



Jean De Meyere / Dr. Luc Desautettes

Undesirable content: the case of disinformation

Introduction

**– The difficult regulation
of disinformation**

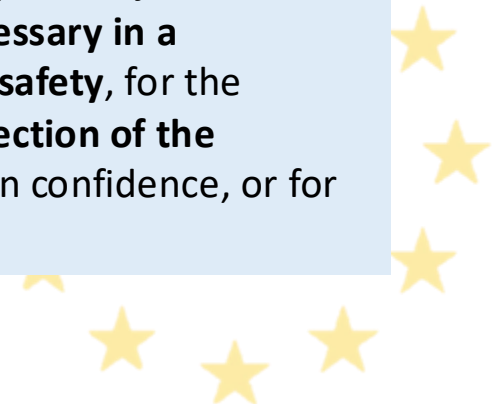


Challenges when regulating disinformation

- Where should the line be drawn ?
- Protection of the **freedom of expression**

Article 10 ECHR – Freedom of expression

1. Everyone has the right to **freedom of expression**. This right shall include freedom to **hold opinions** and to **receive and impart information and ideas without interference by public authority** and regardless of frontiers. (...)
2. The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, **restrictions** or penalties as are prescribed by law and are **necessary in a democratic society**, in the interests of national security, territorial integrity or **public safety**, for the **prevention of disorder** or crime, **for the protection of health or morals**, **for the protection of the reputation or rights of others**, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary.



Regulatory measures against disinformation pushed by social media platforms – T&C as a basis for moderating content

- Online platforms rely on their T&C to moderate content that does not respect their policies
- Exemple: [Google/Youtube ToS regarding misinformation](#)

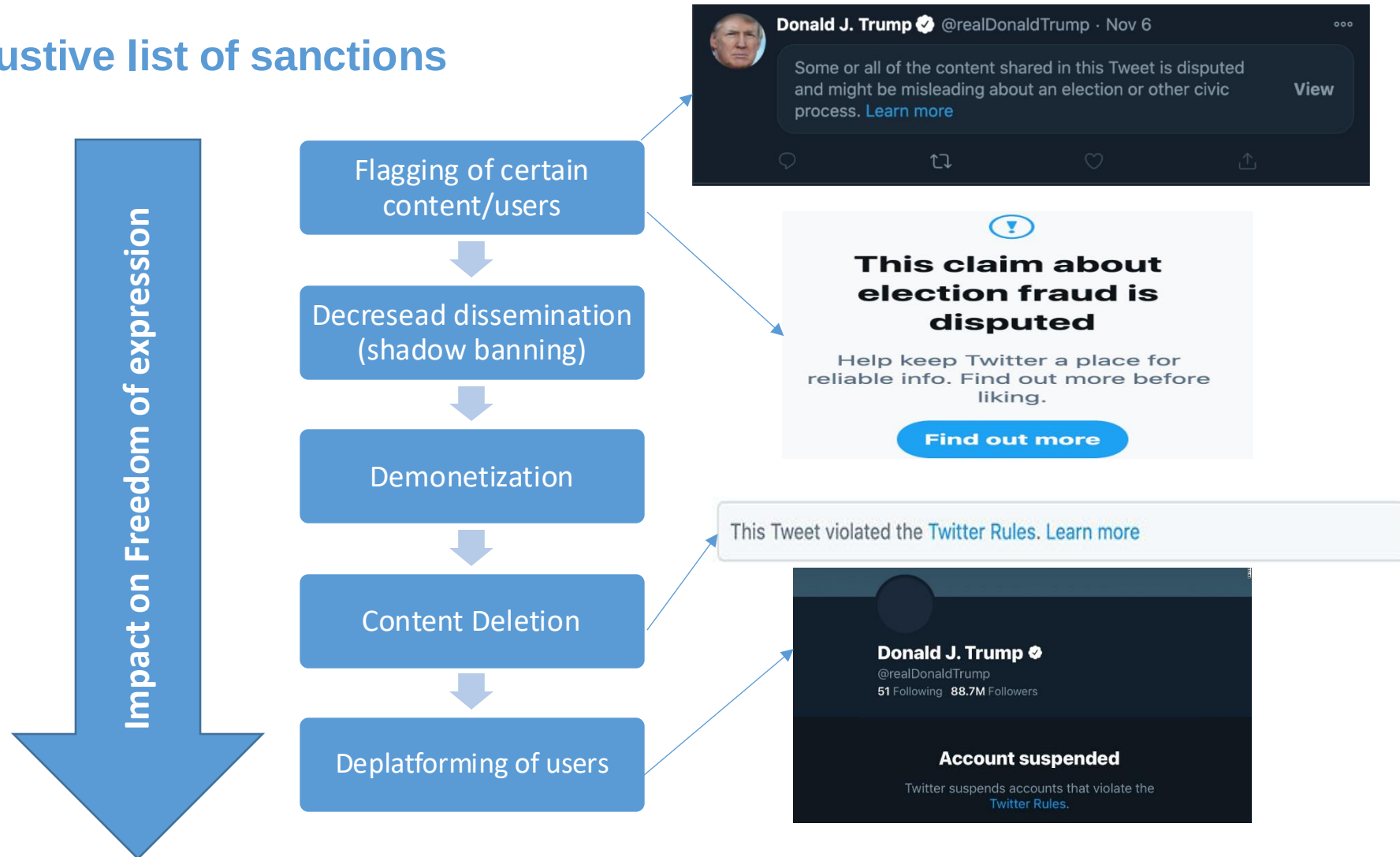
Misinformation policies

Certain types of misleading or deceptive content with serious risk of egregious harm are not allowed on YouTube. This includes certain types of misinformation that can cause real-world harm, like promoting harmful remedies or treatments, certain types of technically manipulated content, or content interfering with democratic processes.

If you find content that violates this policy, report it. Instructions for reporting violations of our Community Guidelines [are available here](#). If you've found multiple videos or comments that you would like to report, you can [report the channel](#).

- Different sanctions can be applied to derogatory content or to the users sharing them
- Critics :
 - Platforms decide themselves and can change their policies at any time
 - Lack of transparency and claims from users
 - Contractualization of freedom of expression

Non-exhaustive list of sanctions



1. Content curation

1.1. Absence of a 'direct' obligation concerning disinformation

Notice and actions mechanisms

Article 16 - Notice and action mechanisms

1. Providers of hosting services shall put mechanisms in place to allow any individual or entity to notify them of the presence on their service of specific items of information that the individual or **entity considers to be illegal content**. (...)
3. Notices referred to in this Article shall be considered to give rise **to actual knowledge or awareness** for the purposes of Article 6 (...).

Article 22 Trusted flaggers

1. Providers of online platforms shall take the necessary technical and organisational measures to ensure that notices submitted by trusted flaggers, acting within their designated area of expertise, through the mechanisms referred to in Article 16, are given priority and are processed and decided upon without undue delay.

-> Only apply to **illegal content**

Illegal content

Art. 3 Definition

(h) ‘illegal content’ means any **information** that, in itself or in relation to an activity, including the sale of products or the provision of services, **is not in compliance with Union law or the law of any Member State which is in compliance with Union law**, irrespective of the precise subject matter or nature of that law;

- Currently no further EU law: misinformation **is awful but lawful**
- Apparently no possibility for further national law:

Recital 9

This Regulation fully harmonises the rules (...), addressing (...) the societal risks that the dissemination of disinformation or other content may generate (...). Accordingly, Member States should not adopt or maintain additional national requirements relating to the matters falling within the scope of this Regulation, unless explicitly provided for in this Regulation (...).

This should not preclude the possibility of applying other national legislation applicable to providers of intermediary services, (...) where the provisions of national law pursue other legitimate public interest objectives than those pursued by this Regulation.

1. Content curation

1.2. Exception – Crisis Response Mechanisms

Crisis Response Mechanisms

Article 36 - Crisis response mechanism

1. **Where a crisis occurs, the Commission**, acting upon a recommendation of the Board **may adopt a decision, requiring** one or more providers (...) to take one or more of the **following actions**: (...)
 - (b) identify and apply specific, effective and proportionate **measures**, such as any of those provided for in **Article 35(1)** or Article 48(2), to prevent, eliminate or limit any such contribution to the serious threat identified pursuant to point (a) of this paragraph; (...)
2. For the purpose of this Article, **a crisis** shall be deemed to have occurred **where extraordinary circumstances** lead to a **serious threat to public security or public health** in the Union or in significant parts of it.

Article 35 – Mitigation risks

1. c) **adapting content moderation processes**, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision- making processes and dedicated resources for content moderation;
- See also recital 108.

Crisis Response Mechanisms

Article 36 - Crisis response mechanism

1. (...)

When identifying and applying measures pursuant to point (b) of this paragraph, **the service provider** or providers **shall take due account of** the gravity of the serious threat referred to in paragraph 2, of the urgency of the measures and of the actual or potential implications for the rights and legitimate interests of all parties concerned, including **the possible failure of the measures to respect the fundamental rights enshrined in the Charter.**

1. Content curation

1.3. The VLOP responsibility

Risk Assessment

Article 34 Risk assessment

1. Providers of very large online platforms and of very large online search engines shall diligently identify, analyse and assess **any systemic risks in the Union stemming from the design or functioning of their service** and its related systems (...)

This risk assessment (...) shall include the following systemic risks: (...)

(b) any actual or foreseeable negative effects for the **exercise of fundamental rights**, in particular the fundamental rights to human dignity enshrined in Article 1 of the Charter, (...), to freedom of expression and information, including the freedom and pluralism of the media, enshrined in Article 11 of the Charter (...).

(c) any actual or foreseeable negative effects **on civic discourse and electoral processes**, and public security;

(d) any actual or foreseeable negative effects in relation to gender-based violence, the protection of public health and minors and serious negative consequences to the person's physical and mental well-being.

Article 34 Risk assessment

2. When conducting risk assessments, providers of very large online platforms and of very large online search engines shall take into account, in particular, **whether and how the following factors influence any of the systemic risks** referred to in paragraph 1:

- (a) The design of their recommender systems and any other relevant algorithmic system;
- (b) Their content moderation systems;
- (c) The applicable terms and conditions and their enforcement;

The assessments shall also analyse whether and how the risks pursuant to paragraph 1 are influenced by **intentional manipulation of their service**, including by inauthentic use or automated exploitation of the service, **as well as the amplification and potentially rapid and wide dissemination** of illegal content and **of information that is incompatible with their terms and conditions**.

Mitigation of risks

Article 35 Mitigation of risks

1. Providers of very large online platforms and of very large online search engines **shall put in place reasonable, proportionate and effective mitigation measures**, tailored to the specific systemic risks identified pursuant to Article 34, **with particular consideration to the impacts of such measures on fundamental rights**.

Such measures may include, where applicable:

- (a) adapting the design, features or functioning of their services, including their online interfaces;
- (b) adapting their terms and conditions** and their enforcement;
- (c) adapting content moderation processes**, including (...), where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision- making processes and dedicated resources for content moderation;
- (d) testing and adapting their algorithmic systems, including their recommender systems;
- (...)
- (k) ensuring that an item of information, whether it constitutes a **generated or manipulated image, audio or video (...)** is **distinguishable through prominent markings** when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.

Code of conduct

Article 45 Codes of conduct

1. The Commission and the Board shall **encourage and facilitate the drawing up of voluntary codes of conduct** at Union level to contribute to the proper application of this Regulation, taking into account in particular the specific challenges of **tackling different types of** illegal content and **systemic risks**, in accordance with Union law in particular on competition and the protection of personal data.
2. **Where significant systemic risk** within the meaning of Article 34(1) **emerge and concern several very large online platforms** or very large online search engines, **the Commission may invite the providers** of very large online platforms concerned or the providers of very large online search engines concerned, and other providers of very large online platforms, of very large online search engines, of online platforms and of other intermediary services, as appropriate, as **well as relevant competent authorities, civil society organisations and other relevant stakeholders, to participate in the drawing up of codes of conduct**, including by **setting out commitments to take specific risk mitigation measures**, as well as a regular reporting framework on any measures taken and their outcomes.

2. transparency obligations toward platforms' monitoring

Reporting obligations – Intermediary Services

Article 15 – Intermediary Services

- Notice mechanisms (art. 16)
- Moderation practices
- Internal complaint-handling systems (art. 20)
- Use of AI for content moderation

Article 24 – Online platforms

- Out-of-court settlements bodies
- Decisions and statements of reasons (art. 17)

Article 42 – VLOPs

- Information on moderation (human resources, training, etc.)
- Risk-assessments (art. 34)
- Mitigation measures (art. 35)
- Audit report (art. 37(4))
- Audit implementation (art. 37(6))

37(6);
(...)

Terms and Conditions

Article 14 – Terms and Conditions

1. Providers of intermediary services shall **include information on any restrictions that they impose in relation to the use of their service** in respect of information provided by the recipients of the service, in their terms and conditions. That information shall include information on any policies, procedures, measures and tools used for the purpose of **content moderation**, including algorithmic decision-making and human review, as well as the rules of procedure of their internal complaint handling system. It shall be set out in clear, plain, intelligible, user-friendly and unambiguous language, and shall be publicly available in an easily accessible and machine-readable format.
2. Providers of intermediary services shall inform the recipients of the service of any significant change to the terms and conditions.
- (...)
4. Providers of intermediary services shall act in a **diligent, objective and proportionate manner in applying and enforcing the restrictions** referred to in paragraph 1, with due regard to the rights and legitimate interests of all parties involved, **including the fundamental rights of the recipients of the service**, such as the freedom of expression, freedom and pluralism of the media, and other fundamental rights and freedoms as enshrined in the Charter.

Statement of reasons

Article 17 - Statement of reasons

1. Providers of hosting services shall provide a clear and specific statement of reasons to any affected recipients of the service for any of the following restrictions imposed on the ground that the information provided by the recipient of the service is illegal content or **incompatible with their terms and conditions**:
(...)

3. The statement of reasons referred to in paragraph 1 shall at least contain the following information:
(a) information on whether the decision entails either the removal of, the disabling of access to, the demotion of or the restriction of the visibility of the information, or the suspension or termination of monetary payments related to that information, or imposes other measures referred to in paragraph 1 with regard to the information, and, where relevant, the territorial scope of the decision and its duration;
(...)

(e) where the decision is based on the alleged **incompatibility of the information with the terms and conditions** of the provider of hosting services, **a reference to the contractual ground** relied on and explanations as to why the information is considered to be incompatible with that ground;
(...)

-> Further transparency for users when action is taken on their content

4. Thank you