Op:Code

Open Code for Hate-free Communication

Monitoring and reporting
illegal hate speech

Shadow monitoring report / 2nd Edition

ActiveWatch

digiQ

INIMÕIGUSTE KESKUS
HUMAN RIGHTS CENTRE
ЦЕНТР ПО ПРАВАМ ЧЕЛОВЕКА

"NEVER AGAIN" ASSOCIATION
www.nigdywiecej.org

MOVIMIENTO
CONTRA LA
INTOLERANCIA

① Methodology

Between 15th of June and 15th of August 2020, OpCode partners, in close cooperation with INACH, have conducted the second shadow monitoring and reporting exercises of illegal hate speech on social media platforms.

The goal of this joint activity is to verify the social media platforms' Code of Conduct compliance in various periods of time when IT companies are not scrutinized by the European Commissions' official Monitoring Exercises.

Due to the contrasting responses and removal rates that were recorded within the first shadow monitoring exercise, the partners have adapted the workflow and the data management in order to assess if the IT companies use arbitrary standards when analysing the reports. Thus, each national team within the consortium has allocated two rapporteurs: one that is potentially identifiable (username, email and IP) within the IT companies' systems – the test group, and the second one that potentially was unknown for the IT companies (with no previous interactions with the platforms and hidden IP) – the control group.
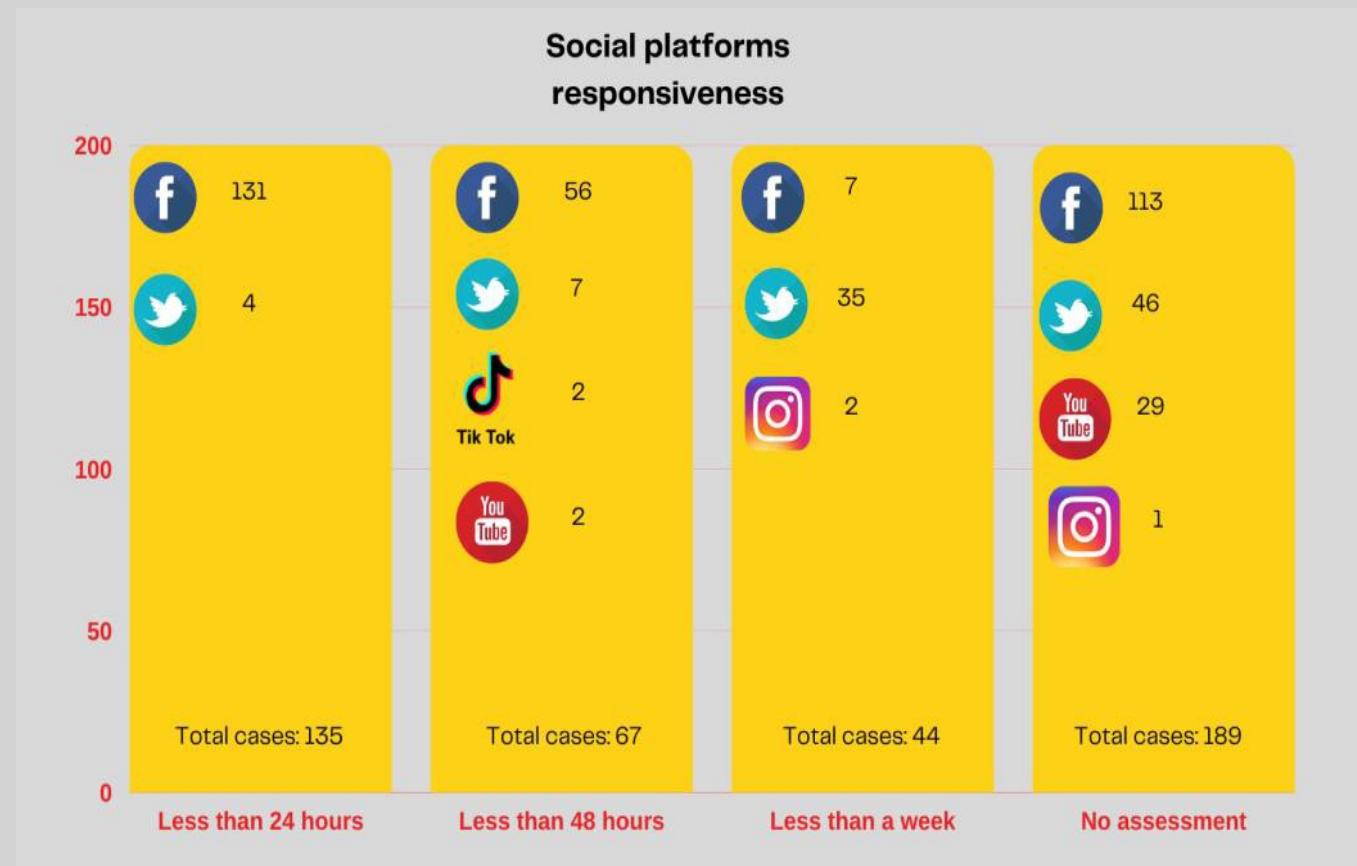
For the monitoring and reporting activities, partners have used the same monitoring and reporting methodology that is being used by the European Commission.

OpCode consortium is formed by: *ActiveWatch (Romania, leading partner), DigiQ (Slovakia), Estonian Human Rights Centre (Estonia), Never Again Association (Poland) and Movimiento Contra la Intolerancia (Spain).*
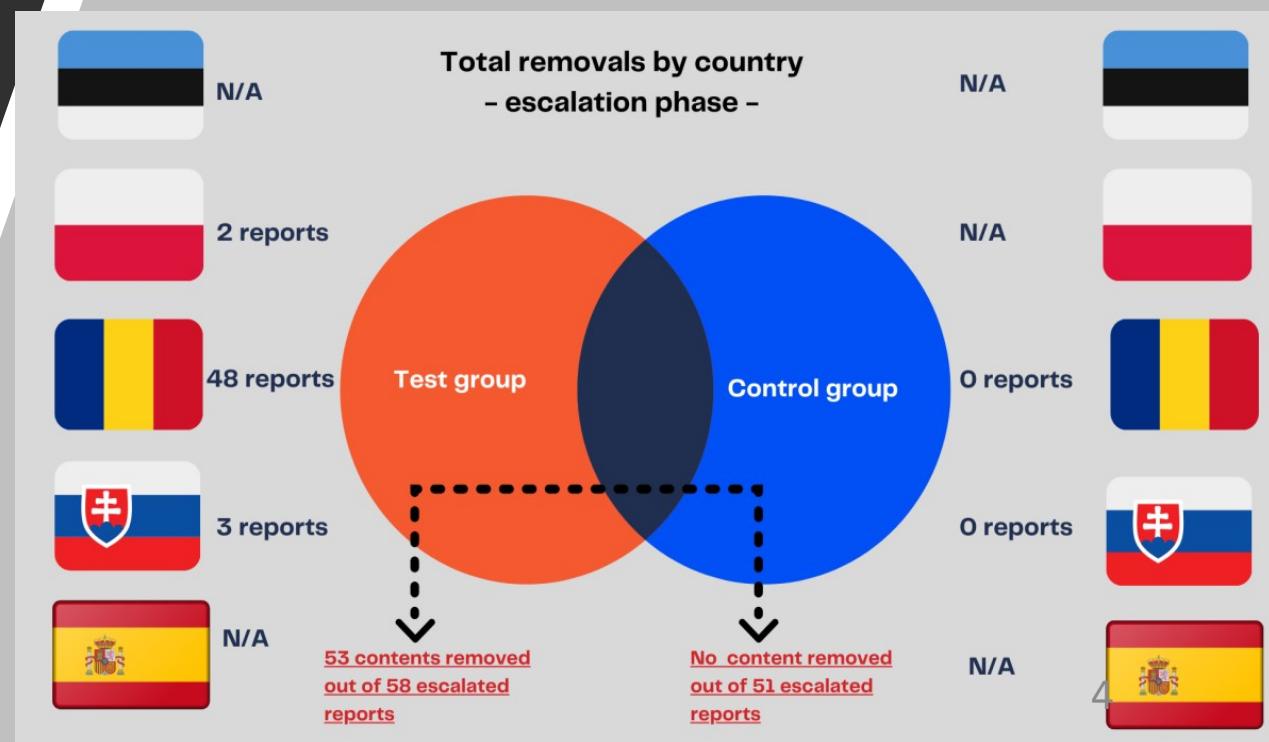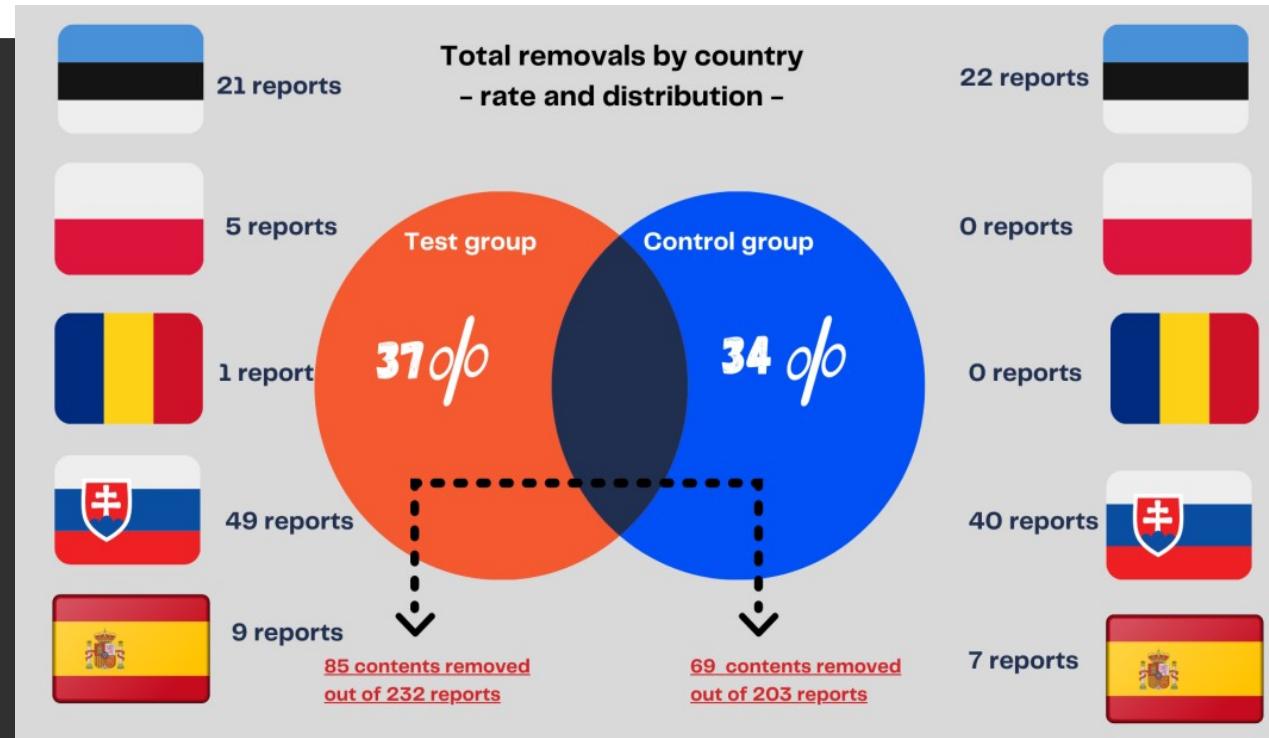
② Conclusions

Almost 43% of the reported content did not receive any feedback from the IT companies.

Almost 90% of the potential illegal and <small>abusive</small> content was identified on Facebook (307 contents), from which the platform managed to assess most of the reported content within 24 hours to one week in 64% of the identified cases. Facebook's performance is lower than in the previous monitoring exercise, yet it should be put in COVID-19 context, when the social platform dedicated most of its resources in other sensitive areas such as fake news or users' misleading conduct in COVID-19 crisis.

Social platforms responsiveness

	Less than 24 hours	Less than 48 hours	Less than a week	No assessment
Facebook	131	56	7	113
Twitter	4	7	35	46
Tik Tok		2		
YouTube		2		29
Instagram			2	1
Total cases	135	67	44	189

Conclusions

Overall, in the first reporting phase, the removal rate was modest and showed no significant treatment between the reporting groups.

Total removals by country – rate and distribution –

		Test group	Control group		
Estonia	21 reports			22 reports	Estonia
Poland	5 reports	**37%**	**34%**	0 reports	Poland
Romania	1 report			0 reports	Romania
Slovakia	49 reports			40 reports	Slovakia
Spain	9 reports	85 contents removed out of 232 reports	69 contents removed out of 203 reports	7 reports	Spain

Total removals by country – escalation phase –

		Test group	Control group		
Estonia	N/A			N/A	Estonia
Poland	2 reports			N/A	Poland
Romania	48 reports			0 reports	Romania
Slovakia	3 reports			0 reports	Slovakia
Spain	N/A	53 contents removed out of 58 escalated reports	No content removed out of 51 escalated reports	N/A	Spain

Conclusions

In the first reporting phase (escalation is not included), Facebook had the highest assessment and removal rates: 68% of the reporting were assessed, and 71% of the assessed content was removed (196 assessments out of 307 submissions, 139 contents removed). Twitter assessed only 52% of the reported content (48 out of 92 reports) and removed 12 contents (25%). YouTube assessed only 2 reports out of 31 (6%) and removed only two contents but without feedback.

Yet, within the escalation phase, IT companies gave higher credit to the rapporteurs that were included in the test group (those users that were identifiable within the social platforms' internal systems).

Given the pandemic context, all social platforms should allocate additional resources to monitor and redress hate speech as this phenomenon is equally dangerous as fake news is and most of the times, they overlap within hate narratives.

The antisemitic narratives and the conspiracy theories are significantly present in social media and prompt responses are mandatory from all social platforms.

Monitoring exercise results

During the second shadow monitoring exercise, all partners have submitted 345 reports of potential harmful and illegal content.

Only 57% of the reported content (246 contents) received feedback from the IT companies. After the escalation, the responsiveness increased to 88%, as 59 escalated reports has received feedback from the social platforms.

Distribution of reports
COUNTRY/IT COMPANY
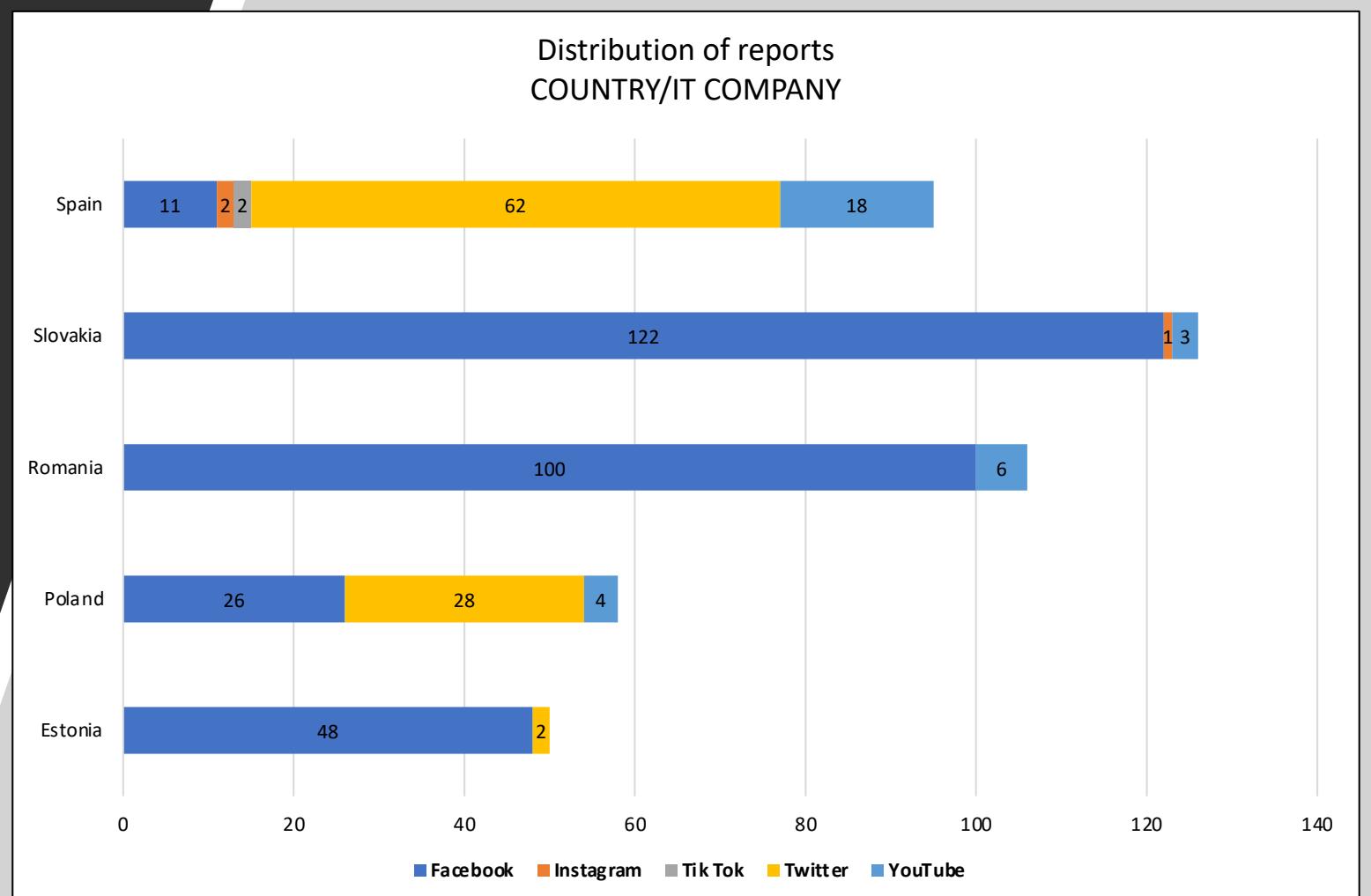
Monitoring exercise results

The highest rates of removal were identified within the Estonian team (43 removals out of 50 reports, with no subsequent escalation) and within the Slovakian team (89 removals out of 126 reports). The lowest response and removal rates were identified within the Polish, Romanian and Spanish teams.

ESTONIA	Assessment rate	Removal rate	First reporting phase
Facebook	98%	91%	47 assessments out of 48 submissions, 43 contents removed
Twitter	0%	0%	2 submissions, 0 assessments

SLOVAKIA	Assessment rate	Removal rate	Frist reporting phase
Facebook	98%	73%	119 assessments out of 122 submissions; 87 out of 119 assessments removed
Instagram	0%	0%	1 submission, 0 assessment/removal
Youtube	0%	67%	3 submissions, 0 assessments, yet 2 of the reported contents were removed

ROMANIA	Assessment rate	Removal rate	First reporting phase
Facebook	2%	50%	2 assessments out of 100 submissions; 1 out the 2 asssessed contents was removed
Youtube	0%	0%	6 submissions, 0 assessments and removals

Poland	Assessment rate	Removal rate	First reporting phase
Facebook	88%	22%	23 assessments out of 26 submissions; after assessment only 5 contents were removed
Twitter	0%	0%	28 submissions, 0 assessments/removals
YouTube	25%	0%	1 assessment out of 4 submissions; 0 removal

Spain	Assessment rate	Removal rate	First reporting phase
Facebook	45%	60%	5 assessments out of 11 submissions, only 3 contents removed
Twitter	77%	25%	48 assessments out of 62 submissions, 12 removed contents
YouTube	6%	0%	1 assessment out of 18 submissions, 0 removals
Instagram	100%	0%	2 assessments out of 2 submissions, 0 removals
Tik Tok	100%	50%	2 assessments out of 2 submissions, 1 content was removed

Monitoring exercise results

As for the duration of assessments, the highest responsiveness was recorded in Estonia and Slovakia, as opposed to the other participating countries within the exercise.

Assessment time	Platform	Estonia	Romania	Slovakia	Poland	Spain
Less than 24 hours	Facebook	47		65	18	1
	Twitter					4
Less than 48hours	Facebook		1	49	4	2
	Tik Tok					2
	Twitter					7
	YouTube				1	1
Less than a week	Facebook		1	4	1	1
	Instagram					2
	Twitter					35
No assessment	Facebook	1	98	4	3	7
	Instagram			1		
	Twitter	2			28	16
	YouTube		6	3	3	17

④ Escalations within the monitoring exercise

Within the escalation phase, the Polish, Romanian and Slovakian teams escalated 109 reports that received no feedback from the social platforms. At the end of the monitoring exercise, only 53 reports were removed by the IT companies, while the other 56 received no feedback not even on the dedicated channels for escalation procedures.

In the escalation phase, Facebook removed 48 out of the 100 reports that were submitted by the Romanian team on the dedicated channel. All removed content has been submitted by the test group (the identifiable member of the team), while the other 52 reports received no feedback. It is worth mentioning that the escalation and the content removal took place in less than 3 hours from the submission.

The Slovakian and Polish teams escalated other 5 reports that eventually were removed by Facebook. The other platforms gave no feedback within the escalation phase.

Grounds of the illegal content

The most frequent categories of illegal content that were reported within this monitoring exercise were:

antigypsyism (most present in Romania and Slovakia), antisemitism (most present in Poland and Romania), racism (high presence in Slovakia) and anti-refugee hatred (Slovakia).

GROUNDS OF ILLEGAL CONTENT	Estonia	Poland	Romania	Slovakia	Spain	Grand Total
Antigypsyism			65	33	1	99
Antisemitism(includingHolocaust denialor revisionism)	2	38	22	2	14	78
Racism	8	16		41	1	66
Anti-refugees hatred	2			39		41
Xenophobia	4		4		24	32
Hatred related to sexual orientation	9	1	8	5	7	30
Anti-Muslim hatred	3		1		18	22
Hatred related to skin colour	15		3		1	19
Other	1		3		9	13
Anti-Arab racism		2	1		7	10
Gender related hatred				6	2	8
Glorification of Nazism or Fascism	1				7	8
Hatred related to origin	5				2	7
Hatred related to ethnicity					1	1
Hatred related to ethnicity, Glorification of Nazism or Fascism					1	1

Grounds of the illegal content

In Estonia, most of the hateful comments targeted people of colour and LGBTQ+ community.

ESTONIA	Facebook	Twitter	Grand Total
Hatred related to skin colour	15		15
Hatred related to sexual orientation	9		9
Racism	6	2	8
Hatred related to origin	5		5
Xenophobia	4		4
Anti-Muslim hatred	3		3
Anti-refuge hatred	2		2
Antisemitism (including Holocaust denial or revisionism)	2		2
Glorification of Nazism or Fascism	1		1
Other	1		1

Grounds of the illegal content

In Poland, *antisemitic* messages and *racist* comments have been the most frequent issues during the monitoring exercise.

POLAND	FACEBOOK	TWITTER	YOUTUBE	GRAND TOTAL
ANTISEMITISM (INCLUDING HOLOCAUST DENIAL OR REVISIONISM)	13	23	2	38
RACISM	9	5	2	16
ANTI-ARAB RACISM	2			2
CONTENT CALLING FOR THE MURDER/ANNIHILATION OF CERTAIN GROUPS/INDIVIDUALS BELONGING OR PERCEIVED TO BELONG TO SUCH GROUPS	1			1
HATRED RELATED TO SEXUAL ORIENTATION	1			1

Grounds of the illegal content

In Romania, messages targeting *Roma community* and *antisemitic comments* occurred mostly during the monitoring.

ROMANIA	FACEBOOK	YOUTUBE	GRAND TOTAL
ANTIGYPSYISM	62	3	65
ANTISEMITISM (INCLUDING HOLOCAUST DENIAL OR REVISIONISM)	19	3	22
HATRED RELATED TO SEXUAL ORIENTATION	8		8
HATRED RELATED TO SKIN COLOUR	3		3
OTHER	3		3
XENOPHOBIA	3		3
ANTI-ARAB RACISM	1		1
ANTI-MUSLIM HATRED	1		1

Grounds of the illegal content
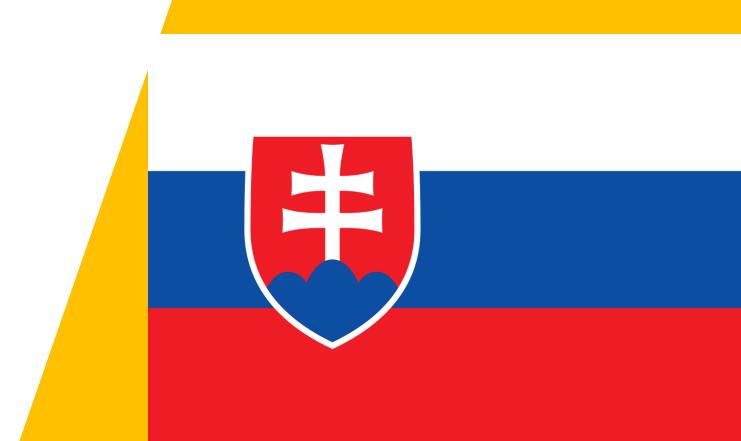
In Slovakia, most of the hateful comments were *racist* or they were *targeting refugees*. Violent messages *against Roma community* were also very substantial in social media.

SLOVAKIA	FACEBOOK	INSTAGRAM	YOUTUBE	GRAND TOTAL
RACISM	41			41
ANTI-REFUGE HATRED	39			39
ANTIGYPSYISM	31		2	33
GENDER RELATED HATRED	6			6
HATRED RELATED TO SEXUAL ORIENTATION	5			5
ANTISEMITISM (INCLUDING HOLOCAUST DENIAL OR REVISIONISM)		1	1	2

Grounds of the illegal content

In Spain, *xenophobic* and *anti-Muslimism* hateful messages occurred mostly, along with a*ntisemitic messages*.

SPAIN	Facebook	Instagram	Tik Tok	Twitter	YouTube	Grand Total
Xenophobia	6			16	2	24
Anti-Muslim hatred	1			13	4	18
Antisemitism (including Holocaust denial or revisionism)	3	1	1	9		14
Other				4	5	9
Anti-Arab racism		1		6		7
Glorification of Nazism or Fascism				1	6	7
Hatred related to sexual orientation	1		1	5		7
Gender related hatred				2		2
Hatred related to origin				2		2
Antigypsyism				1		1
Hatred related to ethnicity				1		1
Hatred related to ethnicity, Glorification of Nazism or Fascism					1	1
Hatred related to skin colour				1		1
Racism				1		1

Typology of the illegal content

6

Almost half (46%) of the reported content consisted in violent and explicit calls for murder and annihilation of individuals and groups that were perceived as undesirable or as a threat for communities (local, national or European). This category occurred mostly in Romania, Slovakia and Estonia.

Dehumanizing and degrading speech against various groups mostly occurred in Poland and in Slovakia.

ILLEGAL HATE SPEECH - TYPOLOGY	Estonia	Poland	Romania	Slovakia	Spain	Total
content calling for the murder / annihilation of certain groups/individuals belonging or perceived to belong to such groups	21	2	74	58	3	158
using degrading, defamatory words/expressions to name certain social groups/individuals belonging or perceived to belong to such groups	13	56	1	44	27	141
content calling for violent acts against certain groups/individuals belonging or perceived to belong to such groups	8		5	19	11	43
condoning/glorifying, denying or grossly trivialising historical events relevant to certain groups	3		18	1	15	37
content calling for limiting the rights of, or otherwise discriminating against certain groups/individuals belonging or perceived to belong to such groups	5		7		19	31
containing expressing insulting overgeneralising statements about certain groups/individuals belonging or perceived to belong to such groups			1	2	10	13
diffusing degrading, defamatory pictures/images relating to certain groups/individuals belonging or perceived to belong to such groups				2	10	12

Typology of the illegal content

In Estonia, more than half of the reported contents (34 out of 50) were calls for murder and annihilation or they portrayed various groups and individuals in a dehumanizing manner.

ESTONIA	Facebook	Twitter	Grand Total
content calling for the murder / annihilation of certain groups/individuals belonging or perceived to belong to such groups	21		21
using degrading, defamatory words/expressions to name certain social groups/individuals belonging or perceived to belong to such groups	11	2	13
content calling for violent acts against certain groups/individuals belonging or perceived to belong to such groups	8		8
content calling for limiting the rights of, or otherwise discriminating against certain groups/individuals belonging or perceived to belong to such groups	5		5
condoning/glorifying, denying or grossly trivialising historical events relevant to certain groups	3		3

Typology of the illegal content

In Poland, almost 97% of the reported content consisted in degrading and dehumanizing hateful speech against various groups and minorities.

POLAND	Facebook	Twitter	YouTube	Grand Total
using degrading, defamatory words/expressions to name certain social groups/individuals belonging or perceived to belong to such groups	24	28	4	56
content calling for the murder / annihilation of certain groups/individuals belonging or perceived to belong to such groups	2			2

Typology of the illegal content

In Romania, 87% of the reported content were explicit and violent calls for murder or they were trivializing Holocaust or glorifying Nazism.

ROMANIA	Facebook	YouTube	Grand Total
content calling for the murder / annihilation of certain groups/individuals belonging or perceived to belong to such groups	73	1	74
condoning/glorifying, denying or grossly trivialising historical events relevant to certain groups	15	3	18
content calling for limiting the rights of, or otherwise discriminating against certain groups/individuals belonging or perceived to belong to such groups	7		7
content calling for violent acts against certain groups/individuals belonging or perceived to belong to such groups	3	2	5
containing expressing insulting overgeneralising statements about certain groups/individuals belonging or perceived to belong to such groups	1		1
using degrading, defamatory words/expressions to name certain social groups/individuals belonging or perceived to belong to such groups	1		1

Typology of the illegal content

In Slovakia, more than 80% of the reported illegal content consisted in violent calls for murder and annihilation of certain groups, along with explicit degrading and defamatory speech against certain groups or individuals.

SLOVAKIA	Facebook	Instagram	YouTube	Grand Total
content calling for the murder / annihilation of certain groups/individuals belonging or perceived to belong to such groups	57	1		58
using degrading, defamatory words/expressions to name certain social groups/individuals belonging or perceived to belong to such groups	42		2	44
content calling for violent acts against certain groups/individuals belonging or perceived to belong to such groups	19			19
containing expressing insulting overgeneralising statements about certain groups/individuals belonging or perceived to belong to such groups	2			2
diffusing degrading, defamatory pictures/images relating to certain groups/individuals belonging or perceived to belong to such groups	2			2
condoning/glorifying, denying or grossly trivialising historical events relevant to certain groups			1	1

Typology of the illegal content

In Spain, almost 65% of the reported content consisted in dehumanizing speech against certain groups or individuals, followed by explicit calls for limitation of rights for these groups or individuals. Content trivializing or glorifying dramatic historical events was also present in hateful messages.

SPAIN	Facebook	Instagram	Tik Tok	Twitter	YouTube	Grand Total
using degrading, defamatory words/expressions to name certain social groups/individuals belonging or perceived to belong to such groups	2	1		22	2	27
content calling for limiting the rights of, or otherwise discriminating against certain groups/individuals belonging or perceived to belong to such groups	1			17	1	19
condoning/glorifying, denying or grossly trivialising historical events relevant to certain groups	2	1	1	4	7	15
content calling for violent acts against certain groups/individuals belonging or perceived to belong to such groups	2			1	8	11
containing expressing insulting overgeneralising statements about certain groups/individuals belonging or perceived to belong to such groups	1		1	8		10
diffusing degrading, defamatory pictures/images relating to certain groups/individuals belonging or perceived to belong to such groups	3			7		10
content calling for the murder / annihilation of certain groups/individuals belonging or perceived to belong to such groups				3		3

7

Country specific challenges

ESTONIA // In the Estonian context, it was essentially impossible to find content which would exceed the necessary threshold to fall under the Estonian Penal Code provision for incitement to hatred, as the provision requires an immediate danger to life, health or property of a person. Regarding the response rate from social media platforms, Facebook was quick to assess the reports and remove content, however, Twitter failed to respond most of the times.

POLAND // The biggest challenge is definitely the unwillingness of the biggest social media platforms to react and remove the reported content. Many of the reports remained not assessed at all. YouTube and Twitter seem to be the worst platforms when it comes to removing hateful or illegal content, they also do not have a clear system of communicating with users reporting the content. Facebook reacts more eagerly to the reports made by the test group, but the removal rate is still alarmingly low.

ROMANIA // The poor assessment rate from Facebook has been the biggest challenge within the monitoring exercise. Eventually the platform removed most of the escalated content, yet only the content that was reported by the test group (the identifiable rapporteur from ActiveWatch) and completely ignored the content that was reported by the control group (the shadow rapporteur from ActiveWatch).

SLOVAKIA // The main challenge of the monitoring exercise was to find out how social networks evaluate the reporting of hate speech from regular reporters and standard users. The social platform removed 70% of reported content. The test group had a slightly higher removal rate of reported posts than the control group. Most of the escalated cases were removed by the social network Facebook, other platforms did not send any feedback.

SPAIN // Certain difficulties were met in finding actual and clear illegal contents, under the provisions of articles 224 and 510 of criminal law. It might be due to the fact that hate speech has been on public debate for years. Currently there is a network of 52 Hate Crimes prosecutors around the country and criminal law is very strict in this concern. It is worth mentioning that online hate speech and intolerance perpetrators have adapted their narratives by using ambiguous semantics in order to avoid legal scrutiny.

Country profiles // Estonia

Current socio-political challenges that could contribute to online and offline radicalization

Estonia held the general election on March 3,2019. Centre Party, Conservative People's Party of Estonia (EKRE) and Isamaa coalition entered office even though the Reform Party had won the most seats. Coalition member EKRE almost tripled its seats and received nearly 18 percent of the vote and got into the government for the first time. The far-right, nationalistic, anti-EU, anti-immigration, anti-LGBT and populist party has changed the Estonian public discourse significantly and made hate against minorities mainstream. EKRE politicians have made, among other things, public misogynistic, anti-Semitic, Russophobe, homophobic and racist comments. They also claim that a shadowy "deep state" secretly runs the country and they openly attack public servants for that reason. In addition, the party has taken aim at their critics in the media and they have their own successful news portal that continuously produces fake or questionable news and also news and opinion pieces that incite to hatred against parts of the population. The party has also attacked NGOs that work in the field of human rights. The inclusion of EKRE in the government has raised fears about the radicalization of social media discourse and self-censorship of the media. For example, there have been two cases of resignations by journalists who claim they were pressured by their editor-in-chief to moderate their coverage about EKRE.

Current radical or extremist movements that engage in online or offline radicalization

There is very little existing research on the topic. The Estonian Internal Security Police has stated in its annual reviews that in general, extremist ideologies have no popular base in Estonia. Based on the information available from public sources, two movements could be considered holding at least partly radical views.

The Soldiers of Odin originates from Finland but has its branch in Estonia. It is an anti- immigrant organization founded in 2015 in Finland by a violent neo-Nazi named Mike Ranta. The group has called for volunteers who are willing "to step out for the defence of our own people in the face of strangers." Facebook has banned the movement's pages several times. The number of members in Estonia is unknown, the group has stated that they have 5000 members and 120 men who are ready to patrol the streets. The movement is known for their "intimidation effect", taking part in anti-immigration and anti-LGBT protests. Last year they sent out a warning message to a bar that had hoisted a rainbow flag and they were also present at an LGBT movie screening to publicly insult organizers and visitors. **Blue Awakening (BA)** - youth wing of EKRE was founded in 2012 by Ruuben Kaalep (now MP) who has a long history of close association with far-right, white supremacist and neo- Nazi groups and figures. BA is known for radical nationalist statements and also for torch processions.During these gatherings the leaders give speeches highlighting the unique nature of Estonians and a call for a new national awakening. The group has also called for ethnic profiling of those with access to state secrets. Last year BA organized firearm trainings for its members and photos from the practice at a firing range with Kaalep appeared on several group members accounts. In 2017 Blue Awakening was thrown out of European Young Conservatives.

Country profiles // Estonia

Existing legal and institutional framework that address online and offline radicalization/ extremism

The Estonian Penal Code includes a provision on prohibiting incitement of hatred. This prohibition does not work in practice due to the wording of the provision, according to which only such incitement of hatred is punishable, which poses an immediate danger to life, health or property of a person. In 2017 and 2018 the provision found no use. Defamation was decriminalized in 2002. Violent extremism and radicalization at policy level is dealt with under the internal security policy, which is the area of responsibility of the Ministry of the Interior and the Internal Security Service in particular. The government has set the policies for combating terrorism in the framework document on combat against terrorism approved in November 2013. In this document it is drawn out that Estonia is paying particular attention in order to prevent the activities of movements and people who spread racial, cultural, religious or affiliation hatred in Estonia. It is stated that earliest signs of radicalism must be prevented, because – as shown by international experience – the growth of radicalism into extremism and then into terrorism is probable and therefore also an undisputed threat to the security of the state. At the same time, there is no specific governmental program to specifically address the issue of prevention of radicalization and violent extremism but there have been some projects by law enforcement authorities and youth organizations. For example, in cooperation with the Estonian Academy of Security Sciences, the Estonian Police and Border Guard are raising the capacity of first line practitioners to detect early signs of radicalization. As a part of the project, the Police issued a manual for local authorities called "Early detection of radicalization and networking".

In addition, it is possible to report on-line hate speech or (potential) signs of radical/extremist content in the Estonian context to web-constables on Facebook. Web-constables are police officers working on the Internet. They respond to notifications and letters submitted by people via the Internet, mainly Facebook, where they have their user profile. Some issues are solved by consultation only, but some of the notifications are forwarded for information or proceedings to relevant police stations.

Online trends andbehavior (e.g.: Internetaccess,demographics,mostpopular [social] media)

There are no significant digital divides in the country and the 2019 Inclusive Internet Index report ranks Estonia 20 out of 100 countries in terms of the affordability of prices for connections. According to Statistics Estonia, as of 2019, 90 percent of Estonian households have an internet connection at home, and it is used daily or almost daily by 98 percent of people between the ages of 16-44. The share of daily or near daily users among people between e ages of 65-74 increased from 71 to 75 percent compared to last year. The most popular website in Estonia is Google, followed by YouTube and Facebook. Instagram is on the 7th and Twitter on the 24th place.

Country profiles // Poland

Current socio-political challenges that could contribute to online and offline radicalization

Since 2016, a significant rise in popularity of radical far-right political parties and groups it has been recorded in Poland. The influence of these organizations can be seen in what happens on the streets (radical manifestations, violence, discrimination) and in the huge amount of hateful content online. Polish government doesn't put pressure on social media platforms when it comes to removing hateful and violent content. After the last parliamentary elections in Poland, 11 far-right politicians have made their way in the Polish Sejm (the lower chamber of the Polish parliament), which has contributed to hate speech spreading in the Polish parliament. Hate speech by politicians and journalists encourage physical violence and discrimination in real life.

Current radical or extremist movements that engage in online or offline radicalization

The most dangerous movement in Poland nowadays is Konfederacja (Confederation) - because they managed to win 11 deputy seats in the Polish parliament. Konfederacja was created by members of several extreme political movements,such as Młodzież Wszechpolska (The All-Polish Youth), KORWIN (named after its leader, Janusz Korwin-Mikke) and ONR (Obóz Narodowo-Radykalny, National-Radical Camp). The mentioned groups have united for the purpose of European and Polish parliamentary elections. All groups are known for actions and demonstrations that incite to hatred such as "The Independence March", hosted annually on 11.11 in Warsaw. In Poland there are also many smaller radical and neofascist organisations.

Country profiles // Poland

Existing legal and institutional framework that address online and offline radicalization/ extremism

The 13th article of Polish Constitution forbids the existence of political parties and other organizations that are based on totalitarian ideology, such as Nazism and fascism. Polish penal code contains two articles on similar matters. Article 256 penalizes propagation of fascism or other totalitarian systems, as well as incitement to hate on the basis of nationality, race, religion and ethnicity. It does not, however, list hatred based on gender, age or sexual orientation as forbidden. Article 257 of Polish penal code states it is illegal to insult a person or a group of people because of their nationality, race, ethnicity and religion. It is worth noting that all of the mentioned articles are used extremely rarely.

Online trends and behaviour (e.g.: Internet access, demographics, most popular [social] media)

In Poland there is widespread access to the Internet and most people use social media. The most popular platform are Facebook and YouTube, followed by Instagram and Twitter. Twitter is used mainly by politicians and journalists. All of these platforms are facing problems like the rising amount of hateful content. It includes comments (homophobic, Islamophobic, racist, sexist, inciting to violence), posts written by public figures that also include severe hate speech, pictures and graphics (including memes) that also carry hateful ideology. Those platforms (especially Facebook) mostly fail at removing hateful content.

Country profiles // Romania

Current socio-political challenges that could contribute to online and offline extremism

Romania is one of the former communist countries and one of the newest EU members that still struggles to overcome the socio-economic gap with western European countries. In 2019, Romania ranked 63rd in the Global Democracy Index issued by The Economist Intelligence Unit[2], within a larger group of countries that could be characterized as flawed democracies. Despite that, Romania has registered significant progresses in the late 20 years. The tormented political life along with the economic disparities and social challenges that came after the accession to the EU (such as massive economic migration within EU borders) could potentially contribute to an upraise in extremism in the near future. No extremist or radical political parties are currently in power or even in the Romanian Parliament. Historically, the most popular extremist party has been Greater Romania Party that has been active in the Romanian Parliament until 2008. Over the years, various political parties with extremist agendas were established but none of them managed to become mainstream. In terms of populist and nationalist agendas, the most visible political party isthe PSD (Social Democratic Party) that has significantly contributed to the deterioration of democratic life between 2016 and 2019, when Romania has been included in the "club" of countries showing illiberal tendencies, alongside Hungary. During the 2019 European and Presidential elections the Romanians gave a strong signal against the populist-nationalist agenda and voted massively for pro-European political parties. Currently, Romania is preparing for general elections and there are no indications that populist narratives will be promoted by mainstream political parties.

Extremist movements that engage in online or offline extremism

There are no official records on the activity of extremist movements in Romania, but there are indications on the existence of civic and religious groups that promote a conservative agenda in Romania. The most active is the Coalition for Family, a conservative network of religious organizations that have in 2018 triggered a national referendum that would have consolidated the ban of LGBTQ marriages3. The New Right Movement is another far-right political party that promotes an ultra-nationalist agenda and is mostly visible with specific public actions against the LGBTQ community and migrants4.

Country profiles // Romania

Existing legal and institutional framework that address online and offline extremism

Romania has several legal provisions both in the criminal and civil code that address extremist manifestation but has limited success in enforcing them. Extremism is criminalized by Law 535/2004 (updated in 2019) on combating and redressing terrorist acts. Other provisions are comprised in Governmental Ordinance 31/2002 that ban xenophobic, racist and Nazis organizations and symbols. The Governmental Ordinance 31/2002 has been updated in 2015 by Law 217/2015 to clearly ban extremist nationalistic movements. Moreover, incitement to hatred is punishable by the Criminal Code under article 369 and hate crimes are considered to be aggravating conditions in establishing a criminal offence. In terms of institutional framework, the Romanian Intelligence Service and the General Directorate for Investigating Organized Crime and Terrorism are the most qualified institutions to monitor and investigate extremist organizations or movements. As for online extremism that is mostly reflected in illegal hate speech, despite the existent legal provisions, Romanian authorities pay little to no attention to this phenomenon.

Online trends and behaviour (e.g.: Internet access, demographics, most popular [social] media)

Romania is a fast-developing country in adopting Internet technologies, with more that 80% of the households connected to broadband internet, but with significant gaps between urban and rural areas. According to Speedtest Global Index, in December 2019 Romania was the 4th country in the world in broadband Internet speed and the 41st in mobile internet speed. Despite this expanding infrastructure and the affordability of devices and services, in 2019 Romania was on the 27th place within the European Digital Economy and Society Index5. Facebook is the most popular platform in Romania, with more than 10 million users and with a market share of more than8 7%. Social media are mostly accessed for entertainment purposes, although youth tend to use social media as an information source as well. Mainstream media, especially television, is still credited as the primary information source for more than 80% of the Romanians. It is worth mentioning that television is starting to decline in viewership – in the last 5 years more than one million Romanians have changed television programs for other media.

1 Full report here

2 The referendum did not meet the legal threshold and was not validated by the Romanian Constitutional Court.

3 Since 2000 The New Right Movement never succeeded to run for local or national elections. Moreover, their activities are not visible in mainstream media.

4 Full report here

Country profiles // Slovakia

Current socio-political challenges that could contribute to online and offline radicalization

Poverty, unemployment, poor housing conditions, alcoholism and other social problems contribute to the boom of radicalism. 16% of Slovak citizens are exposed to social exclusion and poverty. Social unrest is exploited by various populist parties. In addition, the extremist political party Ľudová strana naše Slovensko (ĽSNS - People's Party Our Slovakia) currently has representatives both in the national and the European Parliaments. While, until recently, other political parties have distanced themselves from this party, some parties are currently cooperating with it and thereby legitimizing it.

Current radical or extremist movements that engage in online or offline radicalization

Extremist and racist groups experienced a boom in Slovakia especially in the 1990s. At that time the Internet was not widespread, racist ideology and neo-Nazi propaganda were spread through skinhead movements. Approximately 10 such magazines are published now. The most dominant ones were Blood & Honour Division Slovakia and Slovakia Hammer Skins, which are branches of international neo-nazi organisations. Later, the Slovak National Front established in Trnava (advocating the adoption of anti-communist and non-seizure laws, the death penalty and the law to combat homosexuality), which was divided into a political People's Party and the socially oriented Slovak Community. In 2005, members of the Slovak Community registered the political party Slovak Togetherness - National Party, a year later the Supreme Court dissolved it, as its activities were in conflict with the Constitution of the Slovak Republic. The leader of the party was Marian Kotleba, who currently serves as the chairman of the political party People's Party Our Slovakia. In 2019 his party got into the Slovak parliament with more than 8%, in the presidential election in 2019 he occupied the fourth place with more than 10%. In 2019, a petition was filed to dissolve the party based on spread of racism and extremism both offline and online, however the court dismissed the action.

The main themes presented by right-wing extremists are the criticism of the Slovak government and the membership of the Slovak Republic in Euro-Atlantic structures. The number of right- wing extremists (especially skinheads and neo-Nazis) has been stable since 2004 - around 900 active persons and around 3000 supporters and sympathizers with a risk of growth.

Country profiles // Slovakia

Existing legal and institutional framework that address online and offline radicalization/ extremism

The Slovak legal system recognizes only offenses of extremism, criminal acts of extremism, extremist groups and extremist material. Radicalisation is not defined in the Penal Code. The concept of hate crimes is also not defined in the Slovak Criminal Code. However, the Criminal Code lays down a special motive that includes all crimes motivated by hatred towards any social group. A specific motive is to commit a crime with the intention of publicly inciting violence or hatred against a group of people or an individual because of their membership to a race, nation, nationality, colour, ethnic group, gender, religion or belief. The illegal conduct of right-wing extremists and their supporters can, in principle, include the following offenses under the Criminal Code: an offense committed on a specific motive (with the intention of publicly inciting violence or hatred against a group or individual for their race, a nation, nationality, skin colour, ethnic group, gender, origin or religion if it is a pretext for threatening for previous reasons), a crime committed with a specific motive (national, ethnic or racial hatred or hatred due to skin colour), support and promotion of movements to suppress citizens' rights and freedoms, production of extremist materials, dissemination of extremist materials, possession of extremist materials, defamation of a nation, race and belief, incitement to national, racial and ethnic hatred.

Online trends and behaviour (e.g. Internet access, demographics, most popular)

Each month, 2 500 000 Slovaks use Facebook, 970 000 Slovaks visit Instagram, 3 000 000 Slovaks use YouTube and only 350 000 Slovaks tweet. The average user of social networks in Slovakia is male, up to 24 years of age, with a higher education who lives in a household with income over EUR   1,100 and in a city with more than 100,000 inhabitants. 52% of the population uses social networks for private communication, presentation of opinions or making new contacts. For a further 35% these portals are for entertainment and approximately the same number of people (34%) use them to follow news and news or information about services and goods.

According to the statistics as of 2015, half of the perpetrators of crimes of extremism were aged 18-25. The most common way of committing offences was to publicly show support for various movements suppressing fundamental rights and freedoms, promoting these movements, or by defamation of race, nation and belief.

Country profiles // Spain

Current socio-political challenges that could contribute to online and offline radicalization

After the financial crisis, Spain has been the object of a polarization process in different variables of the political spectrum. From the extreme left and from national populism in some regions, but basically Cataluña. As a consequence of this, the extreme right has gained 52 seats in the Parliament in the last general elections (15.08%). This contributes to increasing the tension and creates a feedback loop in which people with moderate views lose ground. The "we against them" narrative is used as an argument by actors from all parts of the political spectrum. For the right wing "them" are the migrants, the feminists, and the Catalan separatists", for Catalan secessionists "them" is Spain as a whole in a context in which the idea is that they are paying with their taxes the other "lazy" people.

Current radical or extremist movements that engage in online or offline radicalization

The extreme right's main political party is Vox. Even though they operate within the limits of the constitutional and the legal systems their speech is focused on xenophobic and nationalist populism rhetoric. The ultra-right neo Nazis area there are very active in both recruitment and spreading their narratives to radicalized people -mainly youngster, which are their main target group. Football games and RAC (Rock Against Communism) music concerts are extremely popular and viral on YouTube and are also generating consistent online hate speech.

Country profiles // Spain

Existing legal and institutional framework that address online and offline radicalization/ extremism

Since 1995, Spanish Criminal Law includes several articles on hate speech. In 2015, the parliament updated Article 510 in a manner that goes beyond the recommendations of the framework decision on combating racism and xenophobia. It covers many categories of hate speech such as those targeting people based on their gender, sexual orientation, social status, ideology, national origin, disability, etc. Holocaust Denial has also been criminalized since 1995, but the provisions have been updated to make it technically compatible with some requirements regarding the limits of freedom of expression that the Constitutional Court issued in 2007. Some of the institutions that are enforcing the provisions are the Interior Ministry Office on Hate Crimes and a network of 52 Specialized Prosecutors on Hate Crimes (one per each province).

Online trends and behaviour (e.g.: Internet access, demographics, most popular [social] media)

92% of the Spanish population uses the internet. Up to 28 million of the 49.9 million people living in Spain are active on social media. The most used platforms are: YouTube (89%), WhatsApp (87%), Instagram (54%) and Twitter (49%). The average daily time spent online is 5h and 18 minutes, out of which 1h and 39 minutes on social media.