

Navigating content moderation

Solutions for Hosting Service Providers

This fact sheet is part of the FRISCO project’s comprehensive offering of tools and resources supporting Hosting Service Providers (HSPs) to develop effective responses to terrorist content online and fulfil their obligations under the **EU’s TCO Regulation**.¹ This fact sheet focusses on **content moderation** and why it matters for HSPs. The **FRISCO Content Moderation Tool** and the **FRISCO Brochure on technical solutions** presented here aim to support online platforms in choosing the most appropriate and effective content moderation approach and set them up for success.

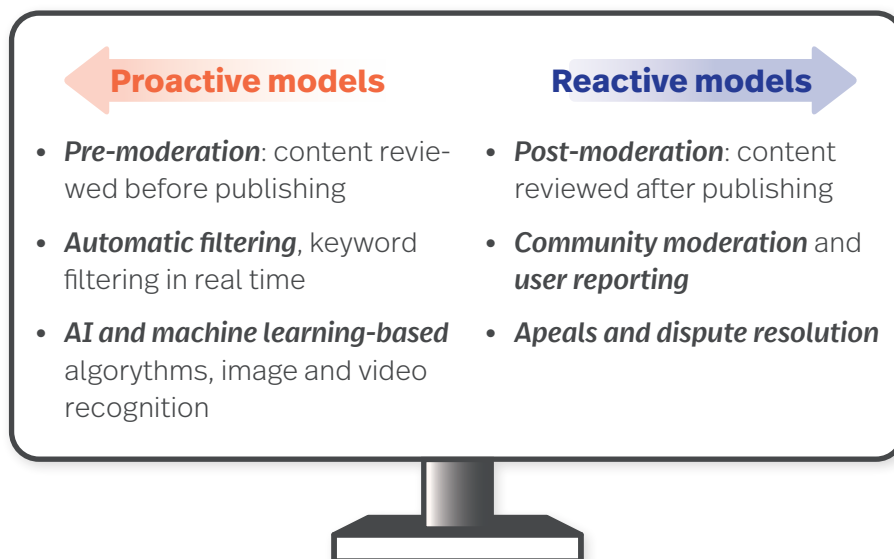
Why is content moderation important?

Content moderation **is a first line of defence** against terrorist content, as it involves preventing harmful content from spreading online and disabling it once it is shared publicly to reduce its reach.

- It entails **the review of user-generated content** to ensure it complies with community guidelines, terms of services, and legal regulations (such as the TCO Regulation).
- It can result in the **removal or restriction of content** flagged as harmful, malicious or abusive by a moderator.
- It can employ a **combination of automated algorithms and human oversight**.
- It requires a **careful balance** between preserving free speech and fostering a safe digital space.



The **EU TCO Regulation** (Art. 5) requires HSPs who have been exposed to terrorist content to put in place **specific proactive measures** to identify and prevent the dissemination of terrorist content. Content moderation is therefore an essential defensive and preventive mechanism.



Content moderation might require significant **technical and human resources** that smaller companies do not have. It is important to know that there are many different content moderation styles. While the TCO Regulation does not prescribe a specific approach, many **HSPs use a combination** of automated moderation alongside user reporting, trusted flagger programs, and industry collaboration efforts.

¹ Regulation (EU) 2021/784 of the European Parliament and of the Council of 29 April 2021 on addressing the dissemination of terrorist content online. <https://eur-lex.europa.eu/eli/reg/2021/784/oj>.

FRISCO Content Moderation Tool



Goal

Optimise and streamline content moderation workflows and processes for HSPs in compliance with the TCO Regulation. Intended for HSPs without in-house technical solutions for managing content moderation.



What is it?

It is a user-friendly trust and safety **content moderation tool** that addresses user-generated content related risks, including terrorist content.



How does it work?

The tool is based on **Tremau's in-house solutions** and tailored to HSPs' needs in relation to the TCO, thanks to the project's findings and resources. This tool is only usable through an **integration with the HSP's system** and provided through a SaaS (Software as a Service) contract with Tremau.

This tool provides a content moderation platform in which incoming flags from different sources of reporting of violative content are aggregated. Moderators can subsequently view the reported cases and make related decisions (i.e, keep or remove content, signal user) based on platform policies. It further allows HSPs to:

- set up an appeal mechanism
- devise a policy configuration and enforcement strategy
- send to users a statement of reasons for removed content, via an automated workflow
- integrate requests/removal orders from law enforcement
- generate transparency reports



Link: <https://friscoproject.eu/content-moderation-tool/>

Contact FRISCO partner Tremau to receive information and test the tool.



FRISCO Brochure: Tools and approaches for small HSPs



The brochure is a collection of further technical tools and software solutions available to HSPs in the field of **detection and moderation of terrorist content**. It aims to familiarise them with the wide range of industry and open-source software applications available to address terrorist content and facilitate full compliance with the TCO Regulation, including:

- Content moderation software
- Tools for detection of harmful/illegal content
- Hash-matching tools



Access the brochure here: <https://friscoproject.eu/publications/>



Partners:

CIVILPOL



IVSZ



TREMAU

Violence Prevention Network



This project has received funding from the European Union's Internal Security Fund (ISF) programme under Grant Agreement No 101080100. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.