



Case Handling in the Context of Content Guidelines and Legislation – A Case Collection

By
Dr. Steffen Eisentraut

INACH



Supported by the Rights, Equality and Citizenship
(REC) Programme of the European Union

Executive Foreword

This publication was written within the framework of the *Research – Report – Remove: Countering Cyber Hate Phenomena* project of the International Network Against Cyber Hate (INACH); funded by the European Commission Directorate-General for Justice and Consumers. The duration of the project is 2016-2017, and its aim is to study, document and report on online hate speech in a comparative and comprehensive way; and to establish structures for a transnational complaints system for instances of cyber hate.

Hate speech is intentional or unintentional public discriminatory and/or defamatory statements; intentional incitement to hatred and/or violence and/or segregation based on a person's or a group's real or perceived race, ethnicity, language, nationality, skin colour, religious beliefs or lack thereof, gender, gender identity, sex, sexual orientation, political beliefs, social status, property, birth, age, mental health, disability, disease.

This report was completed with the participation of the different members of the Network and partners in the project, namely the Zivilcourage und Anti-Rassismus-Arbeit (ZARA) from **Austria**, the Movimiento contra la Intolerancia (MCI) from **Spain**, jugendschutz.net from **Germany**, the Ligue Internationale Contre le Racisme et l'Antisémitisme (LICRA) from **France**, the Inter-Federal Centre For Equal Opportunities and Opposition to Racism from **Belgium** (now called Unia), and the Magenta Foundation from the **Netherlands**; who provided most of the data this report is based upon.

Legal Disclaimer

This publication has been produced with the financial support of the Rights, Equality and Citizenship (REC) Programme of the European Union. The contents of this publication are the sole responsibility of the International Network Against Cyber Hate and can in no way be taken to reflect the views of the European Commission.

Table of content

| | |
|-----------------------------|----|
| 1. Introduction | 3 |
| 2. Twitter | 4 |
| 3. Facebook | 10 |
| 4. YouTube..... | 16 |
| 4. VK.com | 20 |
| 5. Telegram..... | 23 |
| 6. Blogs and Websites | 24 |

1. Introduction

In order to illustrate how cyber hate is dealt with by complaints offices and responsible IT companies, we present ‘typical’ cases in the following. The collection shows how online platforms react to different offences reported by organizations from different countries. All case descriptions include short outlines of the reported content, the communication between complaints office and provider (and other instances, if given) as well as the outcome. The descriptions are presented together with respective national legislations and content guidelines of the platform, if there is one. What can't be shown, however, is a typical case processing for specific complaints (or offences), specific countries or specific IT companies. On the contrary, the collection impressively demonstrates the inconsistency in respect to reporting mechanisms and case handling performed by IT companies.

The following text is structured by different online platforms, which are locations for cyber hate cases, and which by law are responsible for the removal of illegal content after notification. Twitter, Facebook and YouTube were chosen since they are the three biggest social media networks. Additionally we chose VK.com and the mobile instant messenger Telegram in order to include platforms which currently appear as highly problematic when it comes to cyber hate content. Both platforms only respond rarely to complaints and their legal framework is more or less unclear. Finally, blogs and websites – also potential providers of cyber hate – are looked at to complement the case collection.

This case collection could not have been fulfilled without the help of the following organizations who shared their experience:

- Austria: Zivilcourage und Anti-Rassismus-Arbeit (ZARA)
- Belgium: UNIA (Interfederal Centre for Equal Opportunities)
- France: Ligue Internationale Contre le Racisme et l'Antisémitisme (LICRA)
- Germany: jugendschutz.net
- Netherlands: Magenta Foundation
- Spain: Movimiento Contra la Intolerancia (MCI)

2. Twitter

Content guidelines: "Violent threats (direct or indirect): You may not make threats of violence or promote violence, including threatening or promoting terrorism.

[...]

Hateful conduct: You may not promote violence against or directly attack or threaten other people on the basis of race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability, or disease. We also do not allow accounts whose primary purpose is inciting harm towards others on the basis of these categories."

Case in France: Several anti-Semitic, denialist and anti-Muslim tweets on one account

a. What kind of hate speech?

- Anti-Semitic tweet on the Twitter account of the French politician Henry de Lesquen (5.100 followers): "It is the racism of Jews who conduct them to monotheism when they had taken to their gods the "Goyim" [Non-Jews] that they hated."
- Anti-migrant and anti-Muslim tweets on the same account: "I want to stop the immigration-invasion (...) and the abolishment of the right asylum. I would like to develop a great national measure of re-emigration of the extra-European people who don't want to assimilate the French culture".

b. How was the offence reported?

- 17 tweets were reported by using 'report abusive user' form (→ refer to content guidelines).
- Goal: removal of the account.

c. Reaction and further communication?

- Only one tweet removed; account still active and online.
- Licra's next step: contacting the French Twitter contact directly several times.
- Twitter's answer: it is 'borderline'.
- Licra's answer: we made 3 reports to the public prosecutor.

d. Communication with other instances?

- Three reports to the public prosecutor (→ refer to national law):

- One direct proceeding in December 2015.
- One investigation → condemnation to a rehabilitation program in the Holocaust Memorial of Paris.
- One ongoing investigation (for Holocaust denial).

e. What is the outcome?

- Still ongoing process (July 2016): case is taken to court since Twitter's answer is not acceptable.

*French legislation on incitement to hatred and violence: Article 24 of the **Law on the Freedom of the Press of 1881** prohibits anyone from publicly inciting another to discriminate against, or to hate or to harm, a person or a group for belonging or not belonging, in fact or in fancy, to an ethnicity, a nation, a race, a religion, a sex, or a sexual orientation, or for having a handicap. Articles 32 and 33 prohibit anyone from publicly defaming or insulting a person or group for belonging or not belonging, in fact or in fancy, to an ethnicity, a nation, a race, a religion, a sex, or a sexual orientation, or for having a handicap.*

*Holocaust denial: **Act 90-615 to repress acts of racism, anti-Semitism and xenophobia of 1990 (The Gayssot Act)** makes an amendment to Article 24 (Law on the Freedom of the Press of 1881) as follows written: Art. 24 (a). - **those who have disputed the existence of one or more crimes against humanity** such as they are defined by Article 6 of the statute of the international tribunal military annexed in the agreement of London of August 8, 1945 and which were carried out either by the members of an organization declared criminal pursuant to Article 9 of the aforementioned statute, or by a person found guilty such crimes by a French or international jurisdiction [...]*

*Responsibility of ISP: The author of a racist content is not the only one who can be sentenced. Indeed, the **Law on Confidence in the Digital Economy (LCEN)** of 2004 and the Act of 1982 on audiovisual communications have amended the responsibility of the web actors: the ISP and other hosting websites can be considered as responsible of a publication if this one has been duly notified (with a formal notice for promptly removing the publication). Moreover, the LCEN act has defined the obligation for the ISP and hosting websites to provide an accessible and available device for the reporting of hateful contents.*

Case in Austria: Anti-Muslim tweet

a. What kind of hate speech?

- A specific tweet encouraging people to “kill muslims”.

b. How was the offence reported?

- The woman who had witnessed this comment reported a screenshot of this tweet to Twitter herself (→ refer to content guidelines).

c. Reaction and further communication?

- The account is suspended for a certain period of time but is now available again.

d. Communication with other instances?

- The case is forwarded to the reporting centre operated by the Federal Agency for State Protection and Counterterrorism due to a presumed violation of national law (§ 283 Criminal Code – incitement to hatred) (→ refer to national law).
- No feedback (they never give feedback to people who are not party to the possible trial).

e. What is the outcome?

- The reported tweet is removed, the account suspended for a certain period of time but is now available again.

*Austrian legislation on incitement to hatred and violence: Section 283 of the **Austrian Criminal Code** prohibits anyone from publicly inciting to hatred or violence against church communities or religious groups, or a group defined by race, skin color, language, religion, ideology, nationality, origin or national or ethnic descent, sex, handicap, age or sexual orientation or individuals because of their belonging to one of the aforementioned groups. Likewise it prohibits anyone from publicly offending and discriminating the aforementioned groups.*

*Definition of 'public': In 2016 an **amendment to the Austrian Criminal Code** improved the provision on incitement. Previous to this reform, agitation against certain groups had to reach a broad public of no less than 150 people. After the reform was introduced incitement to hatred was redefined. The re-*

quired publicity was reduced to 30 people, and incitement in mainstream media and the internet is now subject to a higher penalty. Dissemination of punishable hateful content (approving of this content) on the internet is now a criminal offence itself.

Case in Belgium: Anti-Muslim profile on Twitter

a. What kind of hate speech?

- Returning Twitter profile, which was filled with anti-Muslim hatred and inciting to violence. After being removed, the author created a new similar page after some time.

b. How was the offence reported?

- UNIA got notice of this returning Twitter profile, they filed a complaint with the authorities about the first page and informed Twitter (→ refer to content guidelines).
- Goal: Because it was a returning profile, UNIA wanted to build a file and hadn't informed the provider.

c. Reaction and further communication?

- Twitter removed the profile.

d. Communication with other instances?

- For the previous page UNIA filed a complaint, but for this new page they were still building the case before filing a new (stronger) complaint (→ refer to national law).

e. Outcome?

- Profile was removed by Twitter after some months without UNIA's intervention.

*Belgian legislation on incitement to hatred and violence: The **Belgian Anti-Racism Law of 1981** prohibits to publicly incite to hatred, violence, segregation or discrimination against a person or a group based on protected criteria (i.e., race, ethnic origin, national origin, nationality, skin colour, ascendance) (article 20), to spread ideas based on race superiority or racial hatred (article 21) and to incite to hatred, violence, segregation or discrimination against a person or a group based on the protected criteria (among others: sexual orientation, religious conviction, and handicap) (article 22). The law furthermore criminalizes to publicly announce the intention to commit any of the aforementioned offences.*

All articles contain the phrase "in the circumstances given in Article 444 of the **Belgian Criminal Code**". Here the definition of 'public' is specified as follows: "either in public meetings or places; or in the presence of several people, in a place that is not public but accessible to a number of people who are entitled to meet or visit there; or in any place in the presence of the offended person and in front of witnesses; or through documents, printed or otherwise, illustrations or symbols that have been displayed, distributed, sold, offered for sale, or publicly exhibited; or finally by documents that have not been made public but which have been sent or communicated to several people."

Another relevant passage is article 442ter of the Criminal Code which refers to harassment, and article 453bis which refers to defamation. In both cases there is a system of aggravating circumstances, if the offences are committed with a biased motivation (racism or any other ground of discrimination).

To sentence someone inciting to hatred, violence, segregation or discrimination, the judge has to prove his malicious/racist intention (so called moral element of the offence). Press related offences based on the anti-discrimination law have to be judged by a jury (the so-called Cour d'Assises), like any other press offence.

*Holocaust denial: The **Belgian Holocaust denial law** makes it illegal to publicly "deny, play down, justify or approve of the genocide committed by the German National Socialist regime during the Second World War". In contrast to offences linked to incitement to hatred and violence, there is no need to prove any malicious/racist intention. Exceptionally, press related offences are judged by a normal criminal court (no jury).*

Case in Germany: Several homophobic and Nazi tweets on one account

a. What kind of hate speech?

- Account includes several tweets that depict Nazi symbols, glorify leading Nazis, and incite to hatred and violence against homosexual people ("good night gay pride", "gas the gays").

b. How was the offence reported?

- Using 'report abusive user' form including three tweets as examples (→ refer to content guidelines)
- Goal: removal of the account.

c. Reaction and further communication?

- The three exemplary tweets were removed, account was still online.
- jugendschutz.net's next step: using the 'law enforcement' form, giving general information about the offences on the account (no separate field in the form for specific examples).
- Twitter's answer: asking to give more specific information on the offences and to send links to the offences.
- jugendschutz.net's answer: sending links of all offences with detailed description and the respective national laws (Section 130 Criminal Code, 'incitement to hatred') (→ refer to content guidelines and to national law)

d. Communication with law enforcement / state authorities

- None.

e. Outcome

- Account is still active, but withheld in Germany.

*German legislation on incitement to hatred and violence: Section 130 of the **German Criminal Code** criminalizes anyone*

"(1) whosoever, in a manner capable of disturbing the public peace 1. incites hatred against a national, racial, religious group or a group defined by their ethnic origins, against segments of the population or individuals because of their belonging to one of the aforementioned groups or segments of the population or calls for violent or arbitrary measures against them; or 2. assaults the human dignity of others by insulting, maliciously maligning an aforementioned group, segments of the population or individuals because of their belonging to one of the aforementioned groups or segments of the population, or defaming segments of the population [...]"

*Youth media protection and hate speech: from a youth protection perspective, taking action against cyber hate aims at protecting young people from potential dangers and disturbances. The relevant provision is the **Interstate Treaty on the Protection of Minors in the Media (JMStV)**, which includes Section 130 of the Criminal Code (see above). jugendschutz.net notifies ISPs and platform operators in Germany and abroad and asks them to remove content violating the laws. Whenever a responsible*

person in Germany can be identified, the case is forwarded to the Commission for the Protection of Minors in the Media who is able to impose sanctions. If the website or service is hosted abroad and no responsible person in Germany can be identified, offences can be forwarded to the Federal Review Board for Media Harmful to Minors. Content can be entered into the list of media harmful to minors, generally referred to as the 'index'. Distributors of an indexed object are no longer permitted to let minors have access to it or to sell, rent out or even present this object in public or to broadcast it. In case of content severely harmful to minors, the object does not need to be entered into the 'index', as the distribution restrictions mentioned above are automatically effective. Strictly prohibited in Germany – even among people over 18 – are incitement to hatred, the denial of the Holocaust, the spreading of propaganda material by forbidden organizations, the depiction of symbols of the National Socialist regime and the glorification of violence.

3. Facebook

Content guidelines: "Facebook removes **hate speech**, which includes content that directly attacks people based on their:

- Race,
- Ethnicity,
- National origin,
- Religious affiliation,
- Sexual orientation,
- Sex, gender, or gender identity, or
- Serious disabilities or diseases.

Organizations and people dedicated to **promoting hatred** against these protected groups are not allowed a presence on Facebook. As with all of our standards, we rely on our community to report this content to us."

Case in the Netherlands: Anti-Semitic comment

a. What kind of hate speech?

- Comment inciting hatred against Jewish people. The Facebook-user in question plainly says that "Jews should be gassed".

b. How was the offence reported?

- Complainant indicates having reported the comment herself, without success.
- Request is sent to Facebook for removal of the comment for being in violation with the FB Statement of Rights and Responsibilities and with the Dutch law (→ refer to content guidelines and to national law).

c. Reaction and further communication?

- Facebook removes comment
- Removal is communicated to complainant

d. Communication with other instances?

- None.

e. Outcome?

- Comment is removed by Facebook.

*Dutch legislation on incitement to hatred and violence: The **Dutch Criminal Code** penalizes*

Art. 137c. [...] who publicly, either orally, or in writing, or by image, intentionally makes a defamatory statement about a group of persons on the grounds of their race, religion or personal beliefs, their hetero- or homosexual orientation or their physical, mental or intellectual disability [...]

Art. 137d. [...] who publicly, either orally or in writing or by image, incites hatred of or discrimination against persons or violence against their person or property, on the grounds of their race, religion or personal beliefs, their sex, their hetero- or homosexual orientation or their physical, mental or intellectual disability [...]

Art. 137e. 1. [...] who, for any reason other than giving factual information: makes public a statement which he knows of or should reasonably suspect to be offensive to a group of persons on the grounds of their race, religion, or personal beliefs, their hetero- or homosexual orientation or their physical, mental or intellectual disability, or incites hatred of or discrimination against people or violence against their person or property on the grounds of race, religion or personal beliefs, their sex, their hetero- or homo-

sexual orientation or their physical, mental or intellectual disability; 2. sends or distributes an object which he knows or should reasonably suspect to contain such defamatory statement or has such in stock for public disclosure or for dissemination [...]

Case in the Netherlands: Denialist post

a. What kind of hate speech?

- Post on the Facebook page of the Muslim Defence League Holland describing the Holocaust as the ‘‘biggest lie in history’’.

b. How was the offence reported?

- Request is sent to Facebook for removal of the post for being in violation with the FB Statement of Rights and Responsibilities and with the Dutch law (→ refer to content guidelines and to national law).

c. Reaction and further communication?

- Facebook removes user
- Removal is communicated to complainant

d. Communication with other instances?

- None.

e. Outcome?

- Entire account is removed by Facebook for making ‘‘real threats or statements of hate’’.

*Holocaust denial: There is no specific law on Holocaust Denial in the Netherlands, but the Supreme Court ruled in 1995 that Article 137 of the **Criminal Code** (see above), which makes illegal the defamation of groups based on their religion or ethnicity, also bans Holocaust denial.*

Case in Belgium: Several anti-Muslim comments on one post

a. What kind of hate speech?

- Vlaams Belang, a far right Belgian party, posted a picture on their Facebook page stating that 69,6% of the Muslims in Belgium think Islamic rules are above the law.
- Several reactions on this post instigate violence based on one of the protected criteria – faith or personal beliefs – in the Belgian anti-discrimination law.

b. How was the offence reported?

- Comments are reported through the complaint form that is implemented by Facebook (→ refer to content guidelines).

c. Reaction and further communication?

- For each of these comments Facebook sent the following answer "We reviewed the comment you reported for containing hate speech or symbols and found it **doesn't violate** our Community Standards."
- Unia sent an extensive e-mail to Facebook's Policy Casework specifying this case and asking for removal of the content.

d. Communication with other instances?

- None.

e. Outcome?

- The worst comments got removed by Facebook at least a week after Unia sent the e-mail to their Policy Casework. Considering the momentariness of social media, the post had already lost its relevance by this time.

*Belgian legislation on incitement to hatred and violence: The **Belgian Anti-Racism Law of 1981** prohibits to publicly incite to hatred, violence, segregation or discrimination against a person or a group based on protected criteria (i.e., race, ethnic origin, national origin, nationality, skin colour, ascendance) (article 20), to spread ideas based on race superiority or racial hatred (article 21) and to incite to hatred, violence, segregation or discrimination against a person or a group based on the protected criteria (among others: sexual orientation, religious conviction, and handicap) (article 22). The law furthermore criminalizes to publicly announce the intention to commit any of the aforementioned offences.*

Case in France: Several anti-Migrant and anti-Muslim posts on one account

a. What kind of hate speech?

- Description of the Facebook page: "Patriots of France come together under the tricolor banner and around the French values... They affirm their commitment to their homeland".
- 113.000 followers.
- On this page, there is an affirmation of patriotic values, defense of police and army, of the Catholic and western values. Muslims and Migrants are used as the scapegoat for the French societal issues. Islam is totally linked with terrorism.
- Example 1: "And there, we can observe the loving, grateful and peaceful behavior of migrants in France. And we are the horrible racists without any heart. The night of the 14th in Paris was like civil war! Share a lot".
- Example 2: Photo of a Muslim woman wearing a tricolor burka with a message "the great replacement".
- Example 3: "Breton kebab: 100% pork!"



b. How was the offence reported?

- Using Facebook report three times (→ refer to content guidelines).
- Goal: removal of the posts.

c. Reaction and further communication?

- Facebook removed the photo of the Muslim woman wearing a tricolor burka.
- Other posts are not removed. Reason: French legislation (see below) is not enough developed for anti-Migrant hate.

d. Communication with other instances?

- None.

e. Outcome?

- One post removed, other hateful posts still online.

*French legislation on incitement to hatred and violence: Article 24 of **the Law on the Freedom of the Press of 1881** prohibits anyone from publicly inciting another to discriminate against, or to hate or to harm, a person or a group for belonging or not belonging, in fact or in fancy, to an ethnicity, a nation, a race, a religion, a sex, or a sexual orientation, or for having a handicap. Articles 32 and 33 prohibit anyone from publicly defaming or insulting a person or group for belonging or not belonging, in fact or in fancy, to an ethnicity, a nation, a race, a religion, a sex, or a sexual orientation, or for having a handicap.*

Case in Austria: Several anti-refugee comments on one post

a. What kind of hate speech?

- The Facebook discussion is started in the context of a video where refugees are beaten by the Italian police. Hundreds of users consequently express their hateful and violent thoughts about refugees and ‘foreigners’ in general.

b. How was the offence reported?

- Almost 100 postings are reported to Facebook by labelling them as ‘hate speech’ (→ refer to content guidelines)

c. Reaction and further communication?

- 5 out of the approximately 100 reported postings are removed by Facebook. This seems to be a random choice.
- Therefore we got in touch with our Facebook contact person in Berlin via e-mail. In her answer she said that this case cannot be examined more thoroughly because the thread has been deleted. As we checked whether this was true, we found out that the thread was still online. Hence we contacted the contact person again but have not received a further reply so far.

d. Communication with other instances?

- The thread is reported to the public prosecutor's office by means of a written statement of the facts including several exemplary postings due to the presumed violation of national law (§ 283 Criminal Code – incitement to hatred). (→ refer to national law)
- No feedback (they never give feedback to people who are not party to the possible trial).

e. Outcome?

- 5 random postings are removed, the rest is still online.

*Austrian legislation on incitement to hatred and violence: Section 283 of the **Austrian Criminal Code** prohibits anyone from publicly inciting to hatred or violence against church communities or religious groups, or a group defined by race, skin color, language, religion, ideology, nationality, origin or national or ethnic descent, sex, handicap, age or sexual orientation or individuals because of their belonging to one of the aforementioned groups. Likewise it prohibits anyone from publicly offending and discriminating the aforementioned groups.*

Protected groups: The 2016 amendment of the Austrian Criminal Code criminalizes inciting to hatred or violence against groups or individuals defined by certain (lacking) attributes like ethnic or national origin. Therefore 'foreigners', 'migrants', 'refugees' or 'asylum seekers' are now protected groups.

4. YouTube

Content guidelines: "We encourage free speech and try to defend your right to express unpopular points of view, but we don't permit hate speech.

Hate speech refers to content that promotes violence or hatred against individuals or groups based on certain attributes, such as:

- race or ethnic origin
- religion
- disability
- gender
- age

- veteran status

- sexual orientation/gender identity

There is a fine line between what is and what is not considered to be hate speech. For instance, it is generally okay to criticize a nation-state, but not okay to post malicious hateful comments about a group of people solely based on their race."

Case in France: Several denialist and Nazi videos on two channels

a. What kind of hate speech?

- On the two YouTube channels of Vincent Reynouard, a French well-known Holocaust denialist, he posts videos of him talking about the false affirmations of the Holocaust (ex: "Treblinka 'material proofs': outright lie and overbearing" and "My challenge: gas me") and advocating the Nazi party (ex: "the National-socialist hope").
- 7,700 followers and 2.1 million views.

b. How was the offence reported?

- Using "Trusted flag report" (→ refer to content guidelines and national law)
- Goal: removal of both channels.

c. Reaction and further communication?

- Videos are removed.
- One channel is removed.

d. Communication with other instances?

- None.

e. Outcome?

- Videos and one channel removed
- Problem: videos are always published by new channels (Meta TV for example), in other countries countering French legislation (no specific laws on Holocaust denial).

Holocaust denial: Act 90-615 to repress acts of racism, anti-Semitism and xenophobia of 1990 (The Gayssot Act) makes an amendment to Article 24 (Law on the Freedom of the Press of 1881) as follows: Art. 24 (a). - those who have disputed the existence of one or more crimes against humanity such as they are defined by Article 6 of the statute of the international tribunal military annexed in the agreement of London of August 8, 1945 and which were carried out either by the members of an organization declared criminal pursuant to Article 9 of the aforementioned statute, or by a person found guilty such crimes by a French or international jurisdiction [...]

Case in Austria: Anti-refugee video

a. What kind of hate speech?

- The YouTube user 'mart hart' frequently posts videos with racist and pro-right-wing-politics content. A specific video spreads hate against refugees.

b. How was the offence reported?

- The respective video was reported to ZARA's counseling unit. This video is reported to YouTube (→ refer to content guidelines).

c. Reaction and further communication?

- The video is removed, the account stays online.

d. Communication with other instances?

- None (no national laws violated).

e. Outcome?

- The video is removed, the account stays online.

*Austrian legislation on incitement to hatred and violence: Section 283 of the **Austrian Criminal Code** prohibits anyone from publicly inciting to hatred or violence against church communities or religious groups, or a group defined by race, skin color, language, religion, ideology, nationality, origin or national or ethnic descent, sex, handicap, age or sexual orientation or individuals because of their belonging to one of the aforementioned groups. Likewise it prohibits anyone from publicly offending and discriminating the aforementioned groups.*

*Protected groups: The 2016 **amendment of the Austrian Criminal Code** criminalizes inciting to hatred or violence against groups or individuals defined by certain (lacking) attributes like ethnic or national origin. Therefore "foreigners", "migrants", "refugees" or "asylum seekers" are now protected groups.*

Case in Germany: Denialist and Nazi videos on one channel

a. What kind of hate speech?

- Channel is used to spread Holocaust denying videos ("I know why they did not know about things that happened. Because that things (Holocaust) didn't happen. So how could they knew?"), videos with forbidden Nazi symbols and music videos that are glorifying National Socialism ("Dear Uncle Adolf", "Adolf Hitler - nothing compares to you").

b. How was the offence reported?

- Using trusted flagging (→ refer to content guidelines and to national law)

c. Reaction and further communication?

- The channel is not available anymore in Germany.

d. Communication with other instances?

- None.

e. Outcome?

- The channel is not available anymore in Germany.
- Problem: Channel is still available in other countries countering German legislation (no specific laws on Holocaust denial and display of forbidden symbols).

*Holocaust denial and glorification of NS regime: Section 130 of the **German Criminal Code** criminalizes anyone*

"(3) Whosoever publicly or in a meeting approves of, denies or downplays an act committed under the rule of National Socialism of the kind indicated in section 6 (1) of the Code of International Criminal Law, in a manner capable of disturbing the public peace [...]

(4) Whosoever publicly or in a meeting disturbs the public peace in a manner that violates the dignity of

the victims by approving of, glorifying, or justifying National Socialist rule of arbitrary force [...]"

4. VK.com

Content guidelines: "When using the Site the User is prohibited from loading, storing, publishing, disseminating, making available or otherwise using any information which:

(a) contains threats, discredits, insults, defiles honor, dignity or business reputation or violates the privacy of other Users or third parties; [...]

(f) propagandizes and/or contributes to racial, religious, ethnic hatred or hostility, propagandizes fascism or racial superiority;

(g) contains extremist materials;

(h) propagandizes criminal activity or contains advices, instructions or guides for criminal activities [...]"

Case in Germany: Two Nazi profiles on VK.com

a. What kind of hate speech?

- Glorification of the Nazi regime, depiction of prohibited symbols like swastika, anti-Muslim posts

b. How was the offence reported?

- Report sent via the standard form to VK support: "I stumbled upon those profiles which clearly violate the terms of our community. According to terms (6.3.4) it is prohibited to post content that (f) propagandizes and/or contributes to racial, religious, ethnic hatred or hostility, propagandizes fascism or racial superiority or (g) contains extremist materials. [...] So please see 1) [name of profile] This profile is full of national socialist glorification and propagandizes fascism. 2) [name of profile] This is a clear violation of ToS! Please keep me informed about the taken measures to prevent further misuse of our community by hate mongers! THX!" (→ refer to content guidelines)

c. Reaction and further communication?

- VK's answer: "I can understand, why you are so worried about this situation. But we should give everyone their right to free speech. If you are sure these people are offenders – please consult the police or lawyers. According to the presumption of innocence it's the law enforcement who can (and must) estimate are they're breaking the law or not."
- jugendschutz.net's answer: "All I ask for is compliance with your own and self-given Terms which say is prohibited to 'propagandizes and/or contributes to racial, religious, ethnic hatred or hostility, propagandizes fascism or racial superiority'

Both profiles do exactly this. They violate your Terms of Service. So please take your own rules serious and take action against that stuff. That's what users expect from you: to act according Terms we have accepted by becoming part of this community.

PS: Freedom of speech cannot be an excuse for accepting inhumanity!"

- VK's answer: "The terms allow us to block somebody but not oblige us to do it. Such question is contradictory, that is why we talk about the police."

d. Communication with other instances?

- None.

e. Outcome?

- Both profiles are not removed.

Case in Germany: Anti-migrant post on VK.com

a. What kind of hate speech?

- Picture of a real built up sniper rifle with a comment saying 'Armed. Project Kanacken (pejorative word for foreigners) prevention has started'.
- Other posts imply that the user is actually in possession of arms and that he is also going to use them; some posts show pictures of Muslim shops and left-wing locations that could be possible targets.

b. How was the offence reported?

- Since the picture of the rifle is alarming and the user potentially announces a crime, jugendschutz.net didn't report to VK but to the Federal Criminal Police Office directly (see below) (→ refer to national law).

c. Reaction and further communication?

- No communication with VK.

d. Communication with other instances?

- The finding was reported to the Federal Criminal Police Office, who sent the file number in reply.

e. Outcome?

- Content is not removed. Two days after reporting, the responsible user posts that he had a visit from the police and that he will leave VK (after already leaving Facebook) since the platform is not safe anymore.

*Russian legislation on incitement to hatred and violence: Since VK is hosted in Russia and apparently does not commit to any national or European laws, a look at Russian legislation is useful. The relevant provision is **Federal Law No. 114 FZ on Counteraction of Extremist Activities of 2002**. Originally a preventive measure against terrorism, it has been used for tackling online hate speech. It criminalizes*

"1) the activity of public and religious associations or any other organisations, or of mass media, or natural persons to plan, organise, prepare and perform the acts aimed at

[...]

- the excitation of racial, national or religious strife, and also social hatred associated with violence or calls for violence;

- the abasement of national dignity;

- the making of mass disturbances, ruffian-like acts, and acts of vandalism for the reasons of ideological, political, racial, national or religious hatred or hostility toward any social group;

- the propaganda of the exclusiveness, superiority or deficiency of individuals on the basis of their attitude to religion, social, racial, national, religious or linguistic identity;

2) the propaganda and public show of nazi attributes or symbolics or the attributes or symbolism similar to nazi attributes or symbolics to the extent of blending;

3) public calls for the said activity or for the performance of the said acts [...]"

Responsibility of ISP: Roskomnadzor is the federal executive body responsible for control and supervision of the media in Russia. "For the purposes of restricting access to the Internet websites containing extremist information, the Russian legislature has established a Unified Automated Information System ('UAIS'). This system allows for the identification of Internet sites containing such information. The Roskomnadzor creates, formats, and maintains the Register: the blacklist. The Register includes domain names, pages indicators of such Internet sites, and their IP addresses that include prohibited information. All information is included in the Register based on the court rulings that have labeled information as extremist. Another ground for including such information into the Register, as previously mentioned, is a request by the General Attorney, sent to the Roskomnadzor. In order to gain access to the list of domain names and IP addresses, a connections services provider is obligated to login to the Roskomnadzor's web service every day. The list is updated twice per day. A matching system identifies websites with extremist information. Within 24 hours after update and identification, a communications service provider is required to restrict access to the prohibited Internet websites if it is found in the blacklist; otherwise, the operator will be liable to the Roskomnadzor, which can withdraw the operator's license. This blocking system works automatically." (Dobryakova 2016)

5. Telegram

No specific rules on hate speech. The possibility of taking down illegal content is described as follows:

"Our mission is to provide a secure means of communication that work everywhere on the planet. To do that in the places where it is most needed (and to continue distributing Telegram through the App Store and Google Play), we have to process legitimate requests to take down illegal public content (sticker sets, bots, and channels) within the app. For example, we can take down sticker sets that violate intellectual property rights or porn bots in countries where pornography is illegal.

User-uploaded stickers sets, channels, and bots by third-party developers are not part of the core Telegram UI. Whenever we receive a complaint at abuse@telegram.org or dmca@telegram.org regarding the legality of public content, we perform the necessary legal checks and take it down when deemed appropriate.

*Please note that this does **not** apply to local restrictions on freedom of speech. For example, if criticizing the government is illegal in some country, Telegram won't be a part of such politically motivated censorship. This goes against our founders' principles. While we do block terrorist (e.g. ISIS-related)*

bots and channels, we will not block anybody who peacefully expresses alternative opinions."

Case in Germany: Hate against members of other faiths on Telegram

a. What kind of hate speech?

- Anti-Shia posts on the Jihadist Telegram Channel 'Garib fi Dunya', for example a picture of a shia-ritual besides a picture of pigs: "Congrats, you're not the descendent of the apes but far worse than any animal." Or "I #hate shias" postings
- Jihadist hate against 'non-believers', in this Channel specifically against Shia's (clashes between Islam's two big sects, the Sunni and the Shia)

b. How was the offence reported?

- Channel was reported by first using the 'report' button on the Telegram App and then, after no reaction, by sending a report via e-mail to abuse@telegram.org (→ refer to content guidelines).
- Goal: removal of the account.

c. Reaction and further communication?

- No reaction.

d. Communication with other instances?

- None.

e. Outcome?

- Channel still online.

6. Blogs and Websites

Case in Spain: Several types of hatred on Blogspot site (resistenciasinlider.blogspot.com)

a. What kind of hate speech?

- Website disseminating offensive material against migrants, Muslims, Jews, LGTB community, homeless and other vulnerable groups. Furthermore, many articles and

opinions defending Aryan race supremacy and incitement to hatred against those collectives as well as promoting ideas whose purpose is rehabilitation of regimes that have committed genocide.

b. How was the offence reported?

- Report to the Special Prosecutor on Hate Crimes of Barcelona (→ refer to national law).
- Demand at court.

c. Reaction and further communication?

- Juridical Process.

d. Communication with other instances?

- Police.

e. Outcome?

- Website was removed.
- Administrator was found guilty of crimes against fundamental rights and public freedom by inciting to hatred and discrimination; and promoting ideas aiming to rehabilitation of regimes that have committed genocide. The perpetrator was condemned to one year in jail and a fine of 6 euros per day during a year.
- No charges have been imposed against Blogspot.

*Spanish legislation on incitement to hatred and violence: **Article 510 of the Spanish Criminal Code** penalizes*

"Any person who publicly encourages, promotes or incites directly or indirectly hatred, hostility, discrimination or violence against a group or part thereof, or against any person by reason of their membership of that group, for reasons of racism, anti-Semitism or other such ideology, religion or beliefs, family situation, membership of an ethnic group, race or nation, national origin, gender, sexual orientation, gender, illness or disability" [...]

"Any person who produces, develops, possesses for the purpose of distribution, provides access to third parties, distributes or sells written or any other kind of material or medium whose content is able to encourage, promote, or incites hatred directly or indirectly against [repeats categories of the previous paragraph]"

*Holocaust denial: The amended **Criminal Code** includes several provisions on Holocaust Denial: "Any person who publicly denies or trivializes or seriously exalts any crimes of genocide, crimes against humanity or against persons and protected in armed or exalts the perpetrators. Any person who infringes the dignity of the person through actions involving humiliation, contempt or damage to the reputation of any of the groups to which the preceding paragraph refers. Any person who exalts or justifies through any published medium or broadcast crimes that had been committed against [repeats the same types of groups as above]"*

Article 607.2: *"The diffusion by any means of ideas or doctrines that deny or justify the crimes in the previous section of this article, or tries the rehabilitation of regimes or institutions which they protect generating practices of such, will be punished with a prison sentence of one to two years."*

Case in France: Racist band website (<https://lecomitedesalutpublic.com/>)

a. What kind of hate speech?

- ‘Les Brigandes’ is an ultra-conservative band made up of seven young women. The music group – whose name pays tribute to the Vendean royalist insurgents – spread a fundamentalist, conspiracy message that is fiercely hostile to any kind of difference. A message in which a white, ultra-catholic, Putin-loving France is idealized. (more info on the group: <http://www.getthetrollsout.org/what-we-do/articles/item/64-are-%E2%80%98les-brigandes%E2%80%99-the-new-stars-of-the-fascist-scene.html>).
- Website hosted on orangewebsite.com (France).
- Additionally the band has a YouTube channel: 2,800 followers and 1.3 million views. (<https://www.youtube.com/channel/UCQsPTAuUF2EMe8y7gnNtQWA>)
- 50,000-100,000 views for most of their songs.
- The band is expanding to other countries: Many videos on YouTube are available with English, German and Russian subtitles.

b. How was the offence reported?

- Report to the online police PHAROS (→ refer to national law).
- Report to pointdecontact.net (→ refer to content guidelines).

c. Reaction and further communication?

- No answer.

d. Communication with other instances?

- Police.

e. Outcome?

- Website and YouTube channel still online.

*French legislation on incitement to hatred and violence: Article 24 of **the Law on the Freedom of the Press of 1881** prohibits anyone from publicly inciting another to discriminate against, or to hate or to harm, a person or a group for belonging or not belonging, in fact or in fancy, to an ethnicity, a nation, a race, a religion, a sex, or a sexual orientation, or for having a handicap. Articles 32 and 33 prohibit anyone from publicly defaming or insulting a person or group for belonging or not belonging, in fact or in fancy, to an ethnicity, a nation, a race, a religion, a sex, or a sexual orientation, or for having a handicap.*

*Responsibility of ISP: The author of racist content is not the only one who can be sentenced. Indeed, the **Law on Confidence in the Digital Economy (LCEN)** of 2004 and the Act of 1982 on audiovisual communications have amended the responsibility of the web actors: The ISP and other hosting websites can be considered as responsible for a publication if they have been duly notified (with a formal notice for promptly removing the publication). Moreover, the LCEN act has defined the obligation for the ISP and hosting websites to provide an accessible and available device for reporting hateful content.*

Case in the Netherlands: Anti-Muslim comment on WordPress site

a. What kind of hate speech?

- Comment inciting hatred and violence, as it calls for ‘violent mass deportation of all Muslims’.

b. How was the offence reported?

- Request is sent to administrator of the site with a request for removal of the above-mentioned expression.

c. Reaction and further communication?

- Comment is removed by administrator.
- Removal is communicated to complainant.

d. Communication with other instances?

- None.

e. Outcome?

- Comment removed.

Wordpress content guidelines: Wordpress emphasizes the responsibility of the users: "If you operate a blog, comment on a blog, post material to WordPress.com, post links on WordPress.com, or otherwise make (or allow any third party to make) material available (any such material, 'Content'), you are entirely responsible for the content of, and any harm resulting from, that Content or your conduct. That is the case regardless of what form the Content takes, which includes, but is not limited to text, photo, video, audio, or code. By using WordPress.com, you represent and warrant that your Content and conduct do not violate these terms or the User Guidelines."

User Guidelines: "We don't tolerate direct and realistic threats of violence. (This doesn't mean that we'll remove all hyperbole or offensive language.)"

*Dutch legislation on incitement to hatred and violence: The **Dutch Criminal Code** penalizes*

Art. 137c. [...] who publicly, either orally, or in writing, or by image, intentionally makes a defamatory statement about a group of persons on the grounds of their race, religion or personal beliefs, their hetero- or homosexual orientation or their physical, mental or intellectual disability [...]

Art. 137d. [...] who publicly, either orally or in writing or by image, incites hatred of or discrimination against persons or violence against their person or property, on the grounds of their race, religion or personal beliefs, their sex, their hetero- or homosexual orientation or their physical, mental or intellectual disability [...]

Art. 137e. 1. [...] who, for any reason other than giving factual information: makes public a statement which he knows of or should reasonably suspect to be offensive to a group of persons on the grounds of their race, religion, or personal beliefs, their hetero- or homosexual orientation or their physical, mental

or intellectual disability, or incites hatred of or discrimination against people or violence against their person or property on the grounds of race, religion or personal beliefs, their sex, their hetero- or homosexual orientation or their physical, mental or intellectual disability; 2. sends or distributes an object which he knows or should reasonably suspect to contain such defamatory statement or has such in stock for public disclosure or for dissemination [...]

Case in Germany: Denialist/Nazi WordPress site

a. What kind of hate speech?

- Depicting Nazi symbols (Section 86a of the Criminal Code), glorifying war (Section 4 (1)7 of the Interstate Treaty on the Protection of Minors in the Media), denial of holocaust (Section 130.3 of the Criminal Code)

b. How was the offence reported?

- Several reports to WordPress via complaint form. (→ refer to content guidelines)

c. Reaction and further communication?

- Wordpress' answer: "Thanks for getting in touch. While WordPress.com does not condone the opinions expressed on the site in question, we will not remove material on the grounds that it is offensive or disturbing to its readers. Please refer to our position on freedom of speech at the link below: <http://en.wordpress.com/freedom-of-speech/>. If you feel strongly about this particular issue, please consider blogging about it. You will, of course, have the same freedoms."

d. Communication with other instances?

- Forwarding the case to the Federal Review Board for Media Harmful to Minors (BPjM).
- No feedback (they never give feedback, success control only possible through controlling if content is still available in search engines).

e. Outcome?

- Website still online, but on the list of media harmful to minors (the 'index').

Wordpress content guidelines: Wordpress emphasizes the responsibility of the users: "If you operate a

blog, comment on a blog, post material to WordPress.com, post links on WordPress.com, or otherwise make (or allow any third party to make) material available (any such material, 'Content'), you are entirely responsible for the content of, and any harm resulting from, that Content or your conduct. That is the case regardless of what form the Content takes, which includes, but is not limited to text, photo, video, audio, or code. By using WordPress.com, you represent and warrant that your Content and conduct do not violate these terms or the User Guidelines."

User Guidelines: "We don't tolerate direct and realistic threats of violence. (This doesn't mean that we'll remove all hyperbole or offensive language.)"

*German legislation on incitement to hatred and violence: Section 130 of the **German Criminal Code** criminalizes anyone*

"(1) whosoever, in a manner capable of disturbing the public peace 1. incites hatred against a national, racial, religious group or a group defined by their ethnic origins, against segments of the population or individuals because of their belonging to one of the aforementioned groups or segments of the population or calls for violent or arbitrary measures against them; or 2. assaults the human dignity of others by insulting, maliciously maligning an aforementioned group, segments of the population or individuals because of their belonging to one of the aforementioned groups or segments of the population, or defaming segments of the population [...]"

Case in Belgium: Anti-migrant and anti-Semitic texts on a blog (skynet blog)

a. What kind of hate speech?

- Several texts on a blog were targeting migrants as being invaders and proposed violent attacks against them as the solution “we have the rights to defend ourselves with weapons if authorities are not doing it”. Other publications of the same author were targeting Jews.

b. How was the offence reported?

- UNIA contacted the provider by mentioning 4 publications that might violate their code of conduct. UNIA also insisted on the fact that these publications could violate Belgian laws by transferring two cases of jurisprudence to them (→ refer to content guidelines and national law).

c. Reaction and further communication?

- The provider advised UNIA of their action and methods of work in a standard email. They insisted on the fact that they are removing content when it is illegal, that they do so when they have effective knowledge of the content and when the complaints are sufficiently precise.
- They also informed of what they did in this specific case: First, they asked the author to modify the incriminated texts. As he did not react after 8 days, second, his blog has been put offline.

d. Communication with other instances?

- None.

e. Outcome?

- Blog is removed.

*Belgian legislation on incitement to hatred and violence: The **Belgian Anti-Racism Law of 1981** prohibits to publicly incite to hatred, violence, segregation or discrimination against a person or a group based on protected criteria (i.e., race, ethnic origin, national origin, nationality, skin color, ascendance) (Article 20), to spread ideas based on race superiority or racial hatred (Article 21) and to incite to hatred, violence, segregation or discrimination against a person or a group based on the protected criteria (among others: sexual orientation, religious conviction, and handicap) (Article 22). The law furthermore criminalizes publicly announcing the intention to commit any of the aforementioned offences.*

Case in Austria: Antiziganist T-shirt on an online selling platform (zazzle.de)

a. What kind of hate speech?

- An online selling platform sells an antiziganist T-Shirt saying ‘Gypsy Hunter Badge – We Kill the Gypsies’.

b. How was the offence reported?

- Several private individuals and a Roma and Sinti organization have already reported the hateful T-Shirt; therefore an additional report from ZARA is not necessary.

c. Reaction and further communication?

- Due to several reports the selling platform removes the T-Shirt from their list of products very quickly.

d. Communication with other instances?

- The profile is reported to the public prosecutor's office by means of a written statement of the facts due to the presumed violation of national laws (Section 283 of the Criminal Code – incitement to hatred, Section 107 of the Criminal Code – dangerous, criminal threats and Section 282 of the Criminal Code – encouragement to committing and approving of a crime).
- No feedback (they never give feedback to people who are not party to the possible trial)

e. Outcome?

- The discriminatory T-shirt and other products offered by the same supplier are no longer offered on the respective online selling platform.

*Austrian legislation on incitement to hatred and violence: Section 283 of the **Austrian Criminal Code** prohibits anyone from publicly inciting to hatred or violence against church communities or religious groups, or a group defined by race, skin color, language, religion, ideology, nationality, origin or national or ethnic descent, sex, handicap, age or sexual orientation or individuals because of their belonging to one of the aforementioned groups. Likewise it prohibits anyone from publicly offending and discriminating the aforementioned groups.*