should contact the Digital Services Coordinator to apply for that status. At any moment it can be decided to revoke the position of Trusted Flagger.

The last presentation of the day was given by **Alexander Verkhovsky** on **Policies and regulations beyond the European Union - in Russia and its sphere of legal influence**. The Russian government defines extremism as: 'a form of political activity that overtly or covertly rejects the principles of parliamentary democracy.' At the beginning, the new laws had some positive sides. Law making became easier because it could be based on this law also for hate crime. The government was actually quite successful in lowering violent hate crimes. However, this definition is very broad. The vague definition lets the government decide what it wants to act against by calling it extremism. It helped produce very wide criminal norms: incitement became the tough criminal act instead of the violence itself. The government can be seen as a violent group as well or the football club for example.

There is a package of new legislation since the war in Ukraine. Many Russian citizens supported these hate speech laws in the beginning because of the reality of Islamic terrorism for example. But now they see it as a repressive mechanism. Nobody knows anymore what is punishable and what not. What can be called hate speech depends against whom the hate is targeted. Therefore, nobody understands where the red line

exactly is. 'Everything for friends, and the law for enemies.' Russian society will need years, in a good scenario, to come out of this.

The first day of the conference was closed by INACH Board member **Steffen Eisentraut,** who reminded us to stay inspired because everything starts with people and ideas.

****END of DAY 1****

DAY 2

The second day of the conference was opened by INACH Board member **Selma Muhič Dizdarevič**, who welcomed everyone and wished everyone a good conference.

The first presentation of the day was given by **Jordy Nijenhuis** on **Conspiracy theories and misinformation - Stories from the European Observatory of Online Hate.** In this world where polarization, hate and disinformation are on the rise, his mission is to offer an alternative to the increasing black and white world we live in. The European Observatory of Online Hate uses a dashboard, set up together with Textgain, that works on early detection of hate in 24 European languages plus Arabic, Russian and Turkish. Currently it monitors 15 social media platforms. Here are two examples of narratives that were found through this dashboard regarding different topics.

1. The war in Ukraine: The Kremlin led propaganda is very active, far right groups will hop on their propaganda and spread it wider. There are different narratives here: pro Russia narratives, anti-Ukraine narratives, narratives on NATO and narratives on the EU.
2. Dutch farmers and their protests against the government restrictions on nitrogen emissions. The example of the Dutch farmers shows how the public debate develops; it polarizes, then radicalizes, after which conspiracy theories jump on it and finally it is amplified by it and spreads widely.

A few conclusions: conspiracy theories are fluid, anti-establishment and antisemitism are at the core of these theories, audiences move from theory to theory and finally, let's not underestimate the power of a good story.

The first Keynote speaker of the day was **Ksenia Termasina** on **'Gender-based hate speech: a Ukrainian approach'.** We all know about the Pyramid of Violence, which starts with problematic and hateful language and can lead to violence and murder. In other words, the issue cannot be underestimated, the impact on mental health is huge. Ukrainian media still portrays women in a sexist manner. There are plenty of examples of it in advertisement or news articles with a title like 'The sexiest female political deputies'. There is a draft law on gender-based hate but it is currently not progressing due to the war in Ukraine. To combat hate speech, there is the cyber police and there is criminal law, but practice shows that the best way to combat hate speech is by education. Finally, feminism is not just about women's position in society but profound changes against discrimination and violence.

After that, the first round of parallel workshops regarding gender-based hate took place. The first workshop was presented by **Elizabeth Cappon** and **Selma Muhič** and moderated by **Cordelia Moore** on the **Normalization of online gender-based hate: how made-to-measure AI solutions and policy making can help tackle this problem**.

Elizabeth explained about the problem of the sexist bias in AI. AI needs training data to be developed but what to do when the training data is flawed? There are the funny mistakes such as pictures that recognize guacamole as a cat or a chocolate chip cookie as a chihuahua. But tests done with AI on: what is Covid?, also generated the words hoax and Great Reset. And the question: what is a Jew? generated the words hate and sh*t. Training data of AI often has a sexist bias which produces an AI that does not recognize sexism. As a result, sexist content stays often undetected and invisible.

An example: the incel culture has as a theory that 80% of men are not getting sex because 80% of women has sex with 20% of the men. In other words, women are not giving free sex to all men equally. On the platform 4Chan 70% all comments mentioning women are

toxic while on twitter it is only 10%. The top four of recurring themes in these toxic comments are: traditionalism, racism or real violent threats on rape and violence. Antisemitism is added to the theory as well because: it is the fault of the Jews; they are trying to break up the traditional Christian family. These are just a few obscure platforms, just a few pathetic people? If one measures the typical incel themes and words on mainstream platforms, one can conclude that these theories are also trickling down there and there is a clear increase according to the data. Besides, there are a number of deadly attacks that are linked to the incel community. Finally, the percentage of severely toxic comments regarding women of colour is much higher: about 80 % compared to 10% to white women. What is the solution? AI should be specifically trained on sexism and gender-based hate and we need to keep pressing for regulation or take down incel and fringe media.

Selma used a qualitative approach to gender-based hate. Harmful hate speech, online violence against women and gender-based online abuse as phenomena are growing rapidly. There are three types of digital feminism: street activism planned through digital media, internet-only movements created on social media and activism fighting against internet-related problems. Strategies to combat gender-based hate include moderating comments, exposing insult, adapting and solidarity. The effects of online gender-based hate, based on studies from all over the world are shocking: it has a silencing effect on women, they are scared of dealing with social media again after and it leads to self-censorship. Women in positions of power like female politicians or public figures receive even more hate. Next to hate speech and threats, users also publish private information about them such as private addresses, private conversations, and pictures of their children. The hate and violence are often even worse for women of colour or women who are part of the LGBT+ community. However, gender-based hate is often not recognized. The experience of the Monitoring Exercise shows that gender-based hate is almost always rejected. It shows the invisibility of sexism and hate. An example of intersectionality is Roma women. There is a lot of research on Roma people but at the same time it is difficult to find any research on the gender dimension regarding Roma women. Therefore, here are some results from research done by Czech Women's Union. Most of the respondents were elderly and had enjoyed secondary education. A quarter of the respondents indicated they feared engaging in online discussions, the vast majority did not know how to hateful content but most of them knew what a hoax was and what fake news entails. Most of them also distrusted news on social media and preferred traditional media like the radio. It is necessary to

think about the vulnerability of women who have an identity that disadvantages them in some way.

The second workshop was presented by **Paula Thijs, Javier Koole** and **Laura Kaun**, and moderated by **Andrea Cox** on **Gender-based hate speech and cyber violence – State of Play and recommended actions.**

Javier and Paula explained the difference between sex, gender, and gender stereotypes. Sex refers to the physical traits of people, gender refers to how people feel and present themselves and gender stereotypes are ideas and expectations of how women and men should behave in society. There is no universal definition of gender based cyber violence yet. Many terms do not cover the full range of behaviour, nor the gendered patterns in cyber violence. Cyber violence is gendered in the sense that women and LGBT+ people are more at risk of receiving it, it is deeply rooted in existing dynamics (ideas of inferiority of women, online violence spills over to offline violence) and there is a digital gender gap (referring to the access and use of digital technologies and internet, the development of the skills needed to use digital technologies and to participate in the design and production and the advancement of women in decision-making roles in the digital sector). Gender based online violence has a severe impact on the physical and psychological health of women and it has a silencing and self-censoring effect. What needs to be done to tackle it? Create gender sensitive interventions, focus on schools and youth care institutions, and focus on the perpetrators, who are mainly male and both strangers and intimate partners. The role of bystanders who blame victims needs to be addressed as well.

Laura explained that the goal of the European Women's Lobby is to bring the feminist perspective to Brussels, to the EU. Hate speech and online violence affect the offline world. Online violence against women and girls is a reality, it is a gendered and prevalent phenomenon. It is part of the continuum of male violence, impacts all women and perpetuates gender

inequality. It affects women of colour and women with disabilities disproportionally. The Covid-19 pandemic led to a spike in online violence.

Recommended actions:

- Prevention (feminist sexuality education, training of teachers and awareness of the definition of gender based cyberviolence).
- Protection (data collection on cyber violence, digital literacy, and accountability of platforms).
- Prosecution (harmonized criminalization of gender-based cyber violence, functioning law enforcement and accessibility of reporting of cyber violence).
- EU legislation (Directive on violence against women and domestic violence, ratify and implement the Istanbul Convention in all member states and seek harmonisation of existing EU legislative instruments).

The second keynote speech was given by **Vladimir Simonko** on **Activism against ignorance: Lithuanian success (?) story.** Vladimir explained that he will present a long-term activist connection between hate speech and hate crime from a personal perspective. He has been an activist since 1993 in Lithuania. Lithuania became independent in 1991, before it was part of the USSR. It took three years to change the law about LGBT+ in Lithuania and to decriminalize it. In 1993 Amsterdam was the best place

to be, that is where informally LGL was established (Lithuanian Gay League). It was decided to publish a gay magazine, Amsterdamas. but the Lithuanian authorities argued that the magazine was seen as an erotic magazine and therefore it would not be allowed to be distributed widely. In other words, the Lithuanian government censored it. It took until 1995 to fight with authorities about making the word gay legal in the name of LGL.

Gay is not a Lithuanian word, that was the justification of the authorities. It was a time of hope and optimistic youth. They wanted rights right now and to change the constitution. Vladimir and his partner came out publicly for the first time in a Lithuanian newspaper, with names and private information. As a consequence, they got recognized and started to realize the meaning of hate speech and hate crime. They experienced beating, hate mail, beating in a shop, threatening phone calls. When they called the police, they suggested they should move to another more 'gay friendly' neighbourhood. In 2007, a smoke bomb was used in a club during the annual conference of ILGA in Vilnius. The perpetrators were never found. An ad on a trolleybus on normalizing it to be gay was taken down because drivers refused to drive it. The 2010 Baltic Pride in Vilnius was the first gay pride. There were 350 protestors and 2000 anti-protestors. The participants were

caged behind security fences so that police could protect them from the hostility of thousands of counter-protesters. In 2018, the office of LGL was attacked and again the perpetrators were never found. Today LGL keeps on fighting against hate speech, online and offline.

After that, the second round of parallel workshops regarding LGBTQ+ hate took place. The first workshop was presented by **Eglė Kuktoraitė** and **Róbert Buzsáki** and moderated by **Tomer Aldub**i on **NGOs fighting against LGBTQ+ hate speech, lessons learned from Lithuania and Hungary.**

Róbert explained that there is criminal law against anti LGBTQ+ hate speech in Hungary. There is section 332 and 333 of the Criminal Code. However, in practice only a very narrow interpretation of these laws are conducted; only incitement liable to provoke violent acts is considered as hate speech and there is the issue of underreporting. Regarding incitement against a community, section 332 of the Criminal Code, only 30 cases were registered since 2013. The effect of political discourse and LGBTQ+ hate speech: the Hungarian prime minister said homosexuals should keep their hands off of children. This statement was trending online and provided ground for hate speech. In 2021, the parliament passed the propaganda law which bans any discussion or information on LGBTQ+ in schools and media. It imposes a ban on every product or advertisement related to LGBTQ+ people that targets people under 18. Again, it provides ground to accept hate speech.

Eglė explained about the situation in Lithuania. The country has a good legal context on hate crimes but there is nothing specific on hate speech. Also, people do not report hate crimes and hate speech to local authorities because they don't think that law enforcement will do anything against it. There are some positive developments: a more elaborate hate speech definition has been established. It also stresses on a proactive role of law enforcement professionals due to the sensitive nature of

hate incidents. When it comes to social media platforms and removal of hate speech, the

examples given by Eglė show that it is very difficult to predict what cases will be removed and what cases will not.

The second workshop was given by **Sara Soares** and **Júlia Mózer** and was moderated by **Valentin Gonzales** on **LGBTQ+ victim protection and challenges, the NGO and IT perspectives.**

Júlia explained that Twitter has brought in a policy update that addresses harassment against women and non-binary people. It is possible now for a user to remove followers and it is possible to unmention yourself if someone has done so and you do not want to be involved in that hateful post (it will still be in the tweet but not anymore on one's wall). Also, a warning appears now before someone posts something potentially hateful, telling them that what they are planning on posting is hateful, so they should reconsider posting it. Many people change their tweet after that warning, so it shows that it is actually working. This warning is only available for now in English, but Twitter is working on expanding it to other languages.

Sara explained that Portugal has changed its hate speech laws in 2012 which better protects victims and defines hate speech better. It also includes gender identity and sexual orientation. However, the law is not applied enough, prosecutors rather use other charges that are easier to prove, such as defamation.

Most victims do not report hate speech. They are afraid to do so and have a lack of trust in the system. As an effect, hate speech is normalized, people accept that it is part of life. Hate speech online always has multiple victims, not just the person targeted by it. At the same time, statistics show that reporting is growing slowly.

The final presentation of the conference was awarding the **Ronald & Suzette Prize** by **Jolie van der Klis** of the Magenta Foundation. INACH's Board member **Jonathan Vick** gave a speech through a pre-recorded video to introduce the Prize and announce the winner: **Alexander Verkhovsky** and the **SOVA centre for Information and Analysis.**

Moderator **Tamás Berecz** closed the conference by thanking everyone for their attendance and participation.

****END of the CONFERENCE****

If you are interested, <u>here is the link</u> to the PowerPoints of the conference.