Analysis

Policing the internet: how Europol takes action against undesirable content online

Kilian Vieth

Translation by Viktoria Langer

July 2017

This analysis was originally published in German in CILIP 112 (March 2017): Alles anti-terror? It is adapted from a masters' thesis, 'Europol Policing the Web: Internet Content & Counter-Radicalization – An Interpretive Policy Analysis Approach'. The author can be contacted at kvieth@stiftung-nv.de.

Europol removes content from the internet. This approach goes beyond regular measures in the fight against terrorism propaganda and mixes police work and media regulation. Should a police agency be responsible for the surveillance and control of Facebook posts and tweets?

Since summer 2015 Europol has maintained an Internet Referral Unit (EU IRU). The IRU searches the internet to analyse and assess content, which is assessed against the provisions of the EU Directive on combating terrorism (2017/541) and may then be rated as inappropriate or dubious. The content is then submitted to the website provider and recommended for removal. This is to prevent the spread of online propaganda and counter radicalisation via the internet. The IRU is deliberately trying to make internet groups aware of online material that does not comply with the terms and conditions of the respective platform (such as Facebook or Twitter). Therefore, the IRU acts as an intersection of private and police media regulation.

The establishment of the new Europol unit was decided in March 2015, only two months after the attacks in Paris on the editorial department of Charlie Hebdo and Jewish supermarket Hypercache, and was based on a political response at the highest level. The European Council called for an EU unit to fight online propaganda in a statement on 30 January 2015, which says that the internet "plays an important role in radicalisation processes" and that "the removal of terrorist and extremist content needs to be enforced". [1] The fast-track implementation of security measures follows a well-established pattern. Security policy

---

[1] EUCO 18/15 (23.04.2015)

players often find themselves under the pressure of demonstrating capacities to act after such attacks, even if it mostly involves symbolic measures.

The IRU builds on the 'Check the Web' project, which was established by the German government. This effort analyses terrorist propaganda online and collects it in a central database which is accessible to all Member States. From its introduction in 2007 until 2015 about 10,000 documents and individuals were registered. [2] However, this includes only the surveillance and analysis of propaganda material; deletion recommendations for providers were not part of 'Check the Web'. The UK was the first Member State to put into practice police measures to fight radicalisation on the internet within its preventive approach against terrorism. The British Counter-Terrorism Internet Referral Unit (CTIRU) was the institutional model for the EU IRU. [3] Since 2010 the British unit has worked on filtering unwanted online content.

Surveillance and deletion

Europol's aim is to reduce access to terrorist and extremist material on the internet with the help of the IRU. The content can include texts, images and videos, but also whole social media accounts or profiles which spread the content in question. So far the unit's work focuses on the monitoring of big platforms like Facebook, YouTube and Twitter, and sites and platforms where propaganda is distributed. Hereby, accounts that translate contents or those which have a vast reach have priority. [4] The Europol unit searches and finds a major part of the content to be removed on its own, but it also accepts content that has been submitted by Member States' authorities and forwarded to the platform. This coordination feature is supposed to avoid multiple requests, i.e. when a specific account is being monitored by a Member State to gather information, another Member State should not be able request the removal of this account. At the same time Europol offers the IRU as a mediator for the Member States which do not have their own unit. However, Europol does not directly accept information provided by citizens; cooperation takes place only between authorities and internet companies (although individuals may be able to submit information to national authorities, as is the case with the UK's CTIRU).

As far as known the IRU has no privileged access to online platforms. It states that it exploits only those possibilities which are open to all citizens: to report content to social media providers. How exactly communication between the EU IRU and the providers works remains unclear, but there is no doubt that this cooperation is to be expanded. For this purpose, the European Commission has initiated the 'EU Internet Forum', to improve the dialogue and the alliance with platform providers such as Google, Facebook and Microsoft. [5]

Far-reaching dual mandate

Even though the decision to introduce a European IRU is clearly connected to the Paris attacks in January 2015, the legal competences of the new unit go beyond the field of counter-terrorism: the EU IRU is not only allowed to monitor and remove "terrorist and extremist" content but also content that connected to "illegal immigration" and the "smuggling of migrants". In April 2015 the European Council decided to include the location and removal of

---

[2] Council document 7266/15 (16.03.2015) p. 3
[3] Council document 1035/15 (17.1.2015) p. 3
[4] Council document 14244/15 (23.11.2015) p. 11
[5] COM (2016) 230 final version 24.04.2016, p. 7

"internet content used by traffickers to attract migrants and refugees" into Europol's mission, in accordance with the respective national constitutions. [6] With this extension of the mandate prior to the actual establishment of the Referral Unit an additional and quite vague legal basis for the removal of information was created. Thus from the start the competences of the EU IRU included the surveillance of online activities of smugglers and traffickers. [7] The Internet Referral Unit therefore cooperates with Europol's Counter-Terrorism Centre (ECTC) and the European Migrant Smuggling Centre (EMSC), which is also hosted by Europol.

The most recent public statistics about the EU IRU's work were published in a report in December 2016. Here it says the Referral Unit had requested 15,421 removals up to October 2016. [8] The content involved was spread across at least 31 platforms and included contributions in eight different languages. [9] 88.9% of the reported content was removed. Due to the unavailability of comparable figures this "success rate" is hard to assess and revise. The interesting part, however, is the amount of reported content that was not removed from the platforms even though Europol had submitted a deletion recommendation, as the crucial question is not about how much content has been removed but on what criteria removal is based.

Therefore, it is worth looking at this process more closely. It is important to know that the contents in question are evaluated in two separated steps. The first evaluation is made by Europol's IRU, which classifies single contributions or accounts as terrorist propaganda or "content used by traffickers to attract migrants and refugees". [10] The content is forwarded to the provider hosting the content, which makes a second evaluation. How and according to which criteria the second evaluation takes place is subject to the company. Usually the companies hereby refer to their community directives (e.g. terms of service). Europol stresses they do not have any influence on final decisions on the removal of content. The agency has no executive powers and while it can report the content, it is not legally allowed to remove it.

Both evaluation processes – Europol's and the platform's – are completely opaque and cannot be assessed in any way. There is no judicial or parliamentary control and no supervision of an ombudsperson. The users have no right of objection or rights of disclosure. The responsibility for removing content is handed over entirely to the companies.

Deliberate privatisation of policies

Not one of these evaluation processes is based on any statutory rules, but simply on corporate community guidelines. Even the IRU goes by companies' community guidelines when evaluating content. This is important because this arrangement is not a random one. On the contrary, one of the reasons for establishing the IRU was that this way more content can be removed than legal requirements would allow.

At the beginning of 2015 the EU Counter-Terrorism Coordinator (CTC) declared in a statement that it would be reasonable to go by the company's terms and conditions because this makes it possible to remove more content than most of the European and national legal frameworks would allow. In the words of the CTC, companies' terms and conditions "often go further than

[6] EUCO 18/15 (23.04.2015)
[7] BT-Drs. 18/9764 (26.09.2016)
[8] Council document 14260/16 (20.12.2016) p. 22
[9] Europol: EU Internet Referral Unit. Year One Report Highlights (22.07.2016), http://www.statewatch.org/news/2016/sep/eu-iru-one-year.htm
[10] EUCO 18/15 (23.04.2015)

national legislation and can therefore help to reduce the amount of radicalising material available online." [11]

What is at stake here is the fundamental right of freedom of speech and the question of who defines the rules for it. Which content is prohibited differs in different countries and on different platforms. Defining the limits of freedom of speech is always a political decision, which in a free society should not be made by opaque, technocratic processes, but through open and democratic debates. One example is the display of female nipples: Instagram and Facebook remove these images instantly, even though in Germany naked breasts in public are not explicitly forbidden. Whereas the display of Nazis symbols is prohibited in Germany, in other countries they can partly be displayed. How freedom of speech is interpreted cannot be generalised and is always subject to a social and political context.

How does Europol make this complex decision about the classification of content? Here, the agency remains silent and refers vaguely to its legal basis. However, this is probably not giving precise criteria for the complex assessment and case-by-case decisions. Europol has quite some freedom when it comes to the categorisation of undesirable content.

An effective approach?

Europol was established to improve the information exchange and the cooperation between national police authorities and to increase the efficiency of European security policy. However, the increasing cooperation with private platforms should be observed critically, because it can already be seen that having come to Europol from the UK, the IRU model is likely to be re-applied on national level to regulate media content. 26 Member States have established national contact points linked to the EU IRU, even though Europol itself admits that the IRU's approach is not very effective. The removal of internet content is like a cat-and-mouse game: if something gets removed it reappears elsewhere and often multiple times. [12] The IRU cannot do a thing about this basic problem, known as the 'Streisand effect'. The removal remains reactive and limited in its effects, also because submitters of propaganda material adapt quickly to the removals. Even if Europol requests more resources for the Referral Unit, it remains unclear if this problem can be solved by more staff and better hardware and software. Europol is probably going to continue the cat-and-mouse game, but with bigger computers, because the capacities and methods of those uploading propaganda are continuously developing.

One approach to make the IRU's work more efficient is the so-called upload filter. These filters are supposed to help detect content that has been removed before, but is slightly altered and uploaded again. In the long run Europol wants to drop its reactive role and enter into prognosis. The 'abuse' of social media is then going to be anticipated and the spreading of terrorist propaganda is going to be prevented in advance. [13] This 'vision' is still up in the air, but its realisation should not be excluded. The preventive filtering of content is technically possible and the use of prognosis technology for police work is a current trend. [14]

---

[11] Council document 1035/15 (17.01.2015) p. 3
[12] Council document 7266/15 (16.03.2015) p. 4
[13] Council document 7266/15 (16.03.2015), p. 5
[14] See, for example, the section 'Empirical perspectives on big data' in the edited collection Exploring the Boundaries of big Data, Netherlands Scientific Council for Government Policy, 2016, https://www.ivir.nl/publicaties/download/1764

This is why we should not only ask about technical efficiency but also about political and social efficiency, because the removal of internet content only treats the symptoms. The security authorities have to count themselves in: if content is removed, which message does that send to the submitter? Does that not enforce self-referential filter bubbles and conspiracy ideologies?

If content gets removed it has to be made sure that there are universal standards for everyone. The work of the EU IRU so far is strongly focused on Islamism and the radicalisation via jihadist ideologies. When it comes to right-wing terrorism in the Ukraine, for example a lot less attention is being paid to the removal of content. To fight "illegal migration" with the help of the IRU goes beyond the usual competences of a police authority as well. The establishment of the IRU was meant to fight terrorism and radicalisation. A convincing narrative: who wants to criticize the removal of brutal videos showing beheadings? However, as soon as the removal infrastructure is in place its area of application can be expanded step-by-step.

Even if the final decision about content removal stays with the platforms the EU IRU is no symbolic initiative. It changes the interaction of public and private actors concerning the regulation of media content and rearranges the criteria. The question is not whether a specific content is legal, but if it is desirable from the perspective of commercial and security policy. Even if the comparison falls short, if the IRU were responsible for newspapers rather than Facebook and Youtube, there would be an immense outcry about the political influence over the media.