



# **COMBATING ONLINE HATE SPEECH AGAINST RELIGIOUS GROUPS AND VARIOUS VULNERABLE MINORITY GROUPS**

COLLECTION OF ESSAYS





## **CONTENTS**

### **FOREWORD**

### **INTRODUCTION**

Supporting and involving religious actors in combating  
hate speech in Hungary

*National Democratic Institute (NDI)*

### **ESSAYS**

The use of social media comes with responsibility

*Ágnes Katalin Kelemen, social historian*

Combating online hate speech against religious  
and various vulnerable minority groups

*Gerzson Dávid Boros, data scientist*

Words are a dangerous weapon and some people are unarmed

*dr. Ágnes Zentai, lawyer*





## FOREWORD

Dear Reader,

The publication you are holding in your hands is quite special, as it was born from the experience of a unique programme, an extraordinary attempt to bring together actors who rarely sit together in such diversity and understanding. We interviewed representatives of religious and minority groups on how they are affected by hate speech on the Internet, what their experiences are in this respect and how we can combat this growing phenomenon. All of the interviewees agreed that they are bothered by the many attacks that members of their communities constantly come under in the online space, and they also feel that minorities are vulnerable, especially to political attacks, and that most of the time no one supports them—neither the authorities, nor other people, nor the “platforms” offer protection against the abusers.

After a year of working together, we have gained a lot of experience, which we share in this publication. The three authors have tried to show different sides of hate speech—a truly multifaceted subject—which increasingly seems to become part of our lives in the long run. I hope these essays provide the readers with some good lessons, and they will follow even our future work on this topic and become an active member of our lively community that fights against hate speech.

Bálint Jósá, editor of the publication





## INTRODUCTION

### **Engaging and Supporting Faith-Based Actors to Combat Hate Speech in Hungary**

As a non-profit, nonpartisan organisation working to support and strengthen democratic institutions worldwide, the National Democratic Institute (NDI) works to further citizen participation, openness and accountability in government and in society. As such, NDI believes firmly that a democratic space unimpeded by hate speech and disinformation gives everyone a better chance to further their interests in inclusive ways, without fear of threat or retaliation for their beliefs or identities. In a wider sense, NDI's work is guided by the principles enshrined in the Universal Declaration of Human Rights and that democratic governance affirms basic human dignity and fundamental freedoms. NDI emphasises making democracy work for all people, and this includes those marginalized because of their religious beliefs or non-beliefs. To this end, NDI programs have been assisting these communities in peacefully claiming their rights, bridging social divides and building a more inclusive political culture. To further this goal, NDI programs have also engaged faith-based organisations in citizen-led activism, including broad-based civic education, community organising and social accountability.

The world of religion, belief, and non-belief is diverse and trends indicate a growing number of people associate with some form of faith. According to a [Pew Research Center demographic study](#), 84 percent of the global population follows a religion and that number is set to increase to 87 percent by 2050. This cuts across the communities where NDI operates. While a significant number of people follow the most practiced religions in the world, such as Christianity, Islam, or Judaism, others follow lesser-known faiths, Indigenous religions, or other belief systems, like Humanism. In NDI's experience, religion can be used to bring people together and foster peace, or misused to promote intolerance and discrimination.



Faith-based organisations and leaders play important roles in their communities. These leaders and organisations are typically trusted among followers and have deeply rooted relationships that often place them at the centre of associational life. They can also access and mobilize networks that span countries, regions, and the globe -- putting faith-based leaders and organisations in unique positions to tackle social and democratic development challenges.

## **Advancing Freedom of Religion or Belief**

Freedom of religion or belief (FoRB) encompasses not just the freedom to hold personal thoughts or convictions, but also to be able to manifest them individually or with others, publicly and privately, without fear of reprisal. It forbids discrimination against individuals who have different beliefs and coercion to make someone hold or change their religion or belief. To promote and protect FoRB, means working alongside faith-based organisations and actors, as well as other human rights defenders, to challenge FoRB violations, such as blasphemy laws, harassment and detention of FoRB activists, hate speech and disinformation, forced conversion, and protection of sacred sites. It means elevating FoRB violations alongside other human rights violations and supporting defenders and activists in countering increasing levels of faith- and religious-based restrictions and attacks on faith groups.

In the Czech Republic, Slovakia, Hungary, and Poland, NDI has worked since 2016 with civic groups and religious leaders to mobilize communities against hate speech and religious and ethnic discrimination. This included organising interfaith dialogues to build partnerships and small grants to support interfaith coalitions undertaking campaigns to educate communities about discrimination and stereotypes, expose young people to new perspectives about religious tolerance, and correct misconceptions that compound fear and intolerance of other religions.

This is why NDI has been proud to partner with and support the work of an innovative coalition of faith-based and civil society organisations in Hungary whose aim is to combat hate and hateful speech online and to engage in dialogue and advocacy on widening the understanding of the public regarding the corrosive effects of hateful speech. Formed in late 2020, the coalition, comprised of RGDS (civil society), *Organization of*



*Muslims in Hungary (Muslim), Association of Hungarian Women Theologians (interfaith), Szim Salom (Jewish), Ambedkar School (Buddhist Roma), and Eszter Haz Association (Jewish)* has been successfully raising awareness of online hate speech among religious and ethnic minority groups and providing advocacy groups with tools to effectively respond.

The work of this group has included: monitoring and analysing hate speech online, in particular that targeted against Muslim, Jewish and Roma people; organising and delivering awareness raising workshops for young people targeted by such hate speech; implementing an online campaign designed to raise awareness of the negative impacts of hate speech and to share counter narratives promoting positive speech.

It is my pleasure to introduce the present study which collects and frames the learning of this group of dedicated individuals and organisations; and advocates for the extension of the concept of hate speech from the strictly legal definition, to encompass hurtful speech online, that can too easily spill over to real life if allowed to fester. The authors examine different sides of this question in their contributions.

Ágnes Kelemen (short few words about her) highlights the responsibility of social media consumers to engage with the effects of their statements online and to critically examine whether these can advertently or inadvertently lead to the discrimination, exclusion and attacks of particular groups. In particular, she draws attention to the important role of organisations, political actors and businesses monitoring their own and their members' online posts. Too often, a lack of engagement with this issue, rather than bad faith, are to blame for the proliferation of racist, anti-Muslim, antisemitic and sexist content. Educating the public and taking the time to notify social media providers such as Facebook about these comments is an arduous but important job that we should all do.

Taking up this thread, Gerzson Boros from the Organisation of Muslims in Hungary, describes how a group of determined activists made it their business to understand and root out hateful speech posted online in Hungarian. Through the NDI-supported project to combat hate speech, a group of like-minded NGOs and individuals established a cooperation in which they "joined forces" to try to do something together to make online hate a little less prevalent in Hungary. Working with a hate speech



monitoring platform established by the Netherlands-based organisation [Textgain](#), they painstakingly followed and reported instances of hate speech, as well as learning along the way what hateful discourses these resulted in that then infiltrated not just online but also mainstream media in Hungary. It was their work that led to the campaign mentioned above, created and distributed by young people who stood up against the hate speech they have experienced online.

The coalition worked throughout with highly qualified media and legal specialists. Among the latter, Ágnes Zentai from the Progressive Jewish Synagogue Szim Salom, a qualified lawyer, was the key person who examined these discourses to understand where they transgressed the law and/or indeed were deemed legal but had a hugely negative impact on divisions and xenophobia in Hungarian society. In her contribution, she writes about this topic and asks where is the border between hate speech, hate crimes and hateful speech? What can be allowed under the guise of freedom of expression? And what is to be sanctioned when it has such negative effects on the lives of individuals and whole communities that it exposes them to attacks that have real repercussions in their lives?

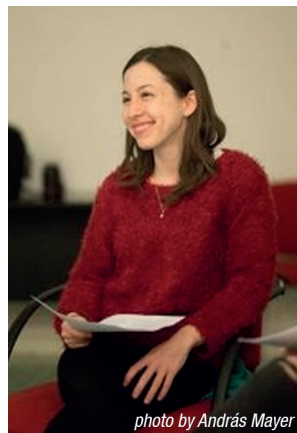
Finally, the project asks the question: how can we all stand up against specific and repeated instances of spreading hate online so that it does not result in a society that is divided, mistrustful and deeply traumatised by everyone having to constantly cover their backs or live in fear?

These are the themes that the coalition will continue to debate and discuss not only with other organisations and communities, but also with public opinion influencers and political actors whose responsibility it is to create a society free of hate speech and where freedom of religion and belief can flourish.





**Ágnes Katalin Kelemen** is a social historian, currently an expert of Hungary as member of an international research project that examines the 20th century history of Central and Eastern Europe from the perspective of refugees. She received her PhD in Comparative History from the Central European University (CEU) in 2019. Previously, she worked as a youth leader at the Szarvas International Jewish camp for five years. She has been involved in the Religious Freedom Programme of NDI, specifically in monitoring online hate speech and developing workshops for young people.



## **The use of social media comes with responsibility**

Drawing on my experience as a historian of the 20th century and Jewish history—also being active in the field of education—I would like to shed light to the great responsibility in the spread of online hate speech of those who create social media content not as individuals, but for institutions, organisations and those who manage such social media platforms. In my view it is a big problem that institutions and organisations often provide a platform for hate speech and comments containing hate speech on social media. In most cases, it is not out of malice, but out of sheer negligence.

As an ordinary social media user, in theory there are increasingly more opportunities to combat online hate speech, since you can report content as hate speech and have it removed from a social media site without confronting the hater in any way. On 31 May 2016 the European Commission and the four large companies that dominate most of the social media (the “Big 4”: Facebook, Twitter, YouTube and Microsoft) set up a code of conduct to help stop the spread of hate speech contents in Europe. They have undertaken to investigate most of the valid reports within one day and to



remove or make inaccessible content that violates national legislation that was transposed from EU law. According to a 2017 survey conducted by Háltér Society, the four companies removed 94% of the content reported by users as hate speech.<sup>1</sup> However, the practice nowadays (in 2021) is the following: unless a large number of people report a specific content all at the same time, the service providers would not qualify it as hate speech, especially if no minority is specifically mentioned in them.<sup>2</sup> I will explain the reasons for this later, but it can only be hoped that algorithms will soon be able to detect even such hidden or concealed content.

Social media providers, however, mostly check content that someone has challenged. Therefore, it is very important that users of social media platforms would be sensitive to hate speech and would always report such content when they encounter it, and no other anomalies would be reported as hate speech if they do not qualify as such. How sensitive and well-informed users become about hate speech largely depends on the institutions and organisations that the users work for or follow on social media.

By 2021, almost all institutions and organisations realised that they need to be present on social media, but few have drawn the appropriate conclusions. On social media, there are many people who can quickly respond to any public content. We can enjoy its benefits, but the consequences are often overlooked when it comes to negative and hurtful comments. It is important that the moderators of a page or group—the people who manage the social media of that institution or organisation—would specify that they will not tolerate any stigmatising related to religion or ethnic origin, the Roma or any other group as it is against the values of the given institution or organisation. Otherwise, certain users would press ‘like’ on hateful comments and the others (the majority) would pretend not to take notice, thereby confirming to the hater that they can freely rampage since the majority silently agrees with them. Hate speech and other types of abuse also spread because of the bystander effect. This means the more witnesses there are to an abuse, the less likely it is that someone will actually help the victim. Peer pressure

---

1 *SThere has been great improvement in action taken by social media against online hate speech* <https://hatter.hu/hirek/sokat-javult-a-kozossegi-media-online-gyul-oletbeszed-elleni-fellepese> (Last accessed on 20.11.2021) Distribution of Online Complaints Per Hate Type

2 Distribution of Online Complaints Per Hate Type  
<https://evidencehub.net/chart/distribution-of-online-complaints-per-hate-type-1.0>  
(Last accessed on 20.11.2021)



will make the majority of us pretend that nothing has happened, and at the same responsibility is divided: “Why should it be me the one who gives a helping hand?”. The bystander effect is being studied by many experts, and there are several ongoing campaigns to stop it.

It is possible that small institutions and organisations do not have a social media manager, but there is an already overworked staff member who is not necessarily skilled in online communication, but they are the ones tasked with the management of the social media profile, which is then reduced to advertising the events of their institution or organisation. For example, cultural institutes in foreign countries often organise commemorations of 20th century events, which are—justly and rightfully—promoted with Facebook posts. The posts themselves are hardly problematic, but the historical traumas of the 20th century (genocides, two World Wars, and dictatorships) that have been processed to very different degrees in the various countries, obviously provoke a wide range of comments from people in the public online world. It could be speech that incites hate or even hate speech.

Speech that incites hate is not always easy to distinguish from hate speech,<sup>3</sup> as it depends on the context whether a statement, besides inciting hatred against a group, puts the members of the group in danger at a given moment. For example, when paramilitary organisations march in Roma neighbourhoods and make racist speeches, it is clear that the Roma residents are in physical danger, as they are effectively held closed and captive in their homes since people who incited others to violence against the Roma are in their close vicinity and they could be attacked at any time. This is would definitely qualify as a *clear and present danger*,<sup>4</sup> since the physical integrity and health of people abused in racist speech can be harmed at any moment in such a situation, and even their lives can be endangered. Unfortunately, there were several examples for this in Hungary in 2006.

---

3 As Ágnes Zentai explained in her essay, hate speech is a hate crime, i.e. a verbal act motivated by severely negative prejudices against a group of people. For more on hate speech, see also the essay on hate speech written by Gerzson Boros.

4 See the essay of this publication written by Ágnes Zentai.



In the online world, the boundaries of space and the presence of members of the vulnerable group are much less obvious. When it comes to public social media content, however, it is reasonable to think that members of the affected group are likely to be present. Furthermore, public hate speech can spread quickly via the Internet, and can put a minority group at risk. It means that there are statements that are considered to be as inciting to hatred in a personal (one-on-one) conversation, but in a public Facebook comment they could be considered as hate speech. Of course, anyone can “report” to Facebook if they think a comment violates the social media platform’s rules and they believe it is hate speech. But in many cases the algorithm fails to acknowledge that it is a valid report, because it does not yet have the large amount of data and contextual knowledge necessary for the artificial intelligence to learn, for example, how to recognise “Gypsy-bashing” or “Jew-bashing”, which is easily spotted by the human brain. The linguistic barrier plays a big role in this, as the meaning of these references (e.g. whether a sentence is humorous or ironic, or whether it is a famous quote or a paraphrase of a famous quote) is defined by cultural references, which are harder to teach to AI than to humans. And there are few native Hungarian-speaking moderators who could help the AI to judge the reported content. Yet such comments are left untouched (because they were not removed, but even reinforced by other users’ “likes”) poisoning the soul of social media user because they portray members of a protected (religious, ethnic, gender, sexual minority) community in a negative light.

In an ideal world, the creator of the post would respond to all problematic comments, pointing out that incitement by the commenter is not in line with the ethical code of the institution or organisation. However, as we do not live in an ideal world, there may not be human resources available to respond to all problematic comments. It is a good practice, and not a very demanding one, to briefly include in the page or group description the values, code of ethics of the institution or organisation, and even more importantly, the moderation principles of the site or group. It would be even better if joining such groups would be subject to answering entry questions, such as “Will you accept and follow the group rules in your posts or comments?”. It is recommended to remind group members of these principles and rules every now and then. The algorithm will anyhow force all site and group operators to create content on a daily basis, and there are not many institutions and organisations that are able to do so, hence the group rule reminders can be very useful even in this respect.



The 444.hu news portal introduced a code of conduct for commenters in early 2016 that contains many sections which are suggested for adoption by all institutions and organisations. One such principle is “Talk to others as you would talk to them face-to-face, and as you would want people to talk and relate to you.”—because as many of us have probably experienced, people can communicate much more aggressively online in writing than in face-to-face communication. Another rule: “Don’t get personal, don’t be aggressive, don’t harass others. Do not make hateful comments about anyone’s gender, religion, ethnicity, physicality, sexuality. Or any other such personal aspects, for that matter.” Another great recommendation for the followers: “Don’t post any hateful content, be it pictures, videos or text.”<sup>5</sup>

Of course, such group rules only achieve their purpose if non-compliance results in appropriate consequences. The most effective way to do this is preferably by checking the group posts and comments twice a day and deleting content that violates the page or group rules. It may be worth explaining a few times why the moderator decided to delete the content that was not in line with the group rules, but if the same user keeps trying to post similar content, it is better to delete problematic posts altogether, without any explanation.

This may sound rather provocative, some may interpret it as a restriction on the freedom of expression (especially the person whose hateful posts are deleted), but in my opinion this is a reasonable restriction for the protection of another important right, the right to human dignity. The more institutional and organisational sites make it impossible to share hate-inciting speech and hate speech on the page and in their groups, the less likely such content will be followed by the vast majority of users. Most users do not inherently hate minorities, but are less vigilant and easily desensitised when hate speech and incitement to hatred become commonplace.

It also means that, unfortunately, we cannot only strive for an ideal world, but we must also remember the tragedies of history—when hate speech gradually led to wars, genocides and other great human catastrophes. It happened even when communication was much slower, before the advent of the Internet. The Internet has led to an incredible acceleration not only of positive content, but also of hate speech, which means that we should have tackled hate speech much earlier, already at the birth of the printing press or already when the first ancient wall graffiti were painted.

---

5 *Commenting rules on 444* <https://444.hu/2016/01/15/kommentelési-szabalyok-a-444-en> (Last accessed on 20.11.2021)





**Gerzson Dávid Boros** data scientist, CEO and founder of Data Science Europe Kft. Gerzson Dávid Boros has been working with machine learning algorithms for almost 10 years. He has worked on more than 20 data science projects, with clients including American, British, Dutch, Russian, Indian, Indonesian and of course Hungarian startups. Gerzson Dávid Boros is a chartered energy engineer, he graduated from the Budapest University of Technology and Economics, and is currently pursuing a Master of Business Administration degree in Artificial Intelligence

at Cumbria University, UK. He is the president of the Hungarian Muslim Youth Society, a committed advocate of critical thinking as well as a percussionist. In the religious freedom programme of NDI he coordinated the monitoring of online hate speech.

### **Combating online hate speech against religious and various vulnerable minority groups**

Together with NGOs and individuals, we created an initiative, in which we joined forces to try to do something together to reduce—even if by a small degree—online hate in Hungary. This very broad topic was narrowed down, and we decided to focus on hate speech against Roma, Jews and Muslims, as these are the most relevant topics for the organisations concerned. Negotiations on the project started in September 2020, while the actual activities were launched in February 2021 and lasted until November 2021.

The project was divided into several phases. The first milestone was the monitoring of online hate speech, which means looking for articles, comments, posts, other writings on the Internet that contain hate speech. In order to perform in a competent manner, we had to learn about the subject, as the team was not trained in this area. We participated in training sessions on the forms and effects of hate speech, the characteristics of online hate speech, the opportunities and limitations of different online platforms, and possible solutions and ways to respond.



Two options were considered for the monitoring activity: either to search the Internet for hate speech using human resources or to use some kind of computerised, automated solution. The former could only have been carried out on a spot-check basis, since there was no time and personnel abundantly available. This would have meant randomly scanning online content to find hate speech, which would have been less efficient and very time-consuming, therefore it was discarded as a solution. We found computer-aided “semi-automatic” hate speech monitoring to be much more effective. It is semi-automatic because the texts are collected by programmed machine algorithms, while it is humans who determine whether the texts contain hate speech or not. At the beginning of our project, there was no algorithm available to us that could determine whether a text contains hate speech in Hungarian without human expertise and intervention. The goal was to create and train such an artificial intelligence algorithm, even a fully automatic one, which is described in detail in this essay. It took a long process to create the semi-automated system, as we had to automate and manage the steps of data collection, data cleaning and data selection.

The data was provided by Textgain, a Dutch company specifically dedicated to automated monitoring of hate speech. The text we received originated from Facebook, Twitter, news portals and other websites, which were filtered with the help of computer programming to reduce the amount of text to be reviewed. In order for Textgain to send us relevant results, and for us to be able to further filter the resulting text fragments through programming, we had to create lists containing a number of words and phrases that refer to hate speech (or are even hate speech in themselves). This list was continuously expanded during the project.

The programmatically selected texts were then arranged into tables and a monitoring team of young volunteers reviewed their content and flagged those that contained hate speech. We could establish that even if something is not legally considered as hate speech, it can still be very harmful and offensive. That is why we introduced the category “hate-inciting speech”. A lesson we learnt was that this kind of monitoring work is psychologically very demanding, it often requires a break and cannot be conducted in the long run.

The most important outcome of this milestone is likely to be the data from the monitoring activity that can be used to teach an artificial intelligence algorithm whether a text contains hate speech or not.



In the second milestone, we organised workshops for young people. The aim was to educate and train young people on online hate speech using non-formal educational methods. The 90-minute workshops were led by two moderators. The goal was for the audience to become aware of why hate speech is a problem, how they can recognise it and what actions can be taken against it.

The programme was advertised on various forums, but unfortunately, with a few exceptions, there were very few people who showed an interest. In our opinion, different conclusions can be drawn from this: either young people are very passive, or they don't understand the importance of this topic, or they are afraid to sign up for such a programme, or they don't even consider hate speech to be a problem. Whichever of the above (or their combination) is true, it is a rather disappointing realisation. Another possible reason could be that we organised the programme online because of the COVID-19 pandemic. However, since schools reopened in June, young people—after all the online learning—probably were not comfortable “sitting in front of the screen” again.

Finally, we held workshops in various camps organised for young people aged between 14 and 18. The programme had a very positive reception. Probably one of the most important outcomes of this milestone was the workshop material developed to educate young people on hate speech that we used to boost their activism against hate speech.

Concurrently with these activities, analyses were carried out on the already monitored data. What we have seen is that there are some terms which are frequently used by Hungarians in hate speech. In the image below, you can find the most frequently used hateful terms in comments, news, blog articles and other online writings about Jews, Muslims and Roma in separate rectangles. Colours are of no particular meaning, but the size of the letters matters: larger words and phrases are more common.





**[Upper left box:** Online hate speech-words most often appearing in hateful online news and comments against Roma, Jewish and Muslim people. Bigger fonts signify more frequent appearance.

**Upper right box:** criminality, murderer, nitty, bastard, stinky, Gypsy, suntanned, gypsy-criminality, rotten, “kolompár” (slang for vxax Roma), orks, thief.

**Lower left box:** foreign-hearted, Zion, Pharisee, bibs and biboldó being pejorative slangs for “Jewish”;

**Lower right box:** Jihadist, terrorism, chador, burka-wearing, migrant, Jihad, terrorist, and migráncs is a “playful version” of migrant]

The third milestone included conducting an online media campaign. It was ran on Facebook, Instagram and TikTok between 23 and 29 September.[1] The videos feature young people acting out a story or talking about basic human values. The videos were produced during the workshop sessions. These young people learnt about filmmaking and—under professional guidance and assistance—they wrote their own stories and recorded these short clips.

Unfortunately, a number of hateful comments were posted under these videos on social media platforms. Such comments were more common under posts where young girls talk about fundamental human rights.

Concurrently with these milestones we worked on creating an artificial intelligence algorithm that can automatically decide—with some accuracy—whether a text contains hate speech. This type of algorithm is



called a classifier algorithm, because it classifies texts into two categories, “hateful” and “non-hateful”. Several different things are required for the creation of such an algorithm. On the one hand, a language model has to be created that can work with Hungarian texts. This was not a problem, as such algorithms have been available before, can be downloaded from the Internet and used for free. On the other hand, a machine learning algorithm has to be trained with data. Previously we had no such data available for training the algorithm. We need the following types of data for our project:

- Text
- Human decision on whether the text contains hate speech or not.

As mentioned earlier, during the monitoring activity the abovementioned data were generated and young people classified the texts into “hate” and “non-hate” categories. This data could already be used to train the algorithm. With the help of various tools, the texts are prepared for the artificial neural network<sup>6</sup>, which is trained with the help of a dataset that represents 67% of all data. Following the training process, it is very important to test the generalisation ability of the algorithm. It means that the algorithm must be able to classify not only texts that it was trained with, but also texts it has never seen before. In order to check the generalisation capabilities of the algorithm, 33% of the data is left as a test dataset.

We managed to create an artificial intelligence algorithm that can classify texts quite accurately. We can adjust the “sensitivity” of the algorithm, which determines the level of severity for the algorithm based on which a text is selected as potentially hateful content. These pre-selected texts are then tediously reviewed by humans to override the decisions made by the AI, thus further training it.

---

6 Artificial neural network: a mathematical model inspired by human neural networks



## Lookout

In the future, it would be worthwhile to develop a data collection algorithm that can automatically scrape content from web pages (so far, the Dutch company Textgain has helped us with this, but we should develop our own solution). Then the existing artificial intelligence algorithm could be used to select texts that may contain hate speech, and these would have to be reviewed by the people doing the monitoring in order to further train the algorithm. It would very important and necessary to respond to such hateful content in some ways. If possible, such content should be reported on the social media platforms or even legal action should be taken. Comprehensive statistics should be compiled, which was impractical for this current project due to the limited data available. Cooperation with other organisations should be made more effective and sustained in the long run. Online campaigns and workshops targeted at young people should be organised on a more regular basis.

## Annex

[1] Links to videos posted on Facebook:

<https://www.facebook.com/PhirenAmenca/videos/328988882301332>  
<https://www.facebook.com/PhirenAmenca/videos/3031041903844939>  
<https://www.facebook.com/PhirenAmenca/videos/832522584096382>  
<https://www.facebook.com/PhirenAmenca/videos/3768385623261155>  
<https://www.facebook.com/PhirenAmenca/videos/362248502236587>  
<https://www.facebook.com/PhirenAmenca/videos/234459205257860>





**dr. Ágnes Zentai** lawyer, mediator, psychodrama leader. Dr. Ágnes Zentai joined the coalition as member of one of the religious associations involved. She served as a judge (in criminal and family law), she taught criminal law in higher education, worked as a lawyer and as a social worker supporting families. Dr. Ágnes Zentai has written a number of articles on the connection between criminal law and the various civil liberties, emphasising that these freedoms are not unlimited, as this would infringe the fundamental rights of others, but that excessive restrictions would

undermine their very essence. In particular, she has devoted much time to the constitutional boundaries of freedom of expression and to hate crimes, both in theory and in practice. This clearly defined her role in the coalition.

**Words are a dangerous weapon and some people are unarmed (János Bródy, Hungarian song writer)** *Hate speech, hate crime, freedom of expression or freedom of opinion and much more.*

How and why did this issue become regulated by law and, in more serious cases, why is it punished with criminal sanctions? There are various approaches and even today there are still many different legal concepts around the world. In Europe the legislation is somewhat more uniform, but very far from being identical. The history and current affairs of a specific country is an important and determining factor in this. Why is it necessary to deal with the issue, with its legal regulation and, within that, criminal law regulation? If I want to be blunt and direct, in one sentence: All genocide starts at the level of words.

However, **freedom of opinion, or more precisely freedom of expression, is a fundamental right in democratic societies and, like all fundamental freedoms, it is a product of the civil revolutions. It is not easy to draw clear boundaries and it is still an important and very difficult regulatory issue for the legislators, but an equally complex one for the enforcement bodies.**

**You can easily acknowledge that its abusive use may cause very serious**



**harm, while its excessive restriction can make this right impossible to enforce. The current Hungarian situation is specific in every respect, and not a bit positive compared to most European countries, but this shall be explained more in detail later.**

In 1789, the Revolutionary Constituent Assembly in France issued the Universal Declaration of Human and Political Rights (which was approved by the King). Shortly afterwards, the first amendment to the then US Constitution was adopted, with the provision still effective today which stipulates that the Congress shall pass no law “on the restriction of the freedom of expression or the press”. Pursuant to this, according to the US Supreme Court, the sanctioning of hate speech cannot, in itself, be limited in the absence of qualifying circumstances. It considers such circumstances to be, in essence, where there is clear evidence of an intention to directly threaten or intimidate persons or groups. Without it, hate speech of any content, uttered before any public, is not punishable by law.

One might think that this has nothing to do with present day Hungary, but unfortunately that is not the case, as it is explained below. The development of European continental law significantly differs from that of the United States in this and many other respects, mainly because of very different historical circumstances.

The French Revolution had a strong impact on the European understanding of liberties because it defined itself in opposition to the existing nobility-based society, unlike in the United States, where the different social development obviously resulted in rather different aspects.

The well-known triple slogan of the French Revolution is liberty, equality, fraternity, with the word equality clearly meaning equality before the law, unlike the misunderstanding of many people. In the previous legal order, different rules applied to people according to their social status, and such rules, of course, were much more favourable to people of higher status.

There is a striking difference between freedom of expression and its limited and unlimited nature in the French legislation and the US legislation, the two of which were adopted with three months between them. The French legislation (dated in 1789!) put the right to freedom of expression as follows: “The free expression of thought and opinion is one of the most precious rights of man; accordingly, every citizen shall be free to speak, write and print, except in cases of abuse of this freedom as provided by law. So thought, opinion is free, it cannot be restricted by anybody, but rules are necessary for its expression.” This was supplemented and clarified in the new (now third) constitution, which came into force in 1795, adding to the



concept of freedoms that “liberty consists in the right to do what does not infringe the rights of others”. Remember, this was written in the late 1700s, in France!

In the course of the next two hundred years, these fundamental concepts seem to have been challenged on a continuous basis. In Hungary, two hundred years later, the opposite process was and is increasingly taking place in the interpretation and consequently, in the application of the law.

For a while, human rights issues were somewhat forgotten even in France. For a long time, the subject was dealt with very differently in the later regimes and system, unlike in the past. Social change always has a fundamental impact on law. Of course, progress in any area is not continuous and unidirectional, as experience clearly shows.

It is no coincidence that the issue of fundamental human rights in Europe very strongly came to the forefront once again, following the Second World War. **The UN General Assembly adopted the Universal Declaration of Human Rights in 1948. This resulted in the document “Covenant on Civil and Political Rights”, which was adopted by over 150 countries in 1966, ratified by Hungary in 1976 and thus made part of Hungarian law. This Covenant is still fundamental today regarding this topic, and it still determines the legislation of the signatories.** Its core and foundation is the idea that every human being is born with equal dignity, and from this derives the fundamental rights to which all human beings are entitled.

[This is where we see ideas already expressed in France during the French Revolution to be reflected, and even though such ideas were disregarded for a while, their influence has grown ever since and it is still powerful today. Namely, that everyone has the right to freedom of expression, but the exercising of this right “carries with it special duties and responsibilities”. They are therefore subject to restrictions, which must be set out by law. It is perhaps not surprising that these restrictive aspects essentially follow the principles developed in the context of the French Revolution. In my opinion the most important question is this: What are the criteria for a legitimate restriction?

According to this document, these fundamental principles are “those which are necessary (a) to respect the rights or reputations of others, or (b) to protect national security or public order, public health or public morals”. It even specifies that “any advocacy of national, racial or religious hatred



which incites discrimination, hostility or violence and which suggests incitement to conduct oneself against others shall be prohibited by law.” However, it does not consider this list to be exhaustive, and individual states may impose additional prohibitions. Although it does not specify the type of law that should regulate this issue, it seems clear that the most serious cases should be regulated by criminal law, while in other cases the rules of civil, administrative, etc. law may be sufficient.

Very importantly, it also **also requires the signatories to “implement their covenantal obligations in a non-discriminatory and effective manner”**.

Non-discrimination, i.e. no distinction can be made between the persons concerned on the basis of gender, colour, nationality, ethnic minority, language, religion, financial status, political affiliation, etc., that is the right of everyone shall be protected, regardless by whom or on behalf of which organisation, etc. the infringement occurred.

An effective system means that there is a system of authorities, courts, etc. in place that perform their respective tasks.

**The current situation in Hungary is in no way compatible with this description, because no matter what legal regulation is in place if the application of the law goes completely against it, i.e. non-discrimination and, above all, effective protection is not guaranteed in any way, rather, it is not guaranteed at all.**

An important EU document is the Charter of Fundamental Rights, which has been binding since 2009 but was adopted by the Council of Europe already in 2000. Following this date only further clarification and consensus regarding its main substance took place. It introduces, among other things, the concept of proportionality with regard to the exercise and enforcement of different rights. This means that certain rights can be subject to the restrictions required to ensure other fundamental rights. In other words, it goes completely against even the de facto unlimited nature of the so-called freedom of expression. The reason for this is that it would make other fundamental rights, such as the right to human dignity, completely unprotected. And then we have not even tackled the very severe social impacts it might have. After all, hate speech without any limits is bound to become increasingly violent and to escalate, especially if it is not met with even the slightest reproach, let alone legal sanction, from the state and the governmental agencies. Although this has not always been the case neither in Hungary.



For a long time this issue was regulated only at the level of criminal law in Hungary. Hate crimes have in common that they are an act motivated by serious negative prejudice against a group of people, the mildest, but not harmless, form of which is the word spoken by someone. This prejudice is imposed on the individuals in the group, suppressing their personal and specific identity, based on which that single individual comes under a negative judgment.

Of course, the law gives priority to those common characteristics that have led to serious consequences several times in the past, such as skin colour, ethnicity, nationality, religion, and then, as time has changed, sexual orientation and gender identity have become increasingly included in the legislation of countries. It is important that this serves primarily the protection of minorities, therefore, it is not applied to those in a significant majority in a specific place, or is applied in a significantly less stringent way, because this would reverse the purpose of the law. After all, throughout history, it is obviously not hate speech against the majority religion, colour, ethnicity, etc. that has led to very serious consequences. At the same time, it is a much lesser offence to human dignity in a given case. And since the equal dignity of all human beings is the very foundation, it does not protect the right of the excluders, the inciters, the threateners to freely express their "opinions".

Unfortunately, in Hungary this has gone offtrack in the application of the law, as you will see in the explanation below. Even though the legislation usually includes a clause that refers to those not listed, it must cover social danger (in terms of criminal law) proportionate to those listed in the legislation in question. In other jurisdictions, comparability based on their own set of criteria is the very foundation.

**So there is a range of hate speech that reaches and ones that does not reach the scope of criminal law, but the point is nevertheless the same. Even if it is not a criminal offence unless it reaches the level of incitement or unless it includes the specific content of another offence, it shall still be considered as hate speech if it is directed at a specific group of people, however phrased, but precisely identifiable and meeting the above criteria, and it contains negative prejudices and strong generalisations against that specific group of people.** It will not matter if the person adds that there are exceptions, or that they have seen before a decent Jew, Gypsy, refugee, etc. In addition to aspects of social peace, tolerance, etc., this is also important in Internet communication,



where, for example, the various social networking sites have their own respective rules, and these must be complied with, otherwise the site will—as expected, but unfortunately not always—impose appropriate sanctions in the event of violation of their rules. Therefore, it may be important for anyone who notices such a situation to report it in the appropriate way.

**The dangers of hate speech and its media coverage, and the importance of combating it with all efforts, are also addressed in other EU documents, which, while not binding, show the EU's commitment in this respect.**

Additionally, already in 1997, the Council of Europe's Committee of Ministers recommended that hate speech should include (in addition to the already mentioned forms, e.g. ethnic hatred, anti-Semitism, xenophobia) aggressive nationalism and aggressive intolerance “expressed through hostility towards” minorities, immigrants, etc. It specifically calls on governments and officials of public administrative bodies and public institutions to refrain from making any statements that can be interpreted as aggressive nationalism or aggressive intolerance to the slightest degree in any statement given to the media, and to publicly distance themselves from such statements. It must be conceded that this is not binding for the member states, but the non-adherence to this rule sends out a very strong message.

**The situation in Hungary took a very different turn.t.**

The Hungarian Criminal Code (abbreviated as Btk. in Hungarian) which was effective at the time of the democratic transition and for many years afterwards (of course with numerous amendments) was Act IV of 1978, which remained effective until it was replaced by the current Act C of 2012. Both of the abovementioned Criminal Codes list among the crimes against human dignity the offences of defamation, slander, and libel (punishable only upon private complaint, i.e. at the request of the rightful claimant), which are worth mentioning in this context and that these are punishable with lenient penalties. These are, in relation to one or more very specific persons, the use of an expression or statement of a fact that may cause defamation, or other such acts, the punishability of which is subject to an additional condition (high publicity, etc.). The same applies when the memory of a dead person is violated. With regard to the statement of facts, the affected person may request proof of the truth under certain conditions. So here the legislative act is very clearly directed at a specific, precisely identifiable person or persons.



In 1989, the legislator amended the previous text, narrowing down the scope of the punishable offences, and removed the offence of incitement against the community from the scope of offences against the state and placed it among the offences against public order.

(1) “Any person who, before the public at large, incites hatred against the Hungarian nation or a nationality, a people, a religion or a race, or against certain groups of the population, commits a criminal offence shall be punishable with up to three years of imprisonment”.

(2) Any person who, before the public at large, uses an insulting or degrading expression against the Hungarian nation, a nationality, a people, a religious denomination or a race, or commits any other such act shall be punishable for misdemeanour (substantially less severe).

In 1992, because of a submitted petition, the Constitutional Court examined the constitutionality of this section and in its Decision No 30 of 1992 upheld the first paragraph unchanged and annulled the second.

The statement of reasons for the first paragraph (in four chapters) were written by Dr András Szabó, professor of criminal law and criminology, and the second by Dr László Sólyom, professor of civil law and then president of the Constitutional Court (of course, all the constitutional judges had staffs and could involve other experts at any time). The Constitutional Court was addressed by the then President of the Supreme Court (now Curia) and the then Chief Prosecutor of Hungary. In their unanimous opinion, this Section 269 of the Criminal Code is in accordance with the effective Constitution. The issue concerned a comparison with Article 61 of the Constitution, which stipulates freedom of expression and freedom of the press.

The reasoning for leaving the first paragraph unchanged (i.e. why it does not infringe freedom of expression and freedom of the press, and more specifically why it meets the conditions that can be considered as a limitation of these fundamental rights if enshrined in law) include such sentences, among others: “1. Humanity has wide historical experience of the potentially harmful nature of hate speech, expressions of contempt and humiliation against certain groups of people”.

Already in 1878, the ministerial reasoning on the Csemegi Criminal Code drew attention to the power of words: “The free communication of ideas, to which mankind owes its most sacred achievements, may become as harmful as fire, which lights and warms, but which, if uncontrolled and unchecked, has very often been the cause of great misfortune, much misery and destruction”.



The so-called Csemegi Code, named after its author, the lawyer Károly Csemegi, is the first Hungarian Criminal Code. It is outstanding even by European standards and is still a point of reference in Hungarian criminal law. It was influenced by the Hungarian Civic Revolution of 1848 and the French ideals of the late 1700s, as already detailed.

The Constitutional Court Decision continues: “The grave historical experience of this century proves that the propagation of racial, ethnic, national, religious view or views that proclaim superiority over others or the lowliness of certain groups, the spread of ideas of hatred, contempt or exclusion endanger the values of human civilisation”.

It is historically proven, as well as by current events, that any manifestation of hatred against specific groups of people is capable of exacerbating social tensions, disturbing social harmony and peace, and in its most serious form, result in violent clashes between different groups in society.” It also refers to the international documents, some of which are mandatory, some of which are recommended, that have already been mentioned above (of course, the ones which were already available at the time).

In defining the concept of incitement to hatred, the decision takes a historic look: “having regard to the fact that incitement to hatred was already a criminal offence in the Csemegi Code, the enforcement authorities can draw on more than 100 years of experience in the judgment of the specific cases”. Here the document quotes previous Curia decisions: “there is agitation only when expressions, remarks, etc. are not addressed to the intellect, but are intended to affect the emotions and passions that are capable of arousing hostility”. And from all this it concludes that “the more serious form of incitement against the public, incitement to hatred, therefore meets the proportionality requirement, it covers only the most dangerous conduct and the factual elements can be clearly interpreted by law enforcement authorities.” Thus, the first paragraph of this offence, that is the one which is still effective under this title in the Criminal Code, was retained.

Then came the last part of the decision, which annulled the second paragraph as an excessive restriction on freedom of expression.

The statement of reasons contains several elements that show that it was written by people who are not experts in criminal law, and it contains a half-sentence, without any explanation, and in the English language, which has subsequently made a ‘career’ with very severe consequences. Namely, the “clear and present danger”, which roughly corresponds to what I wrote in the previous section about the approach and formulation of freedom of



expression in the US, where it is still valid today, and that it is significantly different from the development of continental European law. You may think it means that the restriction is justified if there is an imminent threat to life, physical integrity or health. But if that is what the legislator meant to convey, it should be reflected in the wording of the offence, but it is not there, and the same Constitutional Court decision found no issues in this respect in the first paragraph. In fact, the decision elaborates at length on the concept of incitement, which is light years away from this, while at the same time it is compatible with the development of Hungarian and European law, definitions, and the UN Convention and the EU documents cited, which form part of internal law.

The fact that, having regard to this, the second paragraph has been annulled is a criminal law absurdity, but in itself it would not have caused such a huge problem. The real tragedy came only afterward. In particular, the authorities started to interpret this sentence as if it were a condition of the remaining first paragraph, and we quickly got to the point where not even the most extreme incitement cases were investigated with reference to this section. Or rather, at first there was an investigation, an indictment, but the courts—with reference to the above—acquitted the accused one after another for lack of a criminal act committed, for reasons that are still incomprehensible to me. The situation reached the point where all such charges were rejected already at the investigation stage.

**For years, the legislator has repeatedly tried to supplement the wording of the offence in order to deal with this absurd situation, whereby not even increasingly blatant and highly publicised incitement cases were handled by the law enforcement authorities. That is one issue, but they did not even conduct their own investigations, which would be an obligation in the case of offences brought to their attention.**

In 1996, on the proposal of the Government and by adoption of the Parliament, the criminal offence was amended to include “or commits any other act likely to incite hatred”. The Constitutional Court annulled this amendment in 1999 on the grounds that, according to Decision No 30 of 1992, already described in detail above, a conduct less severe than incitement already constitutes a constitutional restriction on the freedom of expression.

However, this is not true for multiple reasons. On the one hand, in the chapter on the annulment of the second paragraph, there is this sentence “The Constitutional Court has ruled that the dignity of communities may



be a constitutional limit to freedom of expression. The decision does not therefore exclude the possibility that the legislator may provide for criminal law protection that goes beyond the offence of incitement to hatred.” In other words, it says exactly the opposite of what the later decision, the one referring to it, says.

On the other hand, according to the decision an act capable of inciting hatred is in fact also covered by the concept of incitement, on the basis of what has already been quoted at length, irrespective of the fact that in a semantic sense it is a much stronger concept. It is as if the Constitutional Court had taken this decision without having read the earlier decision to which they are referring to. This, in turn, seems to have legitimised a process that is irreversible and is still ongoing to this day, a process that is deteriorating every day, and whose terrible effects we are all feel to an increasing extent. It goes against everything, against historical experience (it must be added here that already at the time of the first decision the Yugoslav Wars were raging, where exactly the same processes were taking place. Here ethnic disputes turned into incitement to the extent that eventually led to war crimes, the deaths and the displacement of a large number of people, only half a century after the Second World War). But it seemingly had no influence neither on the Constitutional Court nor on law enforcement.

**It is of particular interest that the same Constitutional Court in 2000 did not find the criminal act of „violating a national symbol” unconstitutional, which act is literally identical to the second paragraph annulled already eight years earlier. The only difference in the text is that nation, ethnic groups, religion etc. and the group of people belonging to them was replaced with the protected objects of the law, the national anthem, the Hungarian flag, the Hungarian coat of arms, which the Holy Crown have also been added to, and all this is contained in the effective Hungarian Criminal Code. So you cannot say things about these items with impunity that you can freely use to refer to an ethnic or religious group etc.**

In 2003, there was another attempt to amend the text on incitement against a community so that it would curtail this process. It is tragic but not surprising that the Constitutional Court has rejected this amendment without thinking twice, quoting the same arguments as before.



The text adopted by the Parliament replaced the word “instigation” with “incitement”, since in the statement of reasons in the first paragraph of the first Constitutional Court decision the two terms are essentially synonymous, as we have seen in the long explanation on the concept of instigation. There was an addition to the text, namely, “or if the person is calling for the commitment of a violent act”. But to no avail, the Constitutional Court—the members of which, by the way, have undergone several changes in these ten years—only repeated the same thing.

**It is as if the well-known fact did not matter that the poisonous effect of hate speech develops slowly, and by the time it turns to violent action, a long time has passed and many words have been uttered that are increasingly dangerous for the attacked minority, its members being considered as horrible people, treated less and less as humans. At the same time, these decisions totally negate the system, principles, concepts of criminal law and international obligations undertaken by Hungary.**

Two more attempts were made, but in 2008 the Constitutional Court annulled both of them. The statements of reasons were increasingly further away from the law, from reality, from real human rights and from Europe. The Constitutional Court wrote in 2008 that “restrictions on freedom of expression cannot be based on the content of an extremist position, only on its immediate foreseeable consequences”, so as long as the person does not perform an attack immediately after that the words are spoken or they appear in writing, even if the people it is targeted against are present, they consider it still fine and not unlawful. This fact leaves you speechless and shocked.

**Even if you leave the social aspects aside, you still come to the conclusion that only the inciters have human dignity that shall be defended. It was originally human dignity, and by extension the dignity of communities, that set the limits to this freedom, at the level of principles. But those affected, however large their number and serious the harm they suffered, have no effective means to initiate a criminal proceeding.**

**After long debates, the Civil Code was amended in 2007 to slightly compensate for this, and made it possible in such cases to bring an action to court for infringement of personality rights not as an individual, but through membership in the community.**



In this case, not only a member of the community, but also legal persons with specific functions are entitled to present such a petition. In accordance with the nature of civil law, which is substantially different from criminal law, this claim may not be for a penalty, but for other consequences that are substantially less severe than those under criminal law.

**However, even this amendment was annulled by the Constitutional Court in 2008. The Civil Code effective today (Act V of 2013) has brought back this possibility in a significantly narrower form.**

Unfortunately, on the basis of the above, the Constitutional Court practically represented that if neither the people concerned nor those who are attacking them are present, then the right of the person who utters the hate speech or speech inciting to hate in public is more important. This is something that I cannot comprehend, not to mention the increasingly worsening everyday experience we encounter.

**No wonder that after 2010 there has been no attempt by the legislators to find a solution, well, it was to the contrary.** In accordance with the new Criminal Code, Act C of 2012—and of course, before that pursuant to the old and since then also several times amended Criminal Code—incitement against the community (in the currently effective wording instigation), which the legislator placed among the offences against public peace, is as follows: **“Whoever in public incites to violence or hatred against the Hungarian nation, against a national, ethnic, religious group or its members, or against different groups of the population or their members, especially with regard to disability, gender identity, sexual orientation.”**

But it actually does not really matter what wording it contains, because there is no chance of it being applied. At this rate, Hungary will slowly reach the point that we will be actually thankful for it, because at least it is not being used for the very opposite, against those who oppose the inciters. For the time being.

Three more offences should be mentioned among the hate crimes covered by the Criminal Code. Of this, the violation of the national symbol and the absolutely contradictory regulatory absurdity has already been described earlier. The second crime in this form is even more absurd compared to the previous ones.



Its title is the following: **Public denial of the crimes of the National Socialist and Communist regimes.** Even the title is complicated and cumbersome, and the content covers that the title suggests. In other words, if someone denies or relativises the Holocaust, they are undoubtedly committing this crime. But if the person admits that the Holocaust did take place, and he thinks it was a great thing and the problem with it is that it was not fully completed, and does so before the public at large, then this person was just expressing his opinion, exercising the right to freedom of expression.

If the person does not only utters hate speech, incitement to hatred, etc., but by the same motives he also performs actions and deeds, then it qualifies as offence of violence against a member of the community, which is included in the Criminal Code under the chapter on crimes against human dignity and certain fundamental rights.

Pursuant to this: Any person who displays an apparently anti-social behaviour against others for being part, whether in fact or under presumption, of a national, ethnic, racial or religious group, or of a certain societal group, in particular on the grounds of disability, gender identity or sexual orientation, of aiming to cause panic or to frighten others, is guilty of a felony punishable by imprisonment not exceeding three years.

Any person who assaults another person for being part, whether in fact or under presumption, of a national, ethnic, racial or religious group, or of a certain societal group, in particular on the grounds of disability, gender identity or sexual orientation, or compels him by force or by threat of force to do, not to do, or to endure something, is punishable by imprisonment between one to five years.

The law provides for a more severe penalty if the above offence is committed by a group of persons or in a criminal conspiracy, armed, with weapons, causing substantial harm to interest, or by oppressing the victim. As in the case of more serious intentional crimes in general, even the preparation for this offence is punishable.

This offence in itself could be the subject of a separate article, so I will not get into further detail here. However, currently the focus is primarily on verbal or similar hate crimes. This fact contains a number of concepts that have been developed in criminal law but are of no interest to non-lawyers. For all the offences cited, I have only described the statement of facts, but of course they all contain penalties, often quite severe prison sentences, but this is irrelevant to the subject of this article.



Of course, even the facts quoted above have gone through very similar processes, and they have done so in the past. Already ten years ago we were at the point where this was also completely out of control in the application of the law, and there are quite a few concrete court decisions where the attackers became the victims and the attacked became the perpetrators, usually with regard to the Roma.

**And it is visible for everyone that not only the state authorities involved in criminal proceedings do not, but neither does public administration nor the legislature, etc., stand up against incitement, and in many cases they are actively taking the very opposite stance.**

Unfortunately, we all see that the state is not only failing to fulfil this obligation, but in many respects it is leading the way in incitement. It is enough only to think of the various blue billboards that have flooded Hungary (financed by public funds) which attack the refugees, the opposition, or George Soros with an antisemitic argumentation. But you could also include here the portrayal of various civil organisations and NGOs fighting for human rights that are referred to as “queer organisations” or the recent “protect our children” campaign and the related legislation on gender identity and sexual orientation.

**This is the disappointing and rather sad current state of affairs in Hungary regarding the subject.**



Published under the project “Combating online hate speech against religious groups and various vulnerable minority groups”.

