

Tech and Journalism Crisis and Emergency Mechanism (T&JM) - March 2023

The problem

At the start of Russia's invasion of Ukraine, in February 2022, major tech platforms didn't have enough information on who the country's and region's local trustworthy voices were. Despite the fact that platforms collaborated with international partners on numerous news trust initiatives and indicators, they lacked reliable mechanisms to identify professional and trusted local news sources in smaller markets. Ukrainian Public Broadcaster and Euromaidan Press, for example, were neither verified on Twitter nor recognised by Facebook as news publishers, whereas [Russia Today continued to earn advertising revenue on Facebook and YouTube until February 2022](#).

While technology companies have made significant efforts to combat the spread of dis- and misinformation, very few systems are in place to distinguish credible and trusted content creators, such as high-quality journalists and media organisations. At the time of [SempraMedia's Inflection Point International](#) report 62% of the media organisations interviewed¹ were not verified on Twitter and 64% were not verified on Facebook. Overall, only 35% of the media organisations said they had a point person to speak with in connection to the social platforms. Small and local media, investigative journalism organisations, journalists, NGOs, and other professional content creators [often face prohibitive content moderation practices when reporting on current events and topics of public interest](#). "The Chilling: A global study of online violence against women journalists" report found that the percentage of women journalists surveyed who reported online violence to Facebook is 39%, Twitter 26%, Instagram 16%. Also, at least 20% of online gendered violence incidents result in physical violence. Furthermore, many news outlets and journalists struggle to effectively protect their accounts, appeal bogus account suspensions, or to quickly restore wrongly removed or restricted content.

As they lack recognition, their content and accounts are often negatively affected by the platforms' current moderation systems and malicious actors. [Annex II](#) gives examples of such restrictions. In addition, recent [series of articles by Forbidden Stories](#) partners expose the insidious black market for silencing journalists and their stories. There aren't many, easy to access, procedures in place to provide early warnings, an urgent response, or channels for communication in crisis for trusted content and accounts. Without such a mechanism, attacks on journalists will persist and mis- and dis-information will continue to thrive.²

¹ The first study was based on 100 interviews conducted in Argentina, Brazil, Colombia, and Mexico. For the 2021 report, in addition to Latin America, eight more countries were added; 49 digital media organisations from Ghana, Kenya, Nigeria, and South Africa were interviewed; and another 52 from Indonesia, Philippines, Malaysia, and Thailand.

² Recognizing the problem, the European Commission proposed a regulation for establishing a common framework for media services in the internal market, [European Media Freedom Act](#), which aims to address the sector's specificities by including a section on media service provision in the digital environment. Article 17.5 emphasises the

On the other hand, there are users of digital platforms that enjoy privileges. Recommender systems, or algorithms that enable content moderation, benefit those who have achieved a privileged status or recognition. Facebook's XCheck programme has provided special treatment to celebrities, politicians, and other high-profile users, [shielding millions of VIP](#) users from the company's normal enforcement process. While XCheck grew to at least 5.8 million users by 2020, [only selected media companies in the most lucrative markets](#) are designated as news publishers on platforms.

Finally, [majority of content moderation efforts today focus on online speech-related harms and algorithmic moderation of all content](#), with only sporadic measures looking at affirmative action, safety protection and online recognition of credible actors and accounts. [Maria Ressa explained it perfectly with a metaphor](#): "When we focus only on content moderation – it's like there's a polluted river. We take a glass. We scoop out the water. We clean up the water, and dump it back in the river. However, what we have to do is to go all the way to the factory polluting the river, shut it down, and then resuscitate the river."

The proposed solution

Although some media outlets and journalism organisations have received recognition from the platforms, many others have been left behind. This has been especially true for small, independent, and investigative media outlets in nations and regions that aren't viewed as major tech markets. The invasion of Ukraine demonstrated that the current approach is not only ineffective but also potentially harmful to local media and journalists and overall information spaces online.

To address the current bottleneck approach and ensure that credible and professional voices continue to exist and operate freely in digital environments, GFMD proposes a multistakeholder process to establish the Tech and Journalism Crisis and Emergency Mechanism (T&JM), starting with the region of Ukraine and the neighbouring countries³. **This initiative aims to strengthen content and account moderation systems by establishing an emergency and crisis mechanism⁴ for journalists and media organisations and thus safeguard freedom of media and freedom of expression online. The project specifically targets small and medium-sized media, community and investigative journalism organisations and their professional communities.**

The key objective of the Tech and Journalism Crisis and Emergency Mechanism's (T&JM) is to engage journalism organisations, civil society, companies, and academics and experts to develop collaboratively:

1. **Processes and criteria for identification of credible and trusted journalism actors online, their communities and representative groups;**
2. **Crisis and emergency protocols, case escalation criteria, and functioning escalation channels;**
3. **Key elements and processes for establishing a voluntary multistakeholder an emergency and crisis mechanism.**

need for more data on platform interference with news media content by requesting that platforms publicly disclose the number of times they imposed any restriction or suspension on news content creators.

³ Initially focused on Ukraine and the neighbouring countries, this phase of the project is seen as a pilot for possible wider application in other regions.

⁴ This mechanism aims to build on previous efforts such as [Access Now's Declaration of principles for content and platform governance in times of crisis](#), the [Santa Clara Principles](#) and others, in line with the international standards on freedom of expression and other fundamental rights.

The process

The process of developing the T&JM mechanism is multistakeholder and consultative, with the mission of being locally driven and empowering local actors.

The first in a series of consultative meetings and gatherings to discuss the initiative's practical implementation, was **organised by the Global Forum for Media Development (GFMD) and the Baltic Centre for Media Excellence (BCME), in January 2023 in Riga, Latvia** with a particular focus on the actors and experiences from organisations working in Ukraine and the region, as well as exiled media due to the conflict. During the workshop, participants discussed how the initiative could work in practice: the definition of "recognition," the identification of trusted communities, and existing emergency/crisis protocols. The report and list of participants can be seen [here](#).

A second **consultation was held during UNESCO's Internet for Trust Conference in Paris on 21 February 2023, co-organized by GFMD and UCLA's Institute for Technology, Law and Policy**, seeking views from a wider range of partners on how the mechanism should be operationalized, how players should be involved, and what the expectations of such a mechanism were. [This report](#) presents the key takeaways and findings from the consultation.

The roadmap

Draft project roadmap for developing a Tech and Journalism Crisis and Emergency Mechanism (T&JM) includes future consultations in smaller working groups to address the implementation of the pilot, including:

- 1. Define key elements and processes for establishing a voluntary multistakeholder emergency and crisis mechanism.**
 - a. Define the purpose and goals of T&JM:
 - i. Define T&JM's mission and vision and its intended outcomes.
 - b. Determine the scope and composition of organisations setting up T&JM:
 - i. Define the scope of the T&JM, including the platforms that will be included and the stakeholders to be involved and their roles.
 - c. Define and establish the governance framework:
 - i. Develop a governance framework that outlines the roles, responsibilities, and decision-making processes and criteria.
 - ii. Develop transparency, accountability and risk assessment processes and reporting
 - d. Monitoring and evaluation of work of T&JM:
 - i. Establish metrics and KPIs for T&JM performance, and regularly monitor and evaluate the T&JM performance against these benchmarks.
 - e. Optimisation of the mechanism:

- i. Continuously refine and optimise the T&JM practices based on performance feedback, emerging trends, and changing needs.
- 2. Processes and criteria for identification of credible and trusted journalism actors online, their communities and representative groups;**
 - a. Technical expertise and main elements of the mechanism:
 - i. Existing standards and information integrity approaches that could provide a framework or guidance
 - ii. Defining external references, such as news integrity and trust initiatives, professional and ethical self-regulation bodies, and donor and funder audits, that could be used to identify credible and trusted journalism actors online;
 - iii. Defining processes and minimum criteria for demonstrating adherence to professional and ethical journalistic standards;
 - b. Collection of data/evidence around the treatment of credible and trusted journalism actors online, their communities and representative groups
 - i. Identify research to inform the efforts and guide future technical and policy decisions around the identification and relevant issues.
- 3. Crisis and emergency protocols, case escalation criteria, and functioning escalation channels;**
 - a. Create policies and procedures for crisis and emergency escalation and management of cases:
 - i. Develop policies and procedures for crisis and emergency escalation and management of cases that align with the scope and governance framework.
 - b. Develop case management and escalation and monitoring tools and technologies:
 - i. Establish escalation communication channels, procedures and responses from platforms;
 - ii. Develop and implement tools and technologies that can facilitate T&JM work.
 - c. Train T&JM members on policies, processes and procedures:
 - i. Provide training and support to members.

About T&JM

Tech and Journalism Crisis and Emergency Mechanism (T&JM) will **act as a forum where all stakeholders, journalism organisations, civil society organisations, companies and states can discuss, develop recommendations and adopt crisis and emergency protocols, best practices in transparency reporting and account moderation of journalism and media organisations.** It will use a voluntary-compliance approach where all stakeholders sign up to a model that does not create legal obligations and where they voluntarily implement the T&JM's decisions and recommendations.

The T&JM will be a multi-stakeholder process where representatives of the various stakeholders come together to improve the practices of the sector. The regional mechanism should involve and should be led by local and regional actors.

- **Cross-industry collaboration:** Engage a leading group of local and regional organisations representing professional communities of ethical journalism organisations, digital and journalism experts, academics as well as technology companies, and facilitate multi-stakeholder conversations. It will ensure multistakeholderism, open partnerships, and avoid duplication of efforts.
- **Solutions:** Private sector companies, journalists, media and their professional communities, academics, experts, international multilateral organisations and civil society collaborate to share best practices as they develop and implement new solutions through a joint cross-industry mechanism.
- **Civil society engagement and empowerment:** Empower journalism and media communities to participate in policy forums in order to support and promote journalism as a public good. The mechanism should involve and should be led by local and regional actors. It will also link the initiative with international and regional players to collect data, advocate, monitor policies, and supervise the proper functioning of the mechanism.
- **Knowledge, data sharing and transparency:** We work towards the goal that no knowledge and insights are lost, and that experiences and lessons learned at all levels are collected and processed – from global to local, directly from the field as well as in academia.

About GFMD's work in this area

Global Forum for Media Development's (GFMD) core mandate is to promote policies, programmes, strategies, and opportunities for the work of the journalism and media support sector in order to enhance journalism as a public good. We are actively supporting and promoting reputable journalism and civil society organisations on digital platforms as part of our ongoing [work to strengthen digital information ecosystems](#). We are collaborating with Twitter's and Meta's public policy teams to verify the accounts of journalists, news media, and journalism support organisations, particularly those in Ukraine and Eastern Europe. So far, this process has taken the form of ad hoc, informal communication between policy teams at tech platforms and civil society groups like GFMD.

Glossary

Crisis/emergency situations	Following Access Now's Declaration we refer to crisis or emergency situations following the EU Commission recommendation as: areas or situations " <i>in a state of armed conflict, fragile post-conflict areas, as well as areas witnessing weak or non-existing governance and security, such as failed states, and widespread and systematic violations of international law, including human rights abuses</i> "
Tech and Journalism Crisis and Emergency (T&JM)	This multistakeholder mechanism aims to recognise professional communities of practice in the content creators' industry, such as journalists and media organisations, civil rights organisations, academics and other content creators, in order to ensure that credible and professional voices can operate freely in digital settings.
Professional/Trusted Communities of Practice	This category includes associations and groups that share a profession (such as journalists, academics, civil rights defendants, etc.) and comply with the ethical and professional standards of their work.
Trusted Content Creators	Based on the definition of the "Professional/Trusted Communities of Practice" and the concept of 'Trusted Flagger', professionals who are part of these communities of practice and produce online content, adhering to ethical and professional standards, shall be considered Trusted Content Creators for the purposes of this concept note. The term "Trusted Content Creators was coined by scholar Martin Husovec, in an homonymous blogpost in 2021, accessible at: https://husovec.eu/2021/10/trusted-content-creators/
Trusted Flaggers	In the recently adopted EU regulation on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC, EU policymakers took a major step towards recognising communities of practice by creating a system to designate "Trusted Flaggers." Although the number of "trusted flaggers" will be limited, "industry associations representing their members' interests should apply for the status of trusted flaggers," according to the latest version of the DSA text (Recital 46). Furthermore, the DSA allows online platforms to use trusted flaggers "or similar mechanisms to take quick and reliable action against content that is incompatible with their terms and conditions".

Resources and Literature review

Article 19 (2022) [EU: Will the Digital Services Act hold Big Tech to account? - ARTICLE 19](#) (July 05, 2022)

Center for Democracy & Technology (2022) *CDT Comments to the UK Parliament on the UK Online Safety Bill*; June 29, 2022. Available at:
<https://cdt.org/wp-content/uploads/2022/06/2022-06-30-CDT-OSB-Comments-FINAL.pdf>

Digital Security Lab (2022), *Digital Security in the Ukrainian Civil Society Sector during the War* (October 14, 2022). Accessible at: <https://dslua.org/publications/report-digital-security-ukr/>

Frosio, Giancarlo & Husovec, Martin (2020), Accountability and Responsibility of Online Intermediaries, in Giancarlo Frosio (ed.), *The Oxford Handbook of Online Intermediary Liability* (Oxford University Press, 2020), Available at SSRN: <https://ssrn.com/abstract=3451220> or <http://dx.doi.org/10.2139/ssrn.3451220>

Granger, Jacob (2022) [Google News Showcase: Is it viable for the smallest of newsrooms?](#) *Journalism.co.uk*; (July 7, 2022)

Horwitz, Jeff (2021) [Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt](#), *The Wall Street Journal*, (September 13, 2021)

Husovec, Martin, (2021a). *(Ir)Responsible Legislature? Speech Risks under the EU's Rules on Delegated Digital Enforcement* (September 17, 2021). Available at SSRN: <https://ssrn.com/abstract=3784149> or <http://dx.doi.org/10.2139/ssrn.3784149>

Husovec, Martin (2022) *Trusted Content Creators*. LSE Law - Policy Briefing Paper No. 52, Available at SSRN: <https://ssrn.com/abstract=4290917> or <http://dx.doi.org/10.2139/ssrn.4290917>

Husovec, Martin & Roche Laguna, Irene (2022) *Digital Services Act: A Short Primer*, in Martin Husovec and Irene Roche Laguna, *Principles of the Digital Services Act* (Oxford University Press, Forthcoming 2023), Available at SSRN: <https://ssrn.com/abstract=4153796> or <http://dx.doi.org/10.2139/ssrn.4153796>

Radsch, C. (July, 2022) *Making Big Tech Pay for the News They Use*. Washington DC: Center for International Media Assistance. Available at:
<https://www.cima.ned.org/publication/making-big-tech-pay-for-the-news-they-use/>

Ponsford, Dominic (2017) [How the rise of online ads has prompted a 70 per cent cut in journalist numbers at big UK regional dailies](#), *Press Gazette*, (January 24, 2017)

Reporters without Borders (2023) Concept note: [The JTI emergency protocol](#)

Reporters without Borders (2022) [Ukrainian journalism – collateral victim of Big Tech measures | RSE](#)

Strengthened Code of Practice on Disinformation (2022) [The 2022 Code of Practice on Disinformation | Shaping Europe's digital future, 2022 Strengthened Code of Practice on Disinformation | Shaping Europe's digital future](#) (June 16, 2022)

Waterson, Jim (2019) [YouTube blocks history teachers uploading archive videos of Hitler](#), *The Guardian* (June 6, 2019)

Annex I: Monitoring organisations

[7or - 7amleh | Report Violations on Palestine](#)

The Palestinian Observatory of Digital Rights Violations (7or) is the first open-source online platform to monitor, document and follow up on the digital rights violations of Palestinians. The platform offers a tool to report violations and search through the database of violations.

[Digital Monitoring Database - BIRN and SHARE Foundation](#)

BIRN and SHARE Foundation have partnered to monitor digital threats and trends, raise awareness about violations of digital freedom and issue policy recommendations. The Digital Monitoring database covers the state of digital rights in Bosnia and Herzegovina, Croatia, North Macedonia, Serbia, Hungary, and Romania. It documents violations of digital rights and freedoms, with descriptions of the cases and corresponding sources in order to provide facts about the violations, and emphasises the use of technology to highlight social issues in targeted countries. The database categorises cases according to the type of violation, means of attacks, as well as the actors, including both the affected parties and attackers. This is important both for understanding the social context and for monitoring trends in violations of digital rights and freedoms.

[Lumen Database](#)

The Lumen database collects and analyses legal complaints and requests for removal of online materials, helping Internet users to know their rights and understand the law. These data enable us to study the prevalence of legal threats and let Internet users see the source of content removals.

[Mapping Media Freedom - European Centre for Press and Media Freedom](#)

Mapping Media Freedom is a crowd-sourced platform that enables anyone to upload an alert that relates to threats against journalists and media workers across Europe. The platform allows to map and visualise threats, violations and attacks that restrict media freedom. This includes physical attacks, threats of violence made online and offline, legal actions aimed at silencing critical coverage and moves to block access to independent media platforms. All these actions, and many more, restrict media freedom and the Mapping Media Freedom platform helps us gauge the health of the media environment across Europe.

[Mnemonic](#)

Mnemonic gathers, analyses, documents and archives digital content exposing human rights violations. In the context of conflict, they support the establishment of infrastructures to archive the content. Since content is also removed by digital platforms, they offer help reinstating this valuable information online.

Annex II: Cases

Algeria:

- **Facebook blocks the accounts of a number of civil society leaders and protesters involved in mass demonstrations.**

In the context of "the Hirak" in 2019, a popular movement calling for democracy in Algeria, a number of accounts belonging to opposition activists were blocked by Facebook. Civil society organisations as well as independent journalists are active users of social media, with some using it as an alternative platform to traditional media that are otherwise subject to pressure from the Algerian state. In response, protesters gathered in Facebook's headquarters of a number of Western capitals to denounce what one article in the Algerian daily *El Watan* called indirect "censorship" of opposition activists. The latter accused pro-government individuals of instrumentalising Facebook's anti-harassment and hate speech policies to have their posts taken down.

Sources:

<https://www.elwatan.com/a-la-une/comptes-bloques-ou-signales-facebook-censure-t-il-les-opposants-algeriens-21-11-2019>

<https://www.lopinion.fr/international/en-algerie-la-contestation-sur-les-reseaux-sociaux-bo-uscule-la-censure>

<https://www.elwatan.com/edition/actualite/la-presse-en-quete-de-liberte-01-03-2019>

Belgium:

- **Facebook blocks two Euractiv posts that describe Meta's fight against EU data legislation**

On 4 and 5 February 2022, Facebook blocked two posts of the European media network Euractiv, based in Brussels. The blocked posts were stories written by journalist Luca Bertuzzi, who covers digital policy and tech politics for the website. In the specific articles, Bertuzzi investigated how the tech industry was preparing to fight the Data Act, which is an upcoming EU legislative proposal aimed to create binding rules for data sharing. Facebook's mother company Meta is part of several associations behind the initiative. Bertuzzi announced the blockage on Twitter, which was retweeted more than 100 times (retweets and cited tweets combined). He added: "I can't help to think that it shouldn't be about who you know or how much noise you make to get this kind of correction." After the criticism, Facebook republished the stories. Source: <https://www.mapmf.org/alert/24601>

- **Current affairs TV content removed from Facebook**

In March 2021, the Belgian French-speaking broadcaster RTBF posted a current affairs item on Facebook. The video reported on police violence with a possible racist motive in the city of Antwerp that happened in December 2020. Facebook explained that the removal of the report was intended to protect minors.

Bosnia and Herzegovina:

- **Facebook Blocks Page of Bosnian Media Outlet**

The Facebook page of the Bosnian online outlet Direkt-portal.com was blocked for about a month. After the editors of this portal asked for an explanation from the tech giant, it replied that the activities of the site

“do not respect the guidelines regarding false representation and imitation”. The portal is still blocked.
Source: <https://monitoring.bird.tools/data?caseId=556&installation=https:%2F%2Fbih.bird.tools>

China

- **On Chinese social media, censors ban the word “Beijing”, “Xi Jinping” in the lead up to the 2022 Party Congress**

Two recent examples from China illustrate the dynamics of platform interference on the part of social media companies close to the state. On the Chinese social media platforms Baidu and Weibo, state censors banned the words “Beijing” and “Xi Jinping” in the run up to the 20th Party Congress. On Douyin (a version of TikTok that is owned by the same company), as well as on WeChat and Zhihu.com, one tends to find only content produced by state media or organisations close to the government.

Sources:

<https://www.msn.com/en-in/news/world/xi-jinping-has-totally-muzzled-chinese-social-media-how-you-can-t-search-his-name-on-any-platform/ar-AA130QdL>

<https://www.bloomberg.com/news/articles/2022-10-14/china-censors-beijing-after-rare-xi-protest-before-congress>

Colombia:

- During the 2021 protests, some contents related to it were taken down on Instagram: from graphic and explicit images and videos that showed violence to graffiti and poems that addressed the social upheaval. It not only brought up the question of governmental efforts to silence protestors, but also the limits and problems of content moderation in the context of conflict. Instagram’s mother platform Facebook, now Meta, answered back then that it was not a targeted action but rather a global problem with the platform.

After doing field research in the country, the Inter-American Commission on Human Rights concluded that one of the causes for this was that the Colombian public forces conducted digital surveillance and target certain contents as digital terrorism, after fact-checking it themselves. In doing so, the experts reported the content which were taken down by the platforms.

Sources:

<https://www.elespectador.com/judicial/la-censura-digital-el-paro-nacional-y-lo-que-dijo-la-cidh-sobre-el-internet/#>

https://www.oas.org/es/cidh/informes/pdfs/ObservacionesVisita_cidh_Colombia_spA.pdf

<https://www.elespectador.com/tecnologia/que-paso-con-las-publicaciones-en-instagram-durante-el-paro-nacional/>

Denmark:

- **App with TV-content for children (Ramasjang) removed from Google Play app-store**

The Danish Broadcasting Corporation DR encountered issues with Google over the removal and blocking of updates for the Ramasjang app, giving access to children’s content. After ongoing differences over some time, Google decided to remove the app in full in August 2020, without providing DR with further

information or a reason for their decision. After negotiations, followed by an intense public debate about the need to guarantee editorial independence to content providers on online platforms, the app was reinstated in the original form by Google, without further information on change of stance. Source: <https://www.dr.dk/nyheder/kultur/google-slettede-drs-app-til-boern-nu-boejer-techgiganten-sig>

France:

- **Snapchat's community guidelines limit dissemination of sexual education content to young people**

Following new community guidelines published by Snapchat in November 2019, France Télévisions faces problems about the age limitation of sexual education content. The change of Snapchat's community guidelines leads to a situation where France Télévisions' content showing visuals of contraception products (i.e., the pill) and written information on contraception methods is categorized by Snapchat as sexually explicit content "prohibited to under 18s". France Télévisions thus must adapt their content if they want it to reach the audience for which it is intended. The content is originally part of a series of content items called "Sexy soucis" and disseminated via France Télévisions' own online service for young people "Slash".

Germany:

- **Blocking video on terrorist attack in Hanau.**

On 23 February 2021, Facebook and Instagram blocked access to a programme of the German regional public broadcaster WDR (member of ARD). The video, produced for the current affairs programme "Monitor", reported on the terrorist attack in the German city of Hanau, which happened one year ago. The online platforms removed the content item based on their community standards, without giving any further explanations to the broadcaster

Greece:

- **Journalist' account temporarily suspended**

On 12 January 2022, journalist Chris Avramidis was suspended from Twitter, after he reported on mass protests at the Athens University of Economics and Business (ASOEE), during which he denounced the University Police and its invasion of the University. After the suspension, users urgently asked for Avramidis' return with the hashtag #Free_chris_avramidis, stating that any account that expresses an opinion that bothers the government becomes the target of online attacks. That hashtag became Trending Topic on Twitter. Many users also stated the suspension was not an individual but a systematic targeting of censorship of those who succeed against government policy. After a day, Avramidis' account was restored, which the journalist himself announced in a tweet in which he thanked those who protested. "Welcome back. The pressure has paid off. Thanks for your solidarity", he wrote on Twitter, adding: "The problem is not personal, but political. It does not matter if the person of the removed account is named Christos or if it's a waiter named Babi." Source: <https://www.mapmf.org/alert/24687>

Hungary:

- **Facebook Blocks Hungarian Historian's Page**

Facebook has blocked the educational page of the historian Krisztián Ungváry, after a far-right media outlet, Szent Korona Rádió, reported it for hate speech.

Moldova:

- **18 years old Independent Investigative Media in Moldova identified by Facebook as a Russian state media**

Although Ziarul de Garda had already been verified by Facebook and it was one of the first media in Moldova that was blocked by the Russian government when the war started, in July 2022 Facebook started to display a warning message identifying the media outlet as being controlled by the Russian state. This warning included misleading information about how Ziarul de Garda was considered Russian affiliated media by Facebook's own policies and definitions, damaging the activity of the media outlet and negatively impacting on the investigative reporting in Moldova.

Pakistan:

- **Facebook Penalises Accounts of Journalists and News Media Outlets**

Pakistani journalists and digital news media outlets claimed that Facebook permanently disabled, restricted or temporarily suspended the Facebook accounts and pages of some journalists and media outlets after they posted news and commentary about the August 2021 Taliban takeover of Kabul. The posts had reportedly condemned the Taliban takeover, but were still found by Facebook's content moderation system to have violated community standards related to dangerous organisations. The Digital Media Alliance of Pakistan (DigiMAP) issued a public statement urging Facebook to reconsider its policy regarding journalists and media outlets in Pakistan reporting critically on the Taliban in the public interest.

<https://slate.com/technology/2021/09/facebook-twitter-taliban-social-media-content-moderation.html>

<https://digimap.pk/digital-media-alliance-asks-facebook-to-reconsider-its-policy-regarding-journalists-and-media-outlets/>

Palestine:

- **Ongoing Silencing of Israeli Assault Coverage**

Palestinian activists have claimed that social media giants have engaged in digital censorship during the latest Israeli assault on occupied Gaza that left 44 dead, including 15 children. As Palestinians in Gaza were attacked by Israeli rockets, in the digital space, Palestinian rights activists grappled online with social media censorship. Source:

<https://www.middleeasteye.net/news/israel-palestine-gaza-social-media-companies-accusations-silencing-voices>

Romania:

- **Facebook Penalises Website for Using Romanian Dictator's Name**

Pagina de Media, an online publication that covers media industry news, was sanctioned by Facebook for publishing the name of Ion Antonescu, the antisemitic military dictator who ruled Romania during WWII. The punishment consists of a 30-day ban on livestreams and paid advertisement. But the publication complained that it was unjustly punished for reporting impartially on a story about a local TV station which was fined for broadcasting a remembrance ceremony for Ion Antonescu. This isn't the first time that Facebook has penalised Romanian journalists for simply mentioning the name of the former dictator with no regard to the context. In June, the social media network deleted a post by a local journalist who wrote about the renaming of a street that used to bear Ion Antonescu's name. Source: <https://archive.ph/dy12X>

- **Facebook Disables Pages of Major Romanian News Outlets**

Romanian news portal Hotnews.ro said that Facebook had arbitrarily restricted access to its page on the social network without giving it the right to appeal the decision. The partial ban lasted a month-and-a-half

and reduced traffic by 96 per cent. It was motivated by an apparently uncontroversial article on the 77th anniversary of the Allied landings in Normandy. References to Hitler might have prompted the decision. Another prominent Romanian news portal, G4Media, also said its Facebook account had been arbitrarily restricted for weeks. Source: <https://monitoring.bird.tools/data?caseId=1259&installation=https:%2F%2Fro.bird.tools>

Serbia:

- **Twitter Deletes Video about Unsolved Murder Attempt on Journalist**

Twitter deleted a video published by the Independent Journalists' Association of Serbia, NUNS, citing violations of platform rules. The video contained an interview with the daughter of Dejan Anastasijević, the late Serbian journalist who was the target of a bomb attack more than 10 years ago, which remains unsolved. The video was reposted on Twitter the next day in a new tweet. Source: <https://monitoring.bird.tools/data?caseId=3319&installation=https:%2F%2Frs.bird.tools>

- **NGO Twitter Account Disabled**

The Twitter account of Partners for Democratic Change Serbia, a human rights NGO, was disabled for unknown reasons, as the platform did not provide any information, and they had to create another account. Source: <https://monitoring.bird.tools/data?caseId=3332&installation=https:%2F%2Frs.bird.tools>

Spain:

- **Content restricted by YouTube from a Spanish documentary about the Holocaust**

In August 2021 YouTube restricted a documentary by a Spanish journalist, Carlos Hernández, due to breaching their internal policy on violent content and nudity. Despite the documentary aimed to inform students about the reality of concentration camps, where many Spanish were imprisoned, and educate future generations about the risks of fascism, YouTube considered the images were too explicit and inappropriate for teenagers. Source:

https://www.eldiario.es/sociedad/son-imagenes-holocausto-youtube-censuro-inapropiadas-jovenes_1_8287177.html

A similar case arose in 2019 when YouTube's new policy on hate speech made unavailable archive material related to Adolf Hitler that teachers used in their history lectures about the second world war. Source:

<https://www.theguardian.com/technology/2019/jun/06/youtube-blocks-history-teachers-uploading-archive-videos-of-hitler>

Switzerland:

- **Content removed on COVID-19 and Russian police repression.**

In March 2021, YouTube removed two videos produced by the Swiss broadcaster RTS for its programme Géopolitis. A video dealing with the Covid-19 situation in China was taken down by the platform without further explanation. The other reported on police repression in Russia and was classified by YouTube to be appropriate to users of 18 years and older only.

Turkey:

- **Kurdish news agency's Instagram account blocked**

Pro-Kurdish news agency Jin News's account on Instagram -@jinnewshaber- was removed by the platform management with the explanation that the news agency's publications violated the community rules. Source: <https://www.mapmf.org/alert/24490>

- **Mesopotamia News Agency's verified Facebook page removed**

On December 26, 2021, Mesopotamia News Agency's verified Facebook page which had over a hundred thousand followers and had been restricted multiple times previously, was removed by Facebook despite the news agency's repeated complaints against the platform. News agency which publishes critical information, announced that they would continue sharing their news on a new page however also reiterating that Facebook had been ignoring their plea to end censorship against articles and arbitrary restrictions against the platform's Facebook page. Source: <https://www.mapmf.org/alert/24491>

United Kingdom: Novara Media YouTube channel temporarily deleted

On 26 October 2021, the YouTube channel of the alternative media outlet “Novara Media” was temporarily deleted. A statement from Novara Media said: “This morning the Novara Media YouTube channel was deleted without warning or explanation. Novara Media is among the top 50 most watched news and politics channels in the UK.” Their YouTube channel has over 167,000 subscribers. The media group is an independent, left-wing news channel founded in 2011 and it is regulated by IMPRESS. Later on the same day, it was reinstated. “We are now demanding a step-by-step explanation as to how this happened,” Novara Media said in a consequent statement. “Novara Media’s channel was briefly removed after it was flagged, but upon review, it was then immediately reinstated,” A YouTube spokesperson said, “We work quickly to review all flagged content, but with millions of hours of video uploaded on YouTube every day, on occasion we make the wrong call.” Although it is common that videos are removed for violating YouTube policies, it is unusual for a media channel to be pulled. YouTube also had not warned Novara Media beforehand or followed a three-strikes-policy. YouTube did not give details as to why the account was initially flagged. Source: <https://www.mapmf.org/alert/24323>

Georgia

TV Formula’s received a warning from Facebook after its algorithm flagged a quote from Ukrainian President Volodymyr Zelensky as hate speech. The post was merely a translation into Georgian of a post by Zelensky himself.

<https://www.rappler.com/technology/features/ukraine-facebook-battle-coda-story/>